The Functional Neuroanatomy of Action Selection in Schizophrenia

Liana Romaniuk

Doctor of Philosophy
University of Edinburgh
2011

Abstract

Schizophrenia remains an enigmatic disorder with unclear neuropathology. Recent advances in neuroimaging and genetic research suggest alterations in glutamate-dopamine interactions adversely affecting synaptic plasticity both intracortically and subcortically. Relating these changes to the manifestation of symptoms presents a great challenge, requiring a constrained framework to capture the most salient elements. Here, a biologically-grounded computational model of basal ganglia-mediated action selection was used to explore two pathological processes that hypothetically underpin schizophrenia. These were a drop in the efficiency of cortical transmission, reducing both the signal-to-noise ratio (SNR) and overall activity levels; and an excessive compensatory upregulation of subcortical dopamine release. It was proposed that reduced cortical efficiency was the primary process, which led to a secondary disinhibition of subcortical dopamine release within the striatum. This compensation was believed to partly recover lost function, but could then induce disorganised-type symptoms - summarised as selection ”Instability” - if it became too pronounced. This overcompensation was argued to be countered by antipsychotic medication. The model’s validity was tested during an fMRI (functional magnetic resonance imaging) study of 16 healthy volunteers, using a novel perceptual decision-making task, and was found to provide a good account for pallidal activation. Its account for striatum was developed and improved with a small number of principled model modifications: the inclusion of fast spiking interneurons within striatum, and their inhibition by the basal ganglia’s key regulatory nucleus, external globus pallidus. A key final addition was the explicit modelling of dopaminergic midbrain, which is dynamically regulated by both cortex and the basal ganglia. This enabled hypotheses concerning the effects of cortical inefficiency, compensatory dopamine release and medication to be directly tested. The new model was verified with a second set of 12 healthy controls. Its pathological predictions were compared to data from 12 patients with schizophrenia. Model simulations suggested that Instability went hand-in-hand with cortical inefficiency and secondary dopamine upregulation. Patients with high Instability scores showed a loss of SNR within decision-related cortex (consistent with cortical inefficiency); an exaggerated response to task demands within substantia nigra (consistent with dopaminergic upregulation); and had an improved fit to simulated data derived from increasingly cortically-inefficient models. Simulations representing the healthy state provided a good account for patients’ motor putamen, but only cortically-inefficient simulations representing the ill state provided a fit for ventral-anterior striatum. This fit improved as the simulated model became more medicated (increased D2 receptor blockade). The rel-

ative improvement of this account correlated with patients' medication dosage. In summary, by distilling the hypothetical neuropathology of schizophrenia into two simplified umbrella processes, and using a computational model to consider their effects within action selection, this work has successfully related patients' fMRI activation to particular symptomatology and antipsychotic medication. This approach has the potential to improve patient care by enabling a neurobiological appreciation of their current illness state, and tailoring their medication level appropriately.

Acknowledgements

My gratitude goes out to David Willshaw and Stephen Lawrie for their inspirational discussions, incisive comments and patient supervision. I'm indebted to David McGonigle, Andrew Gillies, and Dominic Job for their excellent guidance and tuition on experimental design, fMRI analysis and computational modelling. The model created by Kevin Gurney, Tony Prescott and Peter Redgrave was an invaluable cornerstone of this work. Without Jessika Sussman, this thesis would likely never have happened, and together with Janet Skinner helped me integrate writing a thesis with taking a medical degree. I benefited greatly from the vast clinical experience and advice of Eve Johnstone and David Cunningham-Owens. I'd like to thank Adam Archibald, and Elaine Sandeman, Iona Hamilton, Jenny Boyd-Ellison, Marion Strachan and Angus McLaughlin at the SBIRC for their professionalism and compassion in helping people to be scanned. Stephen Giles made the challenge of managing and securing the subsequent data seem effortless. The wit and wisdom of Prerona Mukherjee, Ruth Philip and Heidi Bonnici made the whole experience a lot of fun. I'd like to thank Lydia Roman and my whole beleagured family for their firm support.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Liana Romaniuk)

Contents

List of Tables

List of Figures

Nomenclature

AAL	Anatomical automatic labelling
ACC	Anterior cingulate cortex
BG	Basal ganglia
BOLD	Blood oxygenation level dependent
Competitiveness	The direct suppression that one action exerts on another
CT	Computerised tomography
EPI	Echo planar imaging
FDR	False discovery rate
fMRI	Functional magnetic resonance imaging
FOV	Field of view
FSIs	Fast-spiking interneurons
FWE	Familywise error
FWE	Familywise error
FWHM	Full Width Half Max
GA	Genetic algorithm
GLM	General linear model
GPe	External globus pallidus

GPi	Internal globus pallidus
HRF	Haemodynamic response function
IFG	Inferior frontal gyrus
ISI	Interstimulus interval
MNI	Montreal Neurological Institute
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
MSNs	Medium spiny neurons
PANSS	Positive and Negative Syndrome Scale
PCP	Phencyclidine
PET	Positron emission tomography
REH	Royal Edinburgh Hospital
ROI	Region of interest
RT	Reaction time
Selectivity	The strictness with which competing actions are suppressed
SNc	Substantia nigra pars compacta
SNR	Signal-to-noise ratio
SNr	Substantia nigra pars reticulata
SPM	Statistical Parametric Mapping
STN	Subthalamic nucleus
TE	Echo time
TR	Scan repitition time
UPDRS	Unified Parkinsons disease rating scale
VOI	Volume of interest
VTA	Ventral tegmental area
WTA	Winner take all

Chapter 1

Introduction

Schizophrenia is a neuropsychiatric disorder whose impact is demonstrated at many levels. Underlying the human experience of this devastating illness are changes in gene expression, receptor efficacy, synaptic plasticity, small-scale network regulation and large-scale network interactions, described in section 2.1. Understandably, comprehensive interpretations of these data have been less forthcoming. Properly conceptualising the themes - and most importantly the commonalities - that connect these sometimes disparate changes is likely to be key to understanding the illness. This work pays particular attention to the negative and disorganised components of the disorder, both of which are arguably less well-researched than positive symptoms such as delusions.

Large-scale computational models offer a viable means of capturing the most salient points of any particular theory. Through iterative development by comparison with empirical data, they may be able to describe the neuropathological processes underlying the illness in a way that balances simplification with relevance. The body of work to date in modelling the neurobiology of schizophrenia are described in section 2.2.

Evidence for the specific neuropathology of schizophrenia remains tenuous, though some key discoveries have been made. Perhaps the strongest amongst these gather around two core themes: first, that there is a dysregulation of the mesolimbic dopamine system, as evidenced by the efficacy of dopamine-blocking antipsychotic medication; and second, that NMDA receptors, and their associated GABAergic regulatory pathways, are dysfunctional within prefrontal cortex. Glutamate-dopamine interactions therefore represent a fertile area for hypothesis testing regarding schizophrenia. One of the more prominent sites for glutamate-dopamine interactions exists at the terminus of frontal cortex's large-scale projection to the striatum. Described

in section 2.3, the striatum is substantially innervated by dopamine neurons, and forms part of a complex of subcortical nuclei known as the basal ganglia (BG). The gross architecture of these nuclei has invited significant speculation as to their computational role within the brain as a whole, with a number of theories converging on the idea of filtering or action selection. Deciding which action to take in a given situation is a problem requiring multiple sources of information to solve [16]. In the absence of a simple reflexive relationship between current behavioral context and a course of action, available options must be evaluated to allow the execution of any single behaviour. This process will engage a distributed cortical network (for recent reviews, see [17, 18, 19]), and computational approaches [20, 21, 22] suggest that centralized systems such as the BG provide an efficient solution to the 'action selection' problem [23], by suppressing competing alternatives [24, 25]. Acting as a key way-station within corticothalamic loops [26], the BG exert a suppressive influence via partially-segregated "channels" [27]. Given sufficient cortical input, the BG's influence is theorised to become focally disinhibitory, removing suppression and allowing the execution of a single course of action.

In this work, the hypothetical consequences of glutamate-dopamine dysfunction were explored within the context of frontostriatal projections and the BG. This was isolated to the specific subfunction of action selection. In order to generate specific predictions regarding BG responses to varying selection demands, the computational model of Gurney and colleagues was used [13, 28], which is described in section 2.4. This specifically quantifies the properties of Selectivity (the strictness with which competing actions are suppressed) and Competitiveness (the direct suppression that one action exerts on another). The model's gating threshold is modulated both dynamically via globus pallidus externum (GPe) and subthalamic nucleus (STN) interactions, and as a consequence of tonic dopamine levels. This architecture provided a number of predictions of BG activity, which were tested using a new perceptual decision-making task, the development of which is described in section 3.2. This built upon previous studies of response inhibition [29, 30, 31, 32], but rather than relying on particular responses being withheld, a two-alternative forced choice paradigm was used to better capture the complexity of selecting between multiple actions. The factor of Competition was varied by using blends of color to introduce ambiguity into the visual cues, which were individually calibrated to the colour perceptual responses of each participant, using software described in section 3.3. Cued actions could be either mechanically incompatible (e.g. flexion or extension around a single joint) or free from direct interference (e.g. flexion of the index and middle fingers), allowing us to vary the factor of Selectivity. The monitoring of these movements required the use of a fibre-optic data glove, described in section 3.4. To explore how cortex interacted with the BG, a means

of assessing effective connectivity was developed, based on Granger causality, described in section 3.5.

The model's ability to account for BG activity in healthy controls was assessed during a first fMRI study, described in chapter 4. This found that activity within some nuclei were well characterised, but in others the model had several shortcomings.

To address these, the model was modified in a biologically-grounded manner to enable it to provide a better fit for the data, as described in chapter 5. This involved incorporating fast-spiking interneurons (FSIs) within the striatum, which were subject to inhibition from GPe. These FSIs acted to inhibit striatal medium spiny neurons (MSNs), thus facilitating a degree of dynamic suppression within the BG's main input nucleus. However, as the objective was to more closely examine how dopamine modulated the selection process, and to instantiate the theory that downregulated cortical activation lead to a disinhibition of subcortical dopamine release, a key addition was made to the model: the explicit representation of dopaminergic midbrain. This also allowed for the exploration of the crucial interacting factor of antipsychotic medication. By modulating "cortical efficiency", and D2 receptor efficacy, the consequences of various illness states were described, providing predictions of how the activation seen in patients would differ from that of healthy controls.

These predictions were then tested during a second fMRI study involving both patients with schizophrenia, and a second group of healthy controls, described in chapter 6. The modified model was first verified against the new control data, and an improved account for striatal activation in particular was found. The specific hypotheses concerning the theoretical underlying neurobiology of schizophrenia, its consequences for BG activation, and the relationships these would have with selection dysfunction were challenged. In order to relate selection dysfunction to the symptoms of schizophrenia, a new metric named "Instability" was derived from the Positive and Negative Syndrome Scale (PANSS), which contrasted elements of negativity and disorganisation (section 6.1.4.4), and was used as a covariate in the major fMRI analyses. Two analytic approaches were adopted: the first regressed the task's 2 x 3 factorial design according to medication dosage and Instability, to see how well the model's qualitative predictions matched the data. A second approach simulated various illness states by varying the modified model's parameters describing cortical efficiency and D2 receptor efficacy, producing predicted BG activation timecourses for each participant's fMRI run. The degree of fit each of these provided for the data was compared, and the modulatory effects of Instability and medication were addressed. This produced a number of findings: models predicting the activity of healthy striatum provided a good account for patients' motor putamen, whereas only

”Ill and medicated” simulations could account for anterior striatum, with the fit improving with increasing patient medication dosage. A triangulation of results was found concerning cortical efficiency, SNR, dopaminergic dysregulation and Instability: the model predicted that cortical inefficiency would lead to dopaminergic dysregulation, and higher Instability. Unstable patients were found to have diminished cortical responses suggestive of reduced SNR and efficiency, and exaggerated activity within dopaminergic midbrain.

This work focuses on the factors influencing and subsequent consequences of tonic dopamine levels, rather than phasic release, which is believed to have more of a bearing on associative learning processes. It therefore makes no direct attempt to comment on the role of phasic dopamine.

Chapter 2

Background and literature review

2.1 Schizophrenia

Schizophrenia is a severe psychotic illness with imposes dramatic changes with long-term ramifications onto a person's life. Its onset is sometimes slow and insidious, sometimes sudden and disorienting, with most patients experiencing erroneous and frequently unpleasant perceptions (hallucinations), while many develop entrenched false beliefs sometime apparently in an attempt to accommodate their increasingly distorted world (delusions). Schizophrenia is usually accompanied by a general decline in cognitive ability. Presentation to a doctor tends to occur only after this vicious cycle has reached terminal velocity, and a florid first episode is in progress. Patients then face the new reality that their original plans for life will have to be significantly altered as certain opportunities are no longer available or attainable, given the primary effects of the illness itself, and the social stigma that diagnosis still entails. The most effective treatment comes in the form of anti-psychotic medication, which is generally effective in damping some but not all of the symptoms [33]. This gap reflects an incomplete knowledge of the neurobiological underpinnings of the disease, which is essential for the development of new, more complete treatment strategies.

This chapter will take a look at how the concept of schizophrenia has developed over time, and how new evidence has driven researchers to clarify and refocus our understanding of this most complex of illnesses.

2.1.1 Development of the concept

The clinical concept of schizophrenia has been adapted and refined over at least two centuries, with particular progress being made in the last one hundred years [34, 35, 36, 37]. To date, all classification systems have been based on those symptoms that patients report to skilled and experienced clinicians. These are alterations of perception, poverty of thought, feeling and expression, and disturbances of cognitive processes: in sum, they are not objectively biological. The pathophysiological processes underlying schizophrenia have not yet been fully characterised, meaning it lacks firm definitional boundaries, especially when comparing individual patients, or seeking a differential diagnosis against bipolar disorder or psychotic depression. Although some have taken this ambiguity to mean that schizophrenia is simply a pragmatic clinical catch-all which should be abandoned [38], most clinicians and researchers alike persevere with this useful if preliminary concept.

2.1.1.1 Origins

As early as 1400BC, the physicians of India and Rome have been documenting symptoms which we might recognise today as schizophrenia [39]. Although some classifications were perhaps less than helpful - for example, Willis' "foolishness" [40], the 17th century gave us syndromes approximating schizophrenia, for example Sauvage's "meloncholia enthusiastica" and "demonomania fantastica" [41]. At the turn of the 19th century, Pinel used the term "demence" or "loss of mind" for what would be the first formal description of schizophrenia as a phenomenon in its own right [42], emphasising the patients' decline in cognitive ability. In 1852, fellow Frenchman Morel noted both the rapid decline and early onset of illness in his patients, adapting the term into "demence precoce" [43] (where precoce means "premature"). Although he attempted a more rigorous means of diagnosis by considering the course of the illness, his contribution was undermined by a fixation with "hereditary degeneracy", usually applied on moral grounds [37]. In 1888, Clouston took the better aspects of Morel's work forward in that he considered illness course and prognosis as paramount when developing his concept of "adolescent insanity" [43]. A similar syndrome was recognised by Hecker as hebephrenia ("youthful insanity") [44], and by Kahlbaum as "paraphrenia hebetica" [45]: Kahlbaum was the first to produce a systematic classification of mental illness syndromes, which was a concept Kraepelin greatly admired and later employed himself.

2.1.1.2 Kraepelin and the symptoms of schizophrenia

Like some before him, Kraepelin was aware of the dangers of a purely symptom-based classification method, and was determined to take a more scientific approach to syndromal classification, placing particular emphasis on a patient's course, prognosis and overall outcome [46, 47].

> "Unfortunately in the field of psychic disturbances there is not a single symptom which is pathognomic for any particular illness. On the other hand we can expect that the composition of the individual characteristics which form the total picture, and in particular the changes which develop in the course of the illness, will not be produced in exactly the same way by any of the other diseases." [48]

In 1917, he ultimately dichotomised psychosis into two main groups: manic-depressive insanity and dementia praecox [49], which he believed could be differentiated according to factors such as premorbid personality, early age of onset, and a generally deteriorating course [50]. Kraepelin's work formed the basis of the modern definition of schizophrenia, and it is around his framework that the symptoms of schizophrenia will be defined here.

- *Avolition, poverty of expression and flattened affect*: For Kraepelin, these formed one of two principal groups of symptoms that characterised the illness. The patient demonstrates a loss of motivation to act on their own initiative; their speech may become lacking in content and quantity, and they may report an absence or slowing of thought itself. Kraepelin elaborates:

  > "...we observe a weakening to those emotional activities which permanently form the mainsprings of volition...Mental activity and instinct for occupation become mute. The result of this highly morbid process is emotional dullness, failure of mental activities, loss of mastery over volition, of endeavour, and ability for independent action..." [48]

- *Incongruence*: This formed the second group of Kraepelin's principal symptoms, and represents a fracturing of the mental processes:

  > "The second group of disorders consists in the loss of the inner unity of activities of intellect, emotion and volition in themselves and among one another....The near connection between thinking and feeling, between deliberation and emotional activity on the one hand, and practical work on the other is more or less lost. Emotions

do not correspond to ideas. The patient laughs and weeps without recognisable cause, without any relation to their circumstances and their experiences, smile as they narrate a tale of their attempted suicide...” [48]

- Auditory hallucinations: Kraepelin placed less emphasis on disturbances of perception, but noted their presence nonetheless:

  ”Sometimes it is only a whispering, ’as if it concerned me’ as the patient says, a secret language, ’taunting the captive’; sometimes the voices are loud or suppressed, as from a ventriloquist, or the call of a telephone, ’childrens’ voices’; a patient heard ’gnats speak.’ Sometimes they shout as a chorus or all confusedly; a patient spoke of ’drumming in the ear’; another heard, ’729,000 girls.’ Sometimes the voices appear to have a metallic sound, they are ’resonant voices,’ ’organ voices,’ or as a tuning fork.” [48]

  In those cases where the voices express meaningful content, this is often of an unpleasant nature. The voices may issue orders (”the ... voice will amuse itself by driving the patient to utter despair in that they approve of his intentions, or order him to make a certain purchase and then berate him for doing so.” [51]); offer criticism (”While the patient is eating, he hears a voice saying, ’Each mouthful is stolen’. If he drops something, he hears, ’If only your foot had been chopped off”’ [51]); provide a running commentary on what the patient is doing (”While getting ready in the morning a patient hears, ’Now she is combing her hair. Now she is getting dressed,’ sometime in a nagging tone...” [51]); or echo the patient’s own thoughts are heard in ambient environmental sounds, leading the patient to believe that everyone can hear them. Also, there is a distinction between sounds that are experienced as originating from an external source, and those for which the patient is aware that they are incorrectly perceived internal thoughts, i.e. pseudohallucinations.

  ”At other times they do not appear to the patients as sense perceptions at all...There is an ’inner feeling in the soul,’ an ’inward voice in the thoughts’; ’it is thought inwardly in me’; yet ’sounded as if thought.”’ [48]

- *Delusions*: These are fixed and false beliefs which remain unmodifiable in the face of overwhelming evidence, and are contrary to those held by the majority of their peers. They are varied in their content, but common themes emerge:

  – Passivity: The patient no longer feels in control of their thoughts and actions, with a likely interpretation being that they are being controlled by an external agency.

”It came to me from the X-ray department. It was nothing to do with me, they wanted it so I picked up the bottle and poured it” [52]. This includes thought insertion, whereby thoughts appear suddenly and unexpectedly in a patient’s mind; and thought withdrawal, when they are suddenly silenced or ”removed” [39].

- Persecution: The patient believes external agencies are working to undermine or harm them. ”There is no kind of human corruption by which one has not sinned against me.” [51]
- Grandiosity: ”The patient is ’something better,’ born to a higher place, the ’glory of Israel,’ an inventor, a great singer, can do what he will.” [48]
- Reference: The patient experiences objects and events in their environment as having a special personal significance. ”Everybody is in the street because of him; each gesture of these people has some significance for him; newspaper advertisements refer to him; the storm was made especially for him; the edition of Goethe’s works in the hospital library is full of hints at him and has been falsified for his sake. A still fairly lucid naturalist rejoices that the small infusoria wave to him when he looks at them through the microscope.” [51]

- *Thought disorder*: A patient’s patterns of thought and speech may be very hard to follow, becoming tangential or entirely loosing sight of a particular sentence’s intended goal. The patient may experience a pressure of thought, flowing rapidly and with sudden changes. Bleuler provides an example of thought disorder from a piece a patient wrote, entitled ”The Golden Age of Horticulture”:

  ”In this historical famous city of the Califs, there is a museum of Assyrian monuments from Macedonia. There flourish plantain trees, bananas [...] and olives. Olive oil is the Arabian liquor-sauce which the Afghans, Moors and Moslems use in ostrich farming. The Indian plantain tree is the whiskey of the Parsees and Arabs. The Parsee possesses as much influence over his elephant as the Moor over his camel.” [51]

Kraepelin’s emphasis was not on the more noticeable symptoms of delusions and hallucinations, but on what would collectively come to be known as negative and disorganised symptoms. These shared a close relationship with the long-lasting cognitive deficits he observed in patients with schizophrenia, notably problems in orienting and maintaining both attention and goal-directed behaviour.

Unlike some of his peers, Kraepelin was particularly concerned with aetiology and the importance of hereditary factors [53], and believed the illness to be the result of a severe, most likely irreparable cortical lesion [48], though of course there was scant evidence for this at the time. In 1911, Bleuler also wished to address the underlying processes of the illness, though he took a more psychodynamic approach, being heavily influenced by Jung and Freud.

2.1.1.3 Bleuler

Although he was operating within a psychodynamic context, Eugen Bleuler imagined the illness to be the result of ”a loosening of mental links between mental contents”, which ironically bears a strong resemblance to today’s theories regarding altered associative learning, albeit without the neural foundation. This was the basis for his new term ”schizophrenia” [51], which he felt was a better description of the person and not just the illness, and represented his attempt to understand the illness’ basis. In this, Bleuler made an important if not ill-founded or at least premature attempt to get to the heart of the disease. He stated:

> ”It is important that we postulate the existence of primary psychic symptoms, even though we do not know exactly what they are, and that the symptomatology of dementia praecox so far described consists to a large extent of secondary symptoms, which are brought about by the reaction of the sick psyche to the complexes (i.e. the affects).

Within Bleuler’s thinking, the more prominent symptoms of hallucinations and delusions were secondary consequences of what he considered to be the core of the illness: a loosening of associations; affective flattening; autism, that is social withdrawal, in some cases a full retreat into the patient’s own world; and ambivalence, the expression of contradictory emotions. Bleuler therefore shares a similar emphasis with Kraepelin, who also prioritised these symptoms. However, Bleuler goes one step further in suggesting that one set of symptoms is driving the development of another set, albeit with little evidence.

The views of Bleuler enjoyed some success, especially in America, where Freudian ideas had landed on particularly fertile ground. An unfortunate consequence of this was an overall broadening of what could be considered to be schizophrenia. Bleuler defined the concept of ”latent schizophrenia” - a state for which a patient displayed some of the underlying core symptoms, without it approaching a condition of clinical concern. Kraepelin himself commented that this was being overstretched: ”But for the vast majority of such disorders which present with anxiety and a lack of self-confidence, such an interpretation must be rejected” [35].

2.1.1.4 Schneider

This lack of specificity, and therefore of clinical usefulness, concerned Kurt Schneider. Although he also assessed the illness both longitudinally and in cross-section [37], his work shifted the means of diagnosing the illness into the here-and-now, producing a checklist of so-called first rank symptoms [54, 52]. He commented:

> ”We shall use both [longitudinal and cross-section information]; as the course itself consists of a sequence and developments of clinical states, their description will always have to precede the former. From this follows unavoidably the task so important for clinical and social psychiatry, to search them for symptoms, which from experience will permit predictions regarding the future course and outcome. This will after all be the questions asked of us.” [54]

For Schneider, the symptoms that both Kraepelin and Bleuler considered to be of paramount importance were too ill-defined, and suffered from poor inter-rater reliability. He wished to identify a set of symptoms that were distinctive and easily-defined, which are summarised here by Johnstone [39]:

- Auditory hallucinations taking any one of three forms:
  - Voices repeating the subject’s thoughts out loud, or anticipating his thoughts
  - Two or more hallucinatory voices discussing the subject, or arguing about him, referring to him in the third person
  - Voices commenting on the subject’s thoughts or behaviour, often as a running commentary
- The sensation of alien thoughts being put into the subject’s mind by some external agency, or his own thoughts being taken away (thought insertion or withdrawal)
- The sensation that the subject’s thinking is no longer confined within this own mind, but is instead shared by, or accessible to, others (thought broadcasting)
- The sensation of feelings, impulses or acts being experienced or carried out under external control
- The experience of being a passive and reluctant recipient of bodily sensation imposed by some external agency

- Delusional perception - a delusion arising fully fledged on the basis of a genuine perception which others would regard as commonplace and unrelated

Schneider's focus on hearing voices is not shared by Kraepelin and Bleuler, and although they would agree on the importance of delusions of passivity, they would regret the loss of avolition, blunted affect and incongruity [34]. Although there is a question mark over the validity of Schneider's schizophrenia, this pragmatic approach of producing a clear, practical checklist proved to be a useful clinical tool. Schneider's legacy is clearly seen in the most popular diagnostic tools: the World Health Organisation's International Classification of Disease (ICD: http://www.who.int/classifications/icd/en/), and the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM, current version: IV [55]). However, the questions surrounding specificity remain [56, 36, 57]. Defending his direction, Schneider stated:

> "...it no longer makes sense to fight e.g. over such questions of whether involutional melancholia "belongs" to the manic-depressive illness or not, but only whether it would be more practical to count it in, or more expedient not to do so. [...] As regards the final biological "right or wrongness" of such operations or fusions of various psychotic pictures (that) will one day, so we hope, be decided by somatic research ... This unless one wishes to abandon clinical psychiatry altogether, the task of finding psychological groups will persist. This however is a classification by useful types: Clinical psychiatry is a pragmatic science." [54]

The real question then, as ever, is to identify the underlying pathophysiological processes underlying schizophrenia. Only on this basis can symptoms be meaningfully delineated.

2.1.2 Emerging evidence for the neuropathology of schizophrenia

As early as 1867, Griesinger believed schizophrenia (which he referred to as "dementia paralytica") had a physical, neural basis [58], in opposition to those favouring "psychical" causes. However, it was Alois Alzheimer, a co-worker of Kraepelin, who first observed glial changes in the postmortem brains of patients with psychosis [59, 60], using the revolutionary neurohistological staining technique he learned from Nissl. The man who succeeded Alzheimer as the director of the Munich Neuroanatomical laboratory - Walther Spielmeyer - noted in a glowing obituary that both Alzheimer and Nissl had won the "victory of anatomy in psychiatry" [61]. However, this did not represent a watershed moment towards a fuller understanding of

the illness' neuropathology, and further progress beyond this finding was not made for many decades [62]. During the first half of the 20th century, American psychiatry neglected the biological approach, and remained distracted by Freudian concepts and psychodynamics in general. Meanwhile in Europe, the one avenue of biological psychiatry that still held sway focused on schizophrenia's hereditary causes, which was tragically misinterpreted and abused to justify sterilisation and euthanasia programmes in several countries, notably Nazi Germany [63]. On a similarly note, leucotomy became an accepted means of treatment for patients, especially in America, often resulting in personality change, or death through haemorrhage or infection. As Norbert Wiener put it, "...prefrontal lobotomy ...has recently been having a certain vogue, probably not unconnected with the fact that it makes the custodial care of many patients easier. Let me remark in passing that killing them makes their custodial care still easier" [64]. Other than being kept comfortable in a psychiatric hospital, the only other interventions available were insulin coma [65] and electoconvulsive therapy [66].

2.1.2.1 A psychopharmacological breakthrough

Biological psychiatry received a massive boost when in 1951 a key discovery was made [67] that effectively put an end to lobotomy and insulin therapy, and provided clues as to the underlying neurochemistry of schizophrenia. Charpentier and Laborit found that chlorpromazine - originally developed as an antihistamine - had a tranquilizing effect on psychotic patients [68], which was later established to be a specific reduction of their symptoms, rather than a simple sedative effect [69].

2.1.2.2 A neuropathological breakthrough

Tentative evidence for gross neural structural changes did exist from early studies using pneumoencephalography [70], though interpretation was hindered as these studies usually did not include controls, due to the unpleasant side-effects of the technique. With time came new imaging techniques that were non-invasive and so much more endurable to participants: computerised tomography (CT) and later still, magnetic resonance imaging (MRI). In 1976, Johnstone and colleagues established that schizophrenia has an organic correlate in the brain using CT imaging [71]. An enlargement of the lateral ventricles was clearly demonstrated in a group of patients with chronic schizophrenia, which has since been replicated in subsequent CT [72, 73] and MRI [74] studies, together with the finding that the third ventricle is similarly affected. Subsequent studies have found that schizophrenia is also associated with decreased

cortical volume, especially in the temporal lobes, and medial structures such as the amygdala, hippocampus and parahippocampal gyrus [74]. It is unlikely that these findings can be entirely put down to confounding factors such as medication or institutionalisation, as they have been observed in first-episode patients [75, 76, 77, 78, 79]. These findings' relevance to the pathological processes of schizophrenia are supported by studies that control for a number of genetic and environmental factors. For example, people at high genetic risk of schizophrenia (i.e. with one or more affected first or second degree relatives) have brain volumes intermediate between those who have illness, and controls with no family history [80]; and in monozygotic twins discordant for schizophrenia, the unaffected twin has larger cortical volumes and smaller ventricles than their sibling [81, 82].

2.1.3 The dopamine hyperfunction hypothesis

The discovery that particular pharmacological compounds possessed antipsychotic properties was of tremendous benefit to patients, although the reasons why they were effective remained ambiguous. However, a number of clues did present themselves in due course.

2.1.3.1 Circumstantial evidence for dopamine's involvement

Although very useful in combating the positive symptoms of schizophrenia, medication did elicit some unfortunate side effects. Principal among these was the inducement of Parkinsonian-type tremor, akinesia and muscular rigidity. Indeed the expected efficacy of new compounds was generally estimated by gauging their propensity to induce such symptoms [83]. It was later found that Parkinson's disease resulted from the catastrophic loss of dopamine-producing cells in the midbrain, which suggested that antipsychotics worked by blocking excess dopamine in schizophrenia [83]. Indeed one of the pharmacological treatments developed for Parkinson's disease, L-DOPA, acts to increase the levels of available dopamine, and has been shown to both precipitate hallucinations in overmedicated Parkinson's patients, and exacerbate the symptoms of people with schizophrenia [83]. Dopamine's involvement was further indicated by the psychotomimetic effects of amphetamine, a stimulant that induces dopamine release [84, 85].

Animals studies have also provided some evidence for the suppressive role of antipsychotic medication on effective dopamine levels. They are associated with increased dopamine metabolite levels, via a mechanism believe to be mediated by dopamine receptor antagonism. Dopamine

suppression is also implicated in the side-effect profile of antipsychotic medication: for example some patients experience hyperprolactinaemia. This hormonal disturbance can lead to loss of libido, as well as menstrual disturbances and osteoporosis in women, and gynecomastia in men. Pituitary secretion of prolactin is normally suppressed by dopamine, indicating that medication disinhibits this process via dopamine receptor antagonism [83].

2.1.3.2 The specific role of striatal D2 dopamine receptors

It was then observed that antipsychotics blocked dopamine receptors in the brain [86]. Indeed, antipsychotic potency shows a near perfect correlation to D2 blockade by nanomolar concentrations of neuroleptics in striatal preparations in vitro [87]. Further, several studies have now shown elevated densities and sensitivity of D2 dopamine receptors in schizophrenia [88].

2.1.4 The glutamate hypofunction hypothesis

The dopamine hyperfunction hypothesis was driven by accidental findings concerning the pharmacology of compounds that proved efficacious in alleviating some of the symptoms of schizophrenia. However, while positive symptoms such as auditory hallucinations and delusions of reference show a relatively high response rate, other experiences such as blunted affect and cognitive dysfunction remain mostly unimproved. D2 receptor dysfunction alone is therefore unlikely to provide a sufficient explanation for the illness. Further clues came from different pharmacological avenues: as well as exacerbating the symptoms of patients [89], blockade of the glutamate receptor NMDA by drugs such as ketamine [90] and phencyclidine (PCP) [91, 92] can induce schizophrenia-like symptoms in healthy controls - most notably including its negative and cognitive aspects. This indication of glutamate's involvement in the disorder has received further support from neuropathological investigations.

2.1.4.1 Altered cortical cytoarchitecture

The initial evidence of neural abnormalities in people with schizophrenia sparked a resurgence in postmortem histological studies, providing access to the microscopic features of schizophrenia [93]. The fruits of this labour however have been somewhat equivocal, with many reported features being contradicted in subsequent studies. Falling into this category are altered lamination of entorhinal cortex ([94, 95, 96], but see also [97, 98]); the disarray of hippocampal

neurons ([99, 100, 101, 102, 103], but see also [104, 105, 106]); and altered hippocampal neuron density ([107, 102], but see also [99, 108, 105, 106]). However, some robust features have emerged: although many hippocampal findings have not proved reliable, it does appear to be the case that patients with schizophrenia have smaller hippocampal neurons [109, 106, 103]. This is also true of prefrontal cortex [110], which also shows evidence of increased neuronal density [111, 112, 113]. Having more tightly-packed neurons is indicative of a reduction of neuropil [114], the fundamental hardware underlying neural connectivity. Subsequent studies have indeed found strong evidence to support a general reduction in dendritic spine density, particularly in dorsolateral prefrontal cortex [115, 116, 117, 118, 119]. This is especially the case for cortical layer 3 [116, 120], which projects to other cortical regions, sending axonal collaterals to other pyramidal neurons in the same layer [121, 122]. As well as this implied deficit in cortico-cortical connectivity, corticothalamic connectivity is also called into question, as layer 3 receives a major input from mediodorsal thalamus [123].

2.1.4.2 Glutamatergic synaptic dysfunction

Decreased dendritic spine density is concordant with the observation that patients with schizophrenia have been found to have reduced levels of synaptic proteins in both the hippocampus and surrounding cortex [124, 125, 126, 127, 128, 129, 130], and dorsolateral prefrontal cortex [131, 132, 115, 133, 134]. In the hippocampus, it is those proteins specifically associated with presynaptic excitatory structures that are especially diminished [128], in concordance with decreased levels of glutamate receptor components [135, 136, 137, 138]. Further evidence of diminished glutamate efficacy in schizophrenia came in the form of decreased levels of glutamate in CSF [139]. However, the failure to replicate this pushed it to the wayside [140]. It was the discovery of the psychotomimetic effects of NMDA antagonists PCP [91, 92] and ketamine [90] that brought the idea back to the fore. As well as eliciting some of the positive symptoms of schizophrenia such as auditory hallucinations, it also brings on thought disorder, emotional blunting and disturbances of working memory [141], which are frequently the symptoms left unaffected or even exacerbated by the majority of dopamine-blocking antipsychotics. This provides an indication as to the importance of glutamate in the pathophysiology of schizophrenia. Although some postmortem evidence suggests that patients have altered glutamate uptake mechanisms [142, 143], and expressed levels of NR1, NR2A [144] and NR2D [145] NMDA receptor subunits in prefrontal cortex, the evidence is somewhat equivocal on this front [146, 147]. This has led some to propose that other mechanisms underlying NMDA receptor signalling must be altered in schizophrenia [148, 149].

In rats, it has been established that NMDA antagonists damage neurons at the cellular level, inducing vacuolation and mitochondrial abnormalities [150, 151] in areas where cytoarchitectonic alterations have been observed in schizophrenia: the cingulate gyrus, temporal lobes, hippocampus and amygdala [152, 153]. However, the neurotoxic effects of NMDA antagonists are mitigated by the co-administration of agents such as muscarinic antagonists or GABA agonists [154]. This led Olney and Farber to propose in their key 1995 paper that NMDA blockage has an overall disinhibitory effect on excitatory pyramidal neurons, by acting to suppress their multiple inhibitory efferent interneurons [155]. They believed this could act through multiple circuits, for example via the anterior thalamus or basal forebrain. Evidence for such large-scale dysfunction is emerging: using magnetic resonance spectroscopy (MRS), patients show increased levels of glutamate in prefrontal and hippocampal areas that covaries with diminished global mental performance, but not antipsychotic medication [156]. A recent study combining MRS with structural imaging in people displaying prodromal symptoms of schizophrenia found that symptomatic people had lower glutamate levels in the thalamus, which directly correlated with insula and medial temporal lobe grey matter volumes [157]. Others have found that levels of N-acetylaspartate (NAA), a putative measure of neural functionality, are reduced in neuroleptic-naive first-episode patients, the degree of which correlates with poor performance on the Auditory Verbal Learning Task.

2.1.4.3 The influence of GABA and dopamine within cortex

The case for this leading to disinhibition of pyramidal cortical cells via interneuron understimulation as argued by Olney and Farber has been strengthened by postmortem findings concerning GABA. Levels of glutamic acid decarboxylase 67 (GAD67), an enzyme key to the synthesis of GABA, is reduced in the interneurons of patients [158, 159], especially those lying within cortical layers 1-5 [158]. Moreover, this seems to be very specific to the subclass of parvalbumin-containing interneurons, that impose a significant inhibitory influence over pyramidal cells, by virtue of their initial axonal and proximal dendritic synapses [149]. It has recently been shown that NMDA antagonism directly leads to both decreased levels of GAD67, and subsequent disinhibition of pyramidal cells. Therefore it appears that NMDA antagonism may have the paradoxical effect of ”releasing the brakes” on primary excitatory cortical activity. The consequences of this are likely to be an upset in ‘slow reverberation’, which in turn is believed to alter working memory.

Furthermore, computational modeling studies of neuronal network activity in the dorsolateral prefrontal cortex suggest that an NMDA receptor activation supports slow reverberating ex-

citation [160, 3]. This pattern of excitation is critical for the sustained firing of dorsolateral prefrontal cortex neurons that is considered to be the cellular basis of working memory [161]. Thus, cortical function is dysregulated and noisy in schizophrenia, impacting on cortico-striatal interactions, and the regulation of dopaminergic neurotransmission [149].

2.1.4.4 Hypofrontality and desynchronisation

One of the most consistent findings in brain imaging studies of patients with schizophrenia is a relative under-activation of prefrontal cortex at rest and especially under task-activated conditions. However, several functional neuroimaging studies reveal that prefrontal cortex is not always hypoactive, but sometimes hyperactive in schizophrenia [162, 163] consistent with a dysfunctional gating process. Indeed in 2003, Manoach presented a compelling case that task demands relative to an individuals location on their load-performance curve will determine whether hypofrontality or hyperfrontality is evident [163], although empirical evidence for this is lacking. Another way of interpreting the same findings, that is in keeping with neurochemical theories and data, is that patients show more activity relative to baseline, but that the response to the task is flatter, i.e. lower SNR leading to less meaningful, differentiable “activation”.

Both chandelier and basket neurons target multiple pyramidal neurons [164], enabling them to synchronize the activity of local populations of pyramidal neurons [165]. Networks of PV-positive GABA neurons, formed by both chemical and electrical synapses, give rise to oscillatory activity in the gamma band range, the synchronized firing of a neuronal population at 30—80 Hz [166]. Interestingly, gamma band oscillations in the human dorsolateral prefrontal cortex increase in proportion to working memory load [167], and in subjects with schizophrenia, prefrontal gamma band oscillations are reduced bilaterally during a working memory task [168]. Thus, a deficit in the synchronization of pyramidal cell firing, resulting from impaired regulation of pyramidal cell networks by PV- positive GABA neurons, may contribute to reduced levels of induced gamma band oscillations, and consequently to impairments in cognitive tasks that involve working memory in subjects with schizophrenia [169]. This is a likely means by which glutamatergic abnormalities can be linked to one of the other key replicated findings in patients, that of disconnectivity between cortical regions [170].

2.1.4.5 Neurodevelopmental aspects

Neuropathological findings indicate that schizophrenia is not accompanied by gliosis and is therefore likely to be a neurodevelopmental disease, with the middle stage of gestation being

a crucial time [171, 93]. As noted by Olney and Farber [155], an animal's vulnerability NMDA antagonist-induced damage is limited to the developmental stages following early adolescence [172]. This important observation is concordant with the knowledge that children rarely show a psychotomimetic reaction to NMDA blockage, allowing these drugs to be routinely used as a paediatric sedatives [173, 174, 175]. Olney proposed that an early, probably prenatal insult would result in a latent flaw that would later undermine the crucial developmental processes of late adolescence, e.g. synaptic pruning. It is at this time when regulatory neutral circuits are being finalised, involving the balancing of inhibitory and excitatory sources.

2.1.5 Summary

Schizophrenia is often a devastating condition which remains defined syndromally and poorly understood. It is however clear that while it has a genetic basis, it is not the result of a "one-hit" event - it is rather a probabilistic outcome, combining developmental vulnerabilities and environmental stressors which impact upon normal neural maturational processes.

While the glutamatergic hypothesis of schizophrenia is in many ways the most parsimonious synthesis of the available literature, problems with dopamine and/or GABA would have a similar effect, or may themselves be a consequence of NMDA dysfunction [176]. There is certainly some replicated evidence for GABAergic disruptions in schizophrenia, such as reduced GABA uptake in hippocampus/amygdala [177] and in the cingulate [109, 178]. Similarly, decreased dopamine in cortex could reduce maintenance (tonic D1) and/or prevent update (phasic D2) resulting in noisy, inefficient, unstable neural representations. Durstewitz & Seamans [179], Lewis & Burgos-Gonzalez [149], Stephan et al. [170] and other major reviews demonstrate that the dopamine hypothesis and glutamate hypothesis of schizophrenia are in the process of being empirically unified and recent developments point towards a complex interaction of the dopaminergic and glutamatergic systems in schizophrenia [180, 181, 141, 182].

The evidence summarised so far supports the two notions that (a) glutamatergic activity within frontal cortex has become dysregulated, in part via GABA dysfunction, resulting in a reduced SNR; and (b) dopamine action is exaggerated within the striatum, mediated via D2 receptors. This breakdown finds an analogue in pharmacological studies: dopamine manipulation is more closely-related to florid symptoms such as hallucinations and delusions, whereas glutamate antagonism more reliably induces negative symptoms and deficits in working memory. However, the key question is how these two pathologies, which are likely to be mutually-exacerbating, interact with one another to produce the characteristic symptoms associated

with schizophrenia.

2.2 Computational models of schizophrenia

The biological evidence concerning schizophrenia comes from a variety of fields. Genetic factors are clearly implicated by schizophrenia's heritability, and have been born out in linkage and association studies. Alterations in the expression of key proteins will impact cellular function, most likely via key receptors and neurotransmitter transporters. This in turn will alter how proximal and distal neurons interact with one another, finally culminating in the pronounced changes in perceptual experience and behaviour that have come to characterise the disorder. Gaining a better understanding of the pathological processes behind schizophrenia demands a broader, multi-level view, but with a focused eye on useful detail. The demands of this kind of consideration can be met by computational modelling, which offers multiple frameworks and concepts that can be used to collapse related factors, and elaborate on those points that seem most pertinent.

Schizophrenia has been considered using computational models in a variety of circumstances, not just related to understanding its neural underpinnings. Artificial neural networks are especially well-adapted to isolating patterns in complex, multifactoral stimuli (this is elaborated on in 2.2.2), and Campana and colleagues [183] have exploited this to differentiate known patients with schizophrenia from healthy controls, based on eye-tracking data (eye-tracking anomalies are well-documented among people with schizophrenia and their relatives [184]). However, it is doubtful that an "automated diagnostic system" would be well-received by the clinicians who might use it. More useful perhaps is the ability to predict a patient's response to medication. Lan and colleagues trained a neural network to indicate patients who would experience clinically-significant weight gain when prescribed antipsychotic medication [185]. Using clinical interview data, selected genotype information and several lifestyle factors, they were able to predict the weight changes of a group patient, after training on a separate sample. Another forward modelling example is provided by Guo et al.: in this case, they depart from neural network methodology and use a Bayesian hierarchical model to predict a patient's response to medication, based on select clinical parameters and a functional neuroimaging brain scan [186]. Medication has also been approached from a bottom-up perspective, specifically by predicting the protein structure and subsequent molecular dynamics of dopamine receptor ligands [187, 188]. Protein-protein interactions, and how they are altered by risk-associated genotypic variation, have also been simulated [189]. Schizophrenia's impact on society has also been

simulated economic models concerning the cost-effectiveness of different medications [190] and early versus late intervention [191].

In this section, all explicit attempts to model the neurobiological aspects of schizophrenia will be reviewed, which range from more abstract cognitive representations, to simplified neurophysiological processes within Hopfield networks, to models that have loose biological substrates, but firm mathematical underpinnings [192]. These vary in their degree of faithfulness to probable biological causes, which is itself a strength and a weakness. Tying things down to specific biological processes allows for sound hypothesis generation, but the available data suggests multiple possible and occasionally contradictory directions such a model should take. Conversely, a broader, more abstract view may be closer to the principles of the illness, but not be tremendously informative, or remotely disprovable.

2.2.1 Core concepts

The computational literature concerning schizophrenia coalesces around four key themes:

- The stability with which cortical activity represents mental phenomena, such as memories or sensory stimuli
- The updating or ”gating” of those cortical representations
- The clarity with which those representations exist versus background activity, i.e. the SNR
- The salience assigned to any given representation

These concepts will be elaborated on in turn, and the models exploring them described.

2.2.2 Cortical stability

Neural network theory posits that collections of interconnected individual ”neurons”, arranged in feedforward/feedback layers, can represent complex inputs by tuning the strength of connections within and across the whole population. Such alterations are guided by feedback of some kind, of which there are several variations (see later). The modification of these weights is broadly analogous to adaptive synaptic strength - or learning - and the resultant networks possess a number of biologically-informative properties. Notably among these is the ability to detect implicit patterns in a series of stimuli, and then use this to correctly classify novel

stimuli. Due to the flexible and unstructured initial stages of the network, they're useful for approximating complex functions, which would otherwise present a considerable challenge to discern explicitly. This property underlies the practical implementation of neural networks in areas such as of visual object or auditory sequence (e.g. speech) recognition. They offer a simple means to capture some salient properties of biological neurons: the effects of excitatory and inhibitory inputs on a given neuron; how these are integrated and resolved into efferent action potentials; their associated rate of firing; synaptic plasticity and the influences that guide it; and how the gathered activity of a neural population can uniquely represent a particular mental phenomenon. They're usually considered to approximate cortical activity within restricted regions, though some models do implement a broader view of the brain, incorporating additional cortical and subcortical modules. These increasingly complex, multi-layer models can perform tasks such as comparing sequentially-presented stimuli, though this is dependent on an essential property: the ability to maintain an initial representation once its originating stimulus has been withdrawn. Indeed, this property of stability plays an important part during the presentation of ambiguous novel stimuli that do not immediately allow the model to settle on an appropriate output.

Being able to maintain a pattern of activation is strongly dependent on the strength of recurrent excitatory connections, and the inhibitory influence of neighbouring neurons within the same layer, as well as neurons from adjacent feedforward and feedback layers. Several groups have developed models that explore the pathophysiology of schizophrenia in these terms, balancing the efficacy of excitatory and inhibitory synapses. In general, their aims have been to relate dysfunction in the biological analogues of these synapses to unstable or "fragile" network activation. The more popular output behavioural analogues are deficits in working memory, which some have extended to include the maintenance of "context" - that is, the broader appreciation for a person's current environment, and their internal states. This "internal model" of the person's perceived world is theorised to be disturbed in schizophrenia. Some have explored the potential implications of this by theorising that this model is in fact the basis for generative models of a person's anticipated effects on their surroundings.

The first explicit attempt to model the pathophysiology of schizophrenia was described by Hoffman in 1987 [193], which has since been elaborated on [194, 195, 196, 2, 197, 198, 199, 1]. These models are based on the notion that schizophrenia results from excessive synaptic pruning during adolescence, based partly on the relatively-restricted age of illness onset [200]. Indeed postmortem evidence suggests that patients with schizophrenia have pronounced loss of dendritic spines, especially in prefrontal regions [116]. Siekmeier and

Hoffman explored the hypothetical effects that lost dendritic spines could have on auditory hallucinations, and more recently semantic priming [199], using an attractor neural network. By eroding connections between neurons in their simulation, they sought to determine if speech percepts could arise in the absence of any phonetic input. They argued that the inherent ambiguities involved in perceiving grammatical sentences from phonemes is partly resolved via the completion a neural network can provide, through a hierarchical input-intermediate-output architecture (figure 2.1). As relationships between words are learned and incorporated into the weights of this network, distinguishing phonemes is additionally guided by semantic context. The intermediate layer was taken to represent auditory working memory, its inputs I described for each neuron x by:

$$I(x) = \sum w_{yx} a_i(y)$$

where $a_i(y)$ is the level of activation of input neuron y, and w_{yx} the weight of each $y \rightarrow x$ connection, analogous to synaptic strength. Each intermediate neuron also received intralayer connections from its neighbours, lagging by one time point, enabling it to consider the previously presented word, i.e., perform a working memory function.

Using an online version of backpropagation learning [201], the network was trained on a limited and partially-degraded vocabulary, forcing the network to rely on its working memory component, with each word eventually producing a unique pattern of activation or attractor state amongst the output neurons. Afterwards, presented phonemes could then be identified as comprising a particular word according to a best fit algorithm - if no fit was found, the network was considered to be perceiving nothing. New sentences were presented to the network, with words being partially degraded to introduce ambiguity, and separated by brief pauses - if the output layer settled to a recognisable attractor within these pauses, the network was said to have ”hallucinated”. Prior to undergoing synaptic pruning, the model was able to identify words with a high degree of accuracy, though this was significantly diminished if order was random, i.e. then the network was unable to exploit the semantic dependencies it had learned. Siekmeier and Hoffman then subjected this network to synaptic pruning, governed by the principle of ”Neural Darwinism”: connections that were deemed too weak, i.e with weights below a certain threshold, were set to zero. This bears some resemblance to ideas surrounding NMDA receptor dysfunction and the glutamate hypofunction hypothesis of schizophrenia. To a point, this had the effect of increasing recognition accuracy; however, this peaked at a pruned level of 30%, after which performance decreased, while there was an increase in ”hallucinated” percepts during periods when inputs were absent (figure 2.2). They contrasted this to an alter-

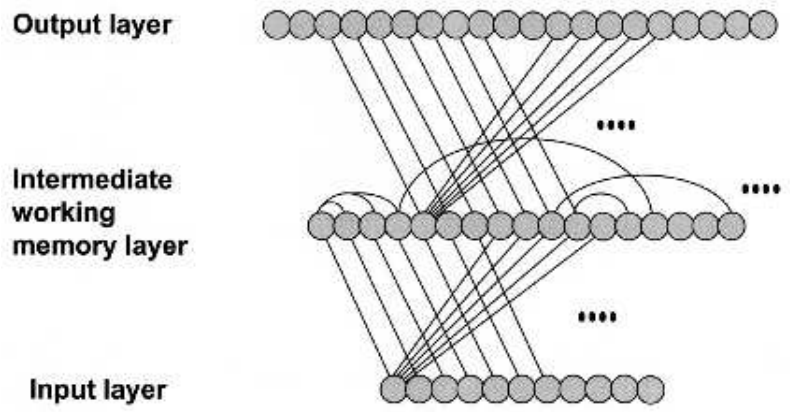


Figure 2.1: A simple schematic of the Hoffman artificial neural network, performing auditory speech perception. Inputs are combinations of phonemes representing words, presented sequentially from a complete sentence. Each input neuron is connected to every intermediate unit, which in turn project to each output unit. In addition, connections remained between intermediate neurons based on activation elicited by the previously presented word. Image taken from Hoffman & McGlashan 2006 [1].

native means of adapting the network over time - neuronal apoptosis though disuse - and found that this contributed to neither improved performance at intermediate levels, nor hallucinated percepts when taken to extremes. The network adopted specific abnormal behaviours - after a considerable degree of pruning, it would cease to act as a cohesive whole, with some subpopulations of neurons acting independently of the rest. This would seem to be a predictable consequence of connectivity degradation, and may indeed be a simple artifact produced by artificial neural networks of limited size: as it loses synapses, its basic capacity is reduced to the point of being insufficient to uniquely encode a given number of words. More interestingly, they observed the development of what was described as ”parasitic foci”: following presentation of an ambiguous word - and sometimes spontaneously - a section of the network would collapse into an attractor that did not represent a learned word, and this attractor could persist during subsequent word presentations. These are effectively overloaded networks, with the underlying architecture no longer having the capacity to encode the number of phenomena demanded of it. The authors proposed this could be analogous to the fixed ideology of delusions, or volitional and attributional disturbances, depending on the cortical area in which they arise.

Hoffman acknowledged the importance of dopamine, and briefly elaborated on its potential

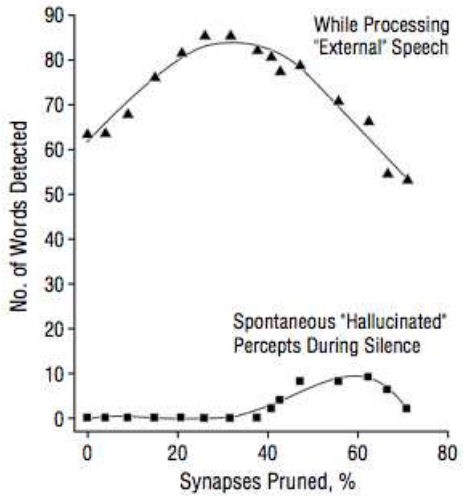


Figure 2.2: Hoffman model behaviour with increasing synaptic pruning. Up to a level of 30%, word recognition accuracy is improved. After this point, performance is degraded, while the incidence of erroneous ”hallucinated” percepts increased during periods of absent input. From McGlashan & Hoffman 2000 [2].

role: he used it to increase the activation threshold of each neuron, effectively damping working memory-related cortical activity. This simplified view of cortical hyperdopaminergia is somewhat at odds with other models of schizophrenia. It did produce similar hallucinations to synaptic pruning, but never performance gains at intermediate levels. They explored the possible interaction between dopamine dysregulation and synaptic pruning by determining which combinations provided a better fit of a sample of behavioural data he had from hallucinating and non-hallucinating patients, and controls. They concluded that synaptic pruning alone provided an insufficient account for the data: that it was either combined with cortical hypodopaminergia (a biologically-plausible possibility), or was entirely accounted for by cortical hyperdopaminergia in isolation (which is incompatible with our understanding of the pathophysiology of schizophrenia). On a more tenuous note, they argued that the pruned network’s tendency to misidentify genuinely-perceived words could underlie delusions of reference, in that patients are simply mishearing the radio, TV, etc. An alternative view would be that inappropriate fluctuations in phasic dopamine give certain innocuous stimuli a special prominence, which the patient then rationalises in a personal way.

The synaptic pruning model provides a plausible account for individual differences in each patient’s prepathological cognitive state, the nature and length of their prodromal phase, the age of illness onset, and the severity of the illness itself. People who begin with fewer synapses and lower brain volume will have ”less to lose” before pruning becomes pathological, leading

to earlier onset and severe illness. Alternatively. if the pace of pruning was abnormally high, one could expect a rapid decline and sudden severe onset. The fluctuations in the modulatory role of dopamine could underlie the episodic nature of the illness.

The models of Hoffman, Siekmeier and McGlashan demonstrate how a biologically-plausible and routine part of developmental plasticity can give rise instabilities in network behaviour, leading to psychotic symptoms when taken to extremes. Others have taken this generalised principle of connectivity erosion and applied it more abstractly to word association tasks, accounting for patient behaviour [202]. Though they highlight the important interaction with dopamine, there is no elaboration as to how it too would become dysregulated. Their emphasis is also very much on cortical abnormalities, whereas neurobiological evidence and medication efficacy demands that subcortical involvement must also be addressed.

An alternative attractor neural network has been developed by Horn and Ruppin, and later elaborated on by Greenstein-Messica [203, 204, 205, 206, 207, 208]. They argue that the essential neurobiological feature of schizophrenia is excessive synaptic plasticity, specifically between neurons within frontal cortex. They initially envisioned this as a compensatory reaction to diminished frontotemporal connectivity [203], but later proposed that it is due to delayed maturation of NMDA receptors, leaving patients with an unusually high proportion of an extra-plastic variant that normally predominates during childhood/early adolescence [208]. Their model contrasts to that of Hoffman and colleagues in that its non-deleterious nature should leave long-term memory retrieval intact - they believed that Hoffman's excessive synaptic pruning would bring about considerable memory deficits, which is not a characteristic feature of schizophrenia. Whereas synaptic plasticity certainly appears to be dysfunctional in schizophrenia [209], neurobiological evidence of synaptic proliferation is thin on the ground [74] (although gross anatomical frontal enlargement has been seen in people with schizophrenia before they become unwell [210]). In the latest version of their model, they explore the concept of synaptic runaway: this is a well-known problem within associative neural networks, whereby previously-learned patterns interfere with the encoding of new patterns, altering synaptic connectivity and so diminishing both learning and retrieval. They train their network with a series of input patterns, and compare two different learning rules: the first is purely dependent on the input pattern and so is not subject to interference from previous encoding; the second is Hebbian in nature (i.e. it alters synaptic plasticity on the basis of mutual pre- and post-synaptic activity), and so is partly dependent on the current activation level of the network itself. The capacity of these networks was established by pushing them to breaking point - they were trained to encode as many input patterns as possible for the network became unstable, and

encoded patterns began to become conflated. This clearly demonstrated synaptic runaway, with the activity-dependent Hebbian network becoming unstable sooner, and having a capacity that was 20% smaller than the activity-independent version. Moreover, they found that this feature was heavily dependent on the activation threshold at which synaptic plasticity could be altered - supporting their view that such dysfunction could be underwritten by unusually malleable immature NMDA receptors. Their method of gradually overloading the network also posed another possibility: the moment of collapse would be delayed, though the foundations for the dysfunction were long-standing - this could underlie the illness' onset during late adolescence, or sooner if the person experienced more notable or stressful events during early life. It is worth noting that their central idea of NMDA hypersensitivity could be considered at odds with the NMDA hypofunction hypothesis, though the implications of NMDA immaturity would vary greatly depending on whether they predominated on pyramidal or interneurons, either dendritically or somatically. The authors comment that synaptic runaway would be additionally augmented by dopamine dysfunction, especially by inappropriate phasic bursts, further potentiating already hypersensitive receptors.

Chen developed a similar model of that of Greenstein-Messica and Ruppin [208]- a binary attractor associative network that was trained to recognise a series of input patterns - and also explored the interaction between an underlying neurobiological anomaly and environmental factors i.e. memory load [211, 212]. His work preceded Greenstein-Messica and Ruppin by almost a decade, though it had a more abstract biological foundation. Chen argued that one of the hippocampus' primary functions is to orthogonalise sensory input, for subsequent processing within associative cortex. Were this function to be compromised, associative cortex would be forced to encode patterns that were unusually correlated. Chen's model demonstrated that associative networks encoding correlated inputs became unstable with significantly lower memory loads than a network exposed to orthogonalised inputs. This bears a strong resemblance to Greenstein-Messica and Ruppin's synaptic runaway - whether the inputs were inherently correlated due to a faulty hippocampus, or effectively correlated due to NMDA-associated excessive involvement of neighbouring neurons, the result was an activity-dependent drive of aberrant synaptic plasticity.

Friston has formalised these principles of erroneous correlation and altered synaptic connectivity using chaos theory and nonlinear dynamics [213]. He drew from evidence that the neuronal activity of patients with schizophrenia has an unusually high dimensional complexity [214], that is, their neuronal dynamics behave as if composed of a larger number of separate fragments. Friston demonstrated this concept by eroding a nonlinear neural network com-

prised of 3 sparsely connected populations. This tallies with Hoffman's "parasitic foci", where reduced synaptic connectivity has given rise to a fragmented network, with some sections behaving quite independently of others. However, it is somewhat at odds with both Chen and Greenstein-Messica/Ruppin, who argue that the underlying neurobiological pathology is born of increased correlation, i.e. a lower-dimensional system.

A more biophysical approach to cortical stability has been taken by Wang and colleagues, culminating in an examination of schizophrenia in 2006 [215, 5, 216, 217, 3]. They constructed small networks much more analogous to those seen within cortex, incorporating excitatory pyramidal neurons and three subtypes of inhibitory interneuron (figure 2.3), with differential equations describing the conductances of multiple key receptor and channels. It worked on the principle that key processes such as working memory occur by virtue of the slow excitatory reverberation that is maintained by strong recurrent connections. Such a network is able to form the attractors crucial to all the previously mentioned models. Of course in order to be useful, these attractors need to be both stable when maintenance is necessary, and able to be destabilised/updated when appropriate. Wang and colleagues describe how this moderation can be enacted by intrinsic networks of GABAergic inhibitory interneurons. This network was able to relate cellular process to a spatial working memory task requiring both a maintenance of a spatial location, then a decision to saccade to that point following a delay, and has born out scrutiny though comparison to non-human primate electrophysiological data. Of relevance to cortical stability and schizophrenia were the following findings:

Stable working memory maintenance is dependent on there being a sufficiently high NMDA:AMPA receptor ratio on recurrent excitatory synapses This seems particularly valid given evidence of both NMDA dysfunction and working memory impairment in schizophrenia. It is worth noting that this is set apart from theories regarding NMDA receptors' role in long-term potentiation - here, it is by virtue of their slow conductance that they're so crucial to setting up the reverberation necessary for attractor maintenance. When Wolf and colleagues took an even more rigorous and detailed look at MSNs within the striatum [218], they too found that the ratio of NMDA:AMPA receptors had profound effects on the ability to entrain to afferent oscillation, which they argued would leave information integration vulnerable to a reduction in NMDA:AMPA. Loh and colleagues considered the role of NMDA in carving the landscape of an attractor network, relating it to the depth of each attractor basin, with less depth resulting in greater instability [219]. Siekmeier and colleagues applied the principle of NMDA hypofunction to the specific entorhinal to hippocampal CA1 circuit, relating the resultant dysfunction to

contextual memory tasks [220, 221]. Spencer used a similar architecture to address a slightly different question: given that gamma-band oscillations have reduced power in patients, on what components are gamma-band oscillations dependent [222]? Recorded electrophysiologically between 30 and 100Hz, these are believed to underlie the formation and selection of cell assemblies [223] (broadly analogous to the attractor basins modelled in artificial neural networks), and are known to be dependent on inhibitory FSIs [224]. He found that reducing the strength of excitatory synapses and decreasing FSI output approximates the changes seen in patient data, whereas decreasing NMDA input into FSIs - potentially a feasible underpinning of schizophrenia - actually increased gamma power, arguing against its involvement in the disease. This ability to produce testable hypotheses based on biophysical simulations is a powerful tool enabling the elimination of otherwise feasible theories.

Stability is modulated by dopamine via D1 receptors in an activity-dependent manner At lower activity levels, dopamine's dominant influence was to enhance NMDA conduction via D1 receptors, further stabilising activation, whilst also partly enhancing the efficacy of inhibitory interneurons, facilitating "winner takes all" manner of operation. At higher activity levels, NMDA receptors were effectively saturated, and dopamine's net influence shifted to enhancing GABA conduction, resulting in faster decay and destabilisation. This pattern of an "inverted-U" response of dopamine within prefrontal cortex is a recurring theme throughout the literature [225, 226] (figure 2.4), and is believed to come about through a relative displacement of the sigmoidal curves describing how NMDA conductance varies according to D1 receptor activation on excitatory pyramidal versus inhibitory interneurons. For interneurons, the response curve is shifted to the right, meaning NMDA conductance is able to remain responsive even at high levels of dopamine receptor activation (figure 2.5). Tanaka significantly expanded on this theme, implementing a closed-loop control system between dorsolateral prefrontal cortex and dopaminergic midbrain [227]. The module representing dorsolateral prefrontal cortex comprised excitatory pyramidal and inhibitory interneurons, which had a reciprocal relationship with simple dopaminergic units. Tanaka drew from evidence that highlighted the importance of tonic (as opposed to phasic) dopamine release, which alters NMDA, AMPA and GABA conductance via a D1 receptor mechanism within dorsolateral prefrontal cortex. This was termed the "operational control hypothesis" (as opposed to the phasic-release "gating hypothesis" which is believed to mediate switching and selection, see section 2.2.3), and provides a more dynamic, continuous means to adapt cortical stability. Tanaka's model allowed an exploration of the relationships between cortico-midbrain connectivity, dopamine release, D1 receptor acti-

vation and dorsolateral prefrontal cortex neuronal activity. They found that the model was able to generate robust representations when either cortico-midbrain connectivity or dopamine release exceeded a particular value. However, things became much less stable in circumstances where cortico-midbrain efficacy was down (which Tanaka described as analogous to the glutamate hypofunction hypothesis of schizophrenia), or dopamine release was diminished: if hypoglutamatergic, stability was impossible unless dopamine release was significantly elevated. This produced a system which became overly-responsive to fluctuations in dopamine, producing paradoxically and potentially-catastrophic increases in dorsolateral prefrontal cortex neural activity. Tanaka speculated that this may explain the confusion surrounding hypo versus hyperfrontality and working memory performance in patients. In later work, Tanaka described how a reduction of GABAergic inhibition could act to induce the catastrophically-elevated dorsolateral prefrontal cortex activity described in his previous model [228]. The relationships between excitatory pyramidal neurons, inhibitory interneurons and D1 receptor activation have been applied to neural networks of semantic priming and its altered performance in patients [229].

Durstewitz and colleagues' series of biophysical models has emphasised the importance of D1 versus D2 receptor-dominated neural networks, representing prefrontal cortex [179]. They describe a dual-state theory of dopamine function: in the D1-dominated state, D1 receptors act to deepen attractor basins and increase stability (in concordance with Wang, Tanaka and others); in the D2-dominated state, D2 receptors flatten the landscape, allowing for greater flexibility at the expense of stability. As to which state the prefrontal cortex finds itself in, this can be determined in the medium-to-long term by receptor expression, or on a more dynamic basis by the levels of dopamine being released: D2 receptors govern synaptic currents when dopamine levels are either high or low, whereas D1 receptors take charge at more intermediate levels. This is believed to be down to receptor localisation with respect to the synaptic cleft, with D2s and D1s being intra and extrasynaptic respectively. This posits another inverted-U response with respect to dopamine - one that can be probed pharmacologically and genetically. Durstewitz highlights findings that pharmacological manipulations that act to increase or decrease efficacy of dopamine release have the expected effects on stability/flexibility, in accordance with transitioning to D1- or D2-dominated states. The same patterns have been observed when considering variants of the COMT gene: this encodes an enzyme that regulates the rate at which dopamine is removed from the synaptic cleft post-release, with val/val carriers having less dopamine available than met/mets. It can therefore be expected that val/val carriers are in the low-dopamine, D2-dominated unstable state, whereas met/met carri-

ers are in the intermediate-dopamine, D1-dominated stable state. Evidence from behavioural and functional studies do indeed support the idea that met/met carriers are better at working memory tasks, and show more focal activation of prefrontal cortex, consistent with the notion of deep, focused attractors and little extraneous cortical noise [230, 231]. Durstewitz argues that in general, the D2-dominated unstable state predominates in schizophrenia, producing positive symptoms such as thought disorder, tangentialism, and other behaviours consistent with a loss of goal maintenance. A compensatory upregulation of D1 expression could shift the system into a more stabilised D1 state, but perhaps at the expense of flexibility, resulting in perseveration and stereotyped thinking. In general agreement with Tanaka, minor fluctuations in dopamine release could result in tremendous shifts between these states.

The work of the biophysical modellers has elaborated on the potential cellular processes underlying the more abstract concepts of Hoffman, Chen and Ruppin. This in turn opens up specific regions, receptors and circuits as being of especial interest for those wishing to understand and treat the illness. It also helps understanding move away from a purely symptom-focused point of view, with its tremendous diversity and multifactoral dependencies, and try to categorise very specific cognitions and behaviours in terms of an underlying neurological dysfunction. Here we find that a person's cognitive flexibility and goal-robustness may be understood in terms of shifts along D1:D2, NMDA:AMPA and excitatory:inhibitory axes.

2.2.3 Gating of cortical activity

Cortical stability concerns the property of reliably entering and remaining within a desired state. A failure in this regard will produce an unwanted spurious representation, or a complete breakdown of meaningful activity. However, the demands placed on cortical systems during normal functioning mean that absolute stability is not useful - they will have to adapt and shift between representations, i.e. be appropriately updated. The mechanisms determining how the current representation should be maintained, and how external influences should be allowed to perturb this, are often characterised as ”gating” mechanisms. Broadly, when the ”gate” is closed, the currently-active pattern of cortical activity remains relatively isolated from sensory input, and so is maintained by intra-layer activity. When the gate is open, sensory input is able to exert

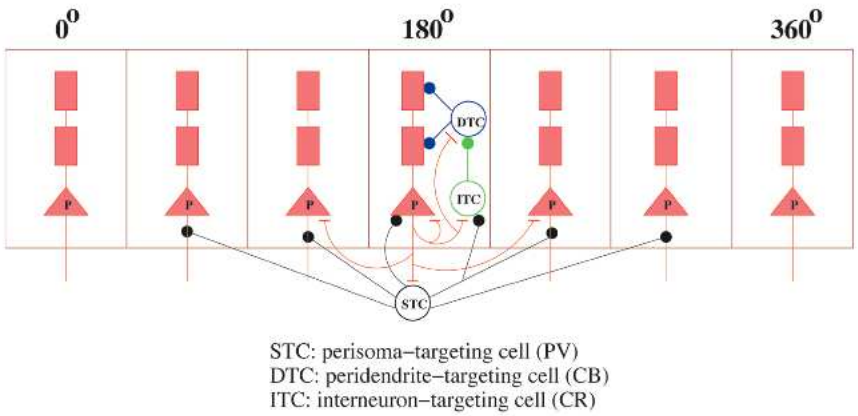


Figure 2.3: Wang's biophysical model of cortical microcircuits subserving spatial working memory [3], taken from Wang 2006 Pharmacopsychiatry. The model is able to perform a delayed saccade task, where the subject shifts their gaze to a point where a dot used to be, after an imposed delay. Each spatial location-encoding neural population is represented by an excitatory pyramidal neuron, having separate somal and two dendritic compartments. Three known varieties of inhibitory interneuron target specific compartments, with STC neurons being analogous to fast-spiking parvalbumin-positive; DTC neurons being calbindin-containing; and ITC being calretinin-containing GABA interneurons.

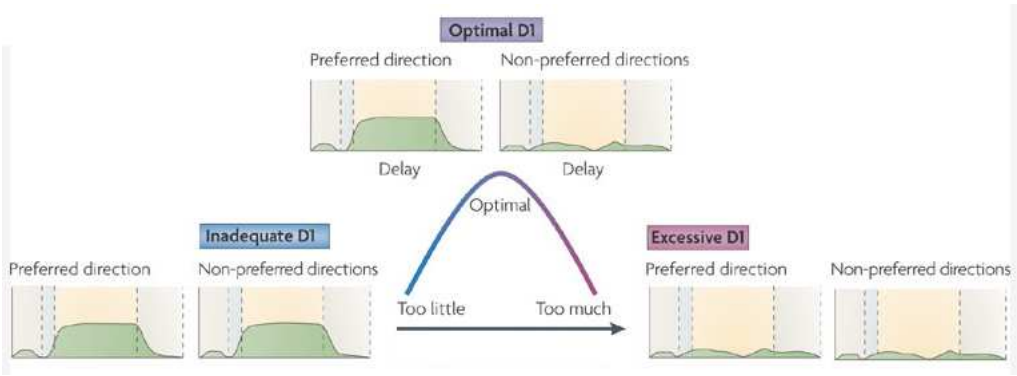


Figure 2.4: The inverted-U influence of dopamine on prefrontal cortical activation. Here, two spatially-selective populations are shown, with one being tuned to the location being maintained in working memory. Modified from Arnsten 2009 Nat. Rev. Neurosci. [4].

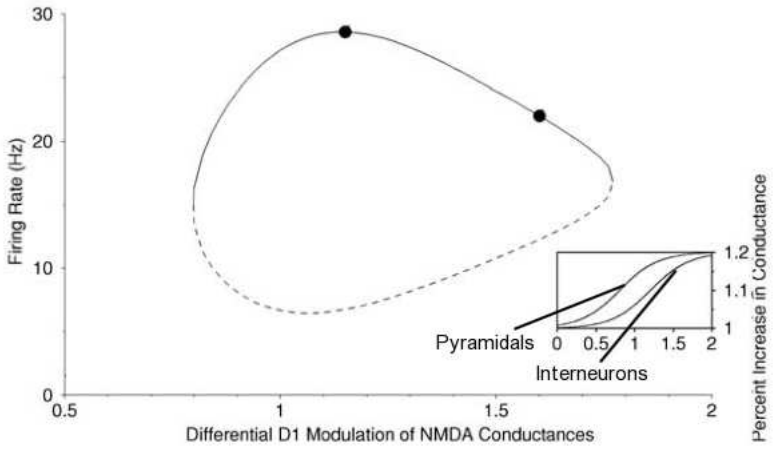


Figure 2.5: Dopamine's inverted-U response with respect to its differential modulation NMDA receptors on excitatory pyramidal versus inhibitory interneurons. Modified from Brunel & Wang 2001 J Comput. Neurosci.[5].

an overriding influence, causing the network to settle into a different, more suitable pattern of activation. Disturbances in the means to dynamically adapt to changing circumstances can be seen as the artificial equivalent of perseveration, distractibility, and the altered/erroneous recall of associations between previously-learned phenomena. Neurobiologically, this has been modelled as the influence that diffuse phasic dopamine release has over the network as a whole.

Deficient cortical dopamine function prevents appropriate updating Cohen, Servan-Schriever and Braver [232, 233, 234, 235, 236, 6] have implemented a series of connectionist feedforward backpropagation networks, with the Braver incarnation [6] (figure 2.7) performing the AX-continuous performance task (AX-CPT (figure 2.6)). This task's key demand is to maintain a presented cue in working memory, for comparison with a subsequently presented probe: if the combination is "A-X", then the participant indicates a target trial has been detected and an appropriate button is pressed. If not, an alternative button is pressed. This creates two measures of performance: context sensitivity (the ratio of A-X hits to B-X false alarms) and context cost (the degree to which reaction times for A-Y trials are slowed compared to B-Y trials). Behaviourally, patients show an overall fragility of context maintenance, in that they are less able to suppress a response to an X probe that has *not* been preceded with an A cue, and do not show a slowing in their responses to Y probes when it was preceded by an A cue. In the model. Braver examines the potential role dopamine has to play using it to "gate" prefrontal cortex: when dopamine fires at a low, tonic rate, prefrontal cortex maintains its current rep-

resentation and is unable to be externally influenced; once dopamine is released as a phasic burst, the context network is able to destabilise and respond to external input, before settling on a new representation. This is in contrast to the previous models explored, which investigated tonic release of dopamine: here the sudden elevation in dopamine acts via synaptic D2 receptors to destabilise established attractors. Inappropriate phasic release of dopamine is a key feature of models concerning schizophrenia and temporal difference learning (see section 2.2.5), but here Braver examined this by introducing random noise to the dopamine gating signal. This had two effects: the first was to ”dilute” the phasic signal itself, hindering its ability to destabilise the context module at the appropriate moment; the second was to elevate tonic activity to the point where spontaneous gating moments could occur, causing the context to be lost or replaced in an unhelpful way. A noisy dopaminergic gating signal could therefore underlie cognitive deficits, perseveration, thought disorder and a loss of goal-directed behaviour. Amos developed a different model that performed the Wisconsin card sorting task [237]: a challenge of cognitive ability that requires both goal-maintenance and set-shifting, and whose performance is understandably reduced in patients. Though Braver et al. had an abstracted unit where context and stimulus were evaluated, Amos explicitly denoted this as striatum, as a first step towards incorporating cortico-BG architecture. He too used dopamine as a gating signal to enable cortical destabilisation and rule updating, though this time its release was triggered by behavioural consequence: as negative feedback following an incorrect trial, which is more aligned with contemporary views of dopamine regarding reinforcement learning.

Monchi and colleagues elaborated extensively on the models of Braver and Amos: they too implemented a feedforward neural network that incorporated the BG and its recurrent parallel loops [238], explicitly rendering loops for posterior parietal, inferior temporal, premotor, orbitofrontal and dorsolateral prefrontal cortex, as well as incorporating a ”punishment module” in the form of amygdala input to limbic striatum (nucleus accumbens), bearing a resemblance to Amos’s model. Monchi characterised the neuropathology of schizophrenia as a dysregulation of the frontolimbic loop, probably mediated by dopamine dysregulation, by reducing the gain of the frontolimbic loop’s striatal node (nucleus accumbens). The model then performed the Wisconsin card sorting task, which produced two results: reducing gain within nucleus accumbens induced a general state of hypofrontality (reduced frontal cortical activity); and in agreement with Braver and Amos, diminished the impact of amygdala’s punishment signal, resulting in an imperturbable, perseverative system. On a related note, Moxon and colleagues developed a model that had similar themes to those of Braver, Amos and Monchi, though localised within different neural subsystems [239]. Instead of examining the gating role of dopamine on cortex,

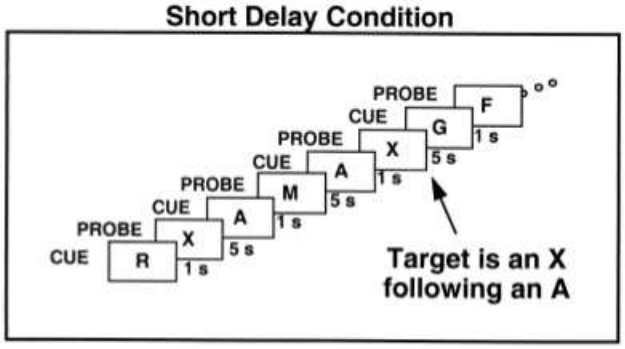
they instead consider cholinergic modulation of hippocampus, and propose that this can act to gate cortical influx into the hippocampal system.

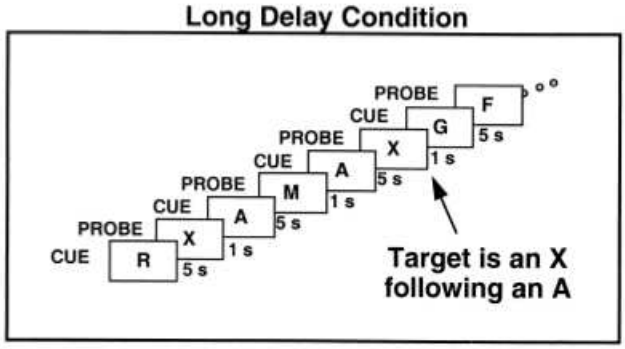
Reducing NMDA conductance enables easier switching As described earlier, Loh et al's model examines how the ratio of NMDA:GABA conductance alters attractor network dynamics: their simulations suggest that the ease with which a distracting stimulus can override an existing attractor and establish its representation as the dominant attractor state increases as pyramidal NMDA conductance decreases [219]. Enhanced GABA flow actually serves to make the system more rigid and resistant to switching.

As discussed earlier (see section 2.2.2), Durstewitz and Seamans contrasted the roles of D1 and D2 dopamine receptors in influencing cortical stability vs flexibility [179] in a biophysically-grounded model of prefrontal pyramidal and inhibitory interneurons. Simulated D2 receptor activation acted to reduce both the NMDA conductances of recurrent cortical excitation (hence reducing attractor basic depth), and the GABA conductances associated with "winner-takes-all" inhibition (introduces noise around the basin edge). They therefore defined the D2-dominated state as unstable, featuring a flatter attractor landscape where less energy is required to induce transitions between states. As D2 receptors have a relatively elevated affinity (versus D1) for dopamine when overall dopamine levels are low, the D2 state is hypothetically more representative of patient neural activity, engendering a system that is easily perturbed into falling into new states, perhaps leading to symptoms such as thought disorder.

2.2.4 Signal-to-noise ratio

Gating is not likely to be the clean-cut switch that artificial networks would ordinarily implement - cortical environments are necessarily noisy, subject to external influences from a variety of distant cortical and subcortical sources, as well as intrinsic intra-cortical randomness. In this setting, gating can be considered to be more continuous rather than discrete, gradually altering the "dimmer switch" with which the light of external influence shines on network activity. However, this has two components to it - the degree to which the desired pattern is potentiated, and the degree to which inappropriate activity is suppressed. Summarised as the SNR [192], several groups have focused on the role that dopamine release may have to play, either by modulating cortical activity through tonic vs phasic release, or by modulating striatal activation,





PROBE	CUE: VALID	CUE: INVALID
TARGET	A-X (70%)	"B-X" (10%)
NONTARGET	"A-Y" (10%)	"B-Y" (10%)

Figure 2.6: The AX continuous performance task, which involves many short-term and dynamically-updated associations. Letters are individually presented as a series of cue-probe pairs. The participant responds by pressing a button for every letter: the task is to press a different button whenever they see an X immediately following an A. A-X was presented most frequently, and the three possible alternatives (names B-X, A-Y and B-Y) much less so. In order to assess context sensitivity, the ratio of A-X hits to B-X false alarms is compared; likewise to assess context cost, they measure the degree to which reaction times for A-Y trials are slowed compared to B-Y trials. Two versions are modelled with differing working memory demands: in the short delay condition, there is 1s between cue and probe; in the long delay condition, this is extended to 5s. From Braver et al. 1999 Biol. Psychiatry [6].

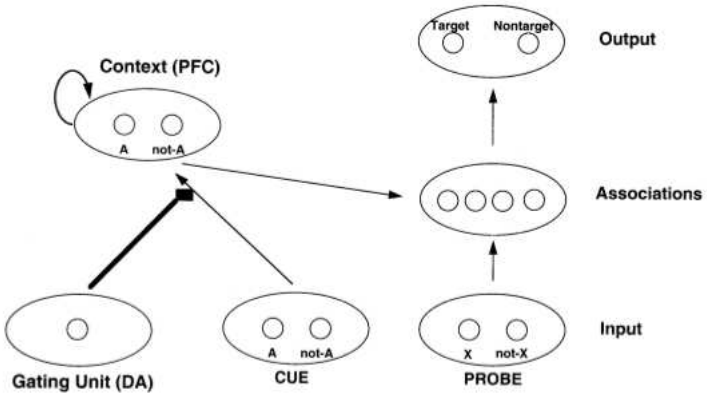


Figure 2.7: A schematic of the Braver model, set up to perform the AX-CPT. The context layer performs a working memory function through the use of recurrent excitatory connections, with competitive selection facilitated by inhibitory interneurons. This provides a representation of the previously presented cue, against which the probe is examined, and a target/nontarget decision made. The ease with which the context module can be updated is modulated by phasic bursts of dopamine from a gating unit. From Braver et al. 1999 Biol. Psychiatry [6].

and the resulting feedforward effects this would have on downstream cortex via corticothalamic loops. This offers important potential insights into schizophrenia, given the strong relationship between dopamine dysregulation and psychotic symptoms.

The Braver model described earlier [6] suggested how random noise introduced into a dopaminergic ”gating” signal could lead to perseveration, as both a diminished phasic response hindered destabilisation/updating, and an increased tonic response augmented maintenance of the current attractor (in line with Tanaka’s operational control hypothesis [227]). Taken together, this ”uninterruptability” would lead to a high but inappropriate SNR. However, the authors propose that on those moments where random noise acts to augment rather than detract from a phasic signal, updating will be made easier than otherwise it would have been, allowing inappropriate distractors to override the current cortical attractor. The inherent unpredictability and dynamic nature of these fluctuations could account for the symptomological diversity both within and between patients.

Taking a rather different approach to many of the models described so far, Peled and colleagues have taken the ideas underlying neural networks, but applied them at a larger scale, instead attempting to describe the connectedness of psychodynamical concepts [240, 241, 7]. They argue that if the brain represents particular phenomena as attractors across subsets of neural populations. The manner in which these are connected synaptically is related to how they are experienced psychologically and behaviourally. By formalising these conceptual

networks, they hope to better analyse the dynamics, and improve psychotherapy. Using abstracted feedforward neural networks, they examined SNR by assessing the influence both threshold gain (the steepness or ”speed” of the response with which post-synaptic potential translates into action potentials) and bias (the offset of that response) have on network behaviour (figure 2.8). A simple model of Rorschach image perception, constructed and tuned based on actual responses from a sample of patients, found that increasing the gain caused the network to alight on more bizarre responses with increasing probability [241]. Another model examining more involved tasks of sequential selection modulated gain and bias in a more dynamic manner - as a neuron remained active, these parameters were altered so that the threshold increased asymtotically, with the neuron becoming less active as a result. This allowed the network to evolve into new states, representing a sequence that was semantically meaningful. They found that fast, high-bias networks struggled to converge, sometimes producing a representation that bore little resemblance to the input (analogous to a hallucination); slow, low-bias networks tended to perseverate on a single network state (analogous to negative symptoms).

These notions of gain and bias are a generalised view of the aims of the biophysically-detailed models developed independently by different authors, concerned with altering the balance of recurrent excitation and competitive inhibition within neural populations. Within Wang’s model [3], any loss in NMDA conductance, or diminished D1 receptor activation, resulted in less competitive inhibition via the intrinsic GABAergic inhibitory interneurons. The resulting network was more vulnerable to interference from distractors. Loh’s less detailed model concurred with these findings, with reduced NMDA or GABA conductance decreasing SNR [219]. Durstewitz described this balance of recurrent excitation:competitive inhibition in terms of its relationship to dopamine receptor activation, with D2 vs D1 dominance producing a network whose states were more vulnerable to interference [179].

2.2.5 Salience and action selection

We have so far examined computational theories concerning pathology that may affect the brain’s ability to reliably maintain and appropriately update what it represents, whilst remaining robust to distractors. However, a final step must occur if this is to be directly related to symptoms: how will behaviour be altered as a result? The models investigating this aspect of the process focus on how learning of the inherent value or ”salience” of particular stimuli may be

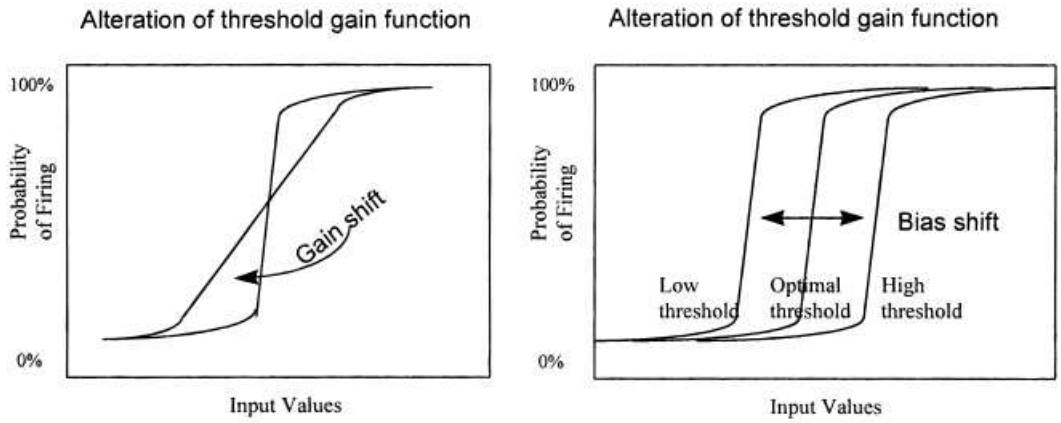


Figure 2.8: The relationships between neural firing, and both threshold gain and bias. From Geva and Peled 2000 JINS [7].

altered, and the cascading effects this has on selecting an appropriate course of action.

Moore and Sellen investigated how tonic subcortical dopamine levels could modulate the contribution new information made to the decision-making process [242]. They abstractly modelled a direct competition between two judgements, for which evidence either for or against each one was gradually accumulated. Dopamine increased the gain of the new information's impact - elevated dopamine levels led to a propensity to "jump to conclusions" - decisions were reached with less evidence, and decisions were more readily reversed upon presentation of non-confirmatory evidence. Moore and Sellen cited behavioural data that this is indeed what patients tend to do - especially those that are delusional - though this is perhaps at odds with the observation that delusions are fixed and robust to contrary evidence.

Grossberg has developed simplified representation of cognitive and limbic brain regions in his CogEM (cognitive-emotional-motor learning) model, instantiating motivational drive, psychological arousal and behaviour (figure2.9) across architecture analogous to prefrontal cortex, sensory cortex and the amygdala [8]. Given a particular perturbation, these elements must reach an activational equilibrium for proper function: Grossberg proposes that the region that represents drive - the amygdala - is unduly depressed in patients, propagating a disturbance through the system. Initially this would produce flattening of affect, but also dysfunctional learning of stimulus salience, and hypofrontality as amygdala provides less input into prefrontal cortex. From this would come two important deficits: a reduced ability to generate and pursue motivationally-relevant goals, and a greater difficulty in ignoring irrelevant stimuli.

This inability to ignore irrelevant stimuli can be directly examined using the latent inhibition

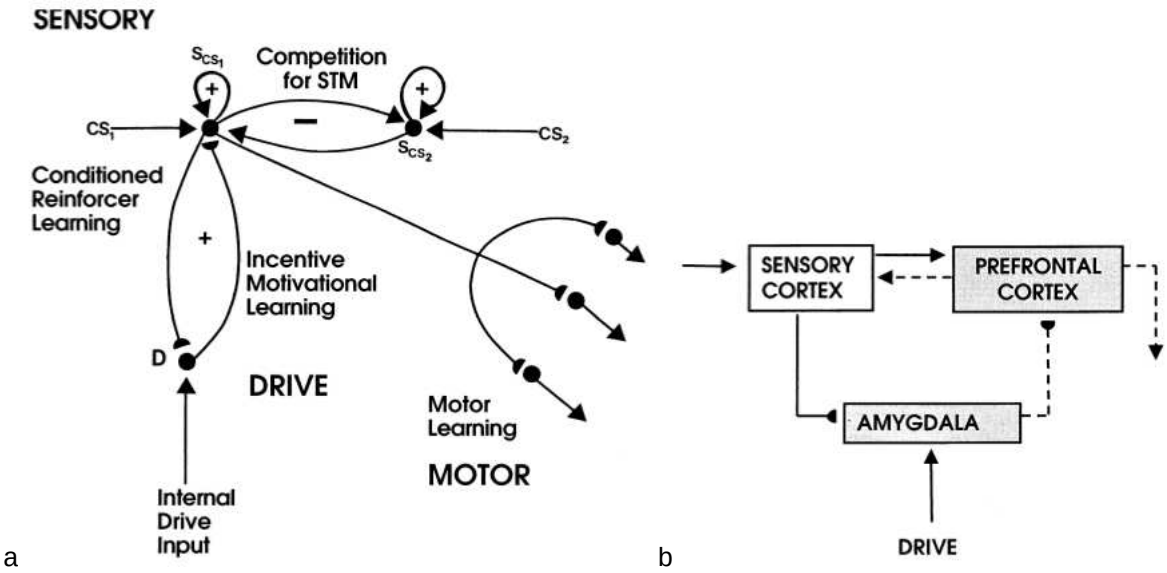


Figure 2.9: (a) The Grossberg CogEM model, relating motivational drive, recognition of sensory stimuli and motor behaviour. The model involved three types of learning: conditioned reinforcer learning (where a conditioned stimulus is associated with a particular outcome); incentive motivational learning (whereby a particular outcome is associated with any correlated sensory events, accounting for processes such as Kamin blocking); and motor learning (generation of tuned motor action plans). CS is conditioned stimulus, STM is short-term memory. (b) The proposed dysfunction in schizophrenia: amygdala activation is depressed, leading to reduced prefrontal cortex activation. From Grossberg 2000 Biol. Psychiatry [8].

behavioural paradigm. This modification of classical conditioning involves pre-exposing the participant to a particular stimulus, which is subsequently used as a conditioned stimulus once training begins. The rate of conditioning is found to be slowed, as there is a ”learned irrelevance” that must first be overcome before the motivational contingency can be learned. However, this doesn't seem to be the case with patients, for whom pre-exposure has no effect - they have not realised the initial irrelevance of the now-conditioned stimulus. Schmajuk modelled this phenomenon using an extended associability network, for which learning was explicitly modulated by novelty [9]. He proposed that the hippocampus is responsible for relating a conditioned and unconditioned stimulus together, but that this is deficient in patients. This then provides less input to ventral striatum, which normally suppresses tonic release of dopamine: the subsequently elevated dopamine potentiates the evaluation of the conditioned stimulus (CS), preventing the appreciation of its diminishing novelty.

Smith and colleagues also looked at latent inhibition in patients [243], but using a model derived from reinforcement learning: temporal difference learning [244]. This updates the assessment of value for a given stimulus according to the difference between its associated expected reward and the reward actually received: the prediction error. This signal has a close correlation with the activity seen in dopaminergic regions of the brain, and has led many to believe that representing prediction error provides a good account for phasic dopamine release [245]. This learning is back-propagated over time, so that the earliest stimulus to be associated with the reward takes on all predictive value. In itself, this does not provide an account for how latent inhibition is set up during pre-exposure, as nothing can be updated when no outcome is present. However it is quite informative of how latent inhibition is affected by medication: they model dopamine blockade as proportionally reducing the prediction error signal, leading to a restoration of latent inhibition, which is indeed seen in patients (and is why this is a popular animal model of psychosis).

In addition, Smith and colleagues developed an earlier model that still explored the role of dopamine, but using a different version of reinforcement learning that involved building an explicit internal representation of the world's contingencies. They focused on the functional significance of tonic dopamine whilst performing a conditioned avoidance task [10, 243]. Instead of focussing on temporal difference learning, they instead created an explicit internal model that represents the various states an agent might find itself in during conditioned avoidance, which is both updated by trial and error, and used to select the most appropriate action (figure 2.11). This internal model is argued to be maintained within prefrontal cortex, in much

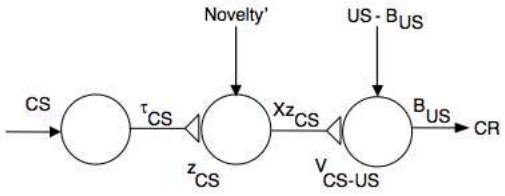


Figure 2.10: A simplified view of Schmajuk's model of latent inhibition. The degree to which a conditioned stimulus (CS) was associated with an unconditioned stimulus (US) is modulated by novelty. This is instantiated over a complex network including the amygdala, ventral striatum, ventral tegmental area (dopaminergic innervation), hippocampus, thalamus and prefrontal cortex. From Schmajuk 2005 Neurosci Biobehav. Rev. [9].

the same way as the previously considered models maintained working memory - it is therefore dependent on there being cortical stability. Smith and coleagues' main interest is in the final stage of action selection, and how this is influenced by tonic dopamine levels (again this is in contrast to temporal difference models that focus on phasic dopamine release). To select the most rewarding course of action, the model steps through the available actions, and aggregates the rewards associated with their probable state transitions over several time steps. Dopamine acts as a temporal discounting factor - lower dopamine levels means that less emphasis is placed on temporally-distal states. They argue that a chronic hypodopaminergic state within the cortex of patients could induce negative symptoms such as avolition, as distal rewards factor less into determining whether or not to commission an action. They also suggested that this lack of tonic dopamine would lead to a disinhibition of phasic dopamine, with aberrant bursts adversely altering the estimates of state-action transitions, broadly analogous to delusion formation [246].

Murray and colleagues took a significant step forward in directly relating temporal difference modelling to the neural activation of patients using fMRI [247]. Participants performed a simple task of instrumental conditioning, which involved learning which of two cues was most likely to yield a financial reward. The fMRI data was then regressed according to a mathematically-derived prediction error signal. This provided a good account for midbrain activation in healthy controls, concordant with the view that phasic dopamine is the neurobiological correlate of this signal. The same could not be said for patients, who instead showed an abnormally high response to the neutral, non-rewarding stimulus. This is in line with the observation that motivationally-relevant stimuli are not sufficiently attended to, and irrelevant stimuli cannot be ignored.

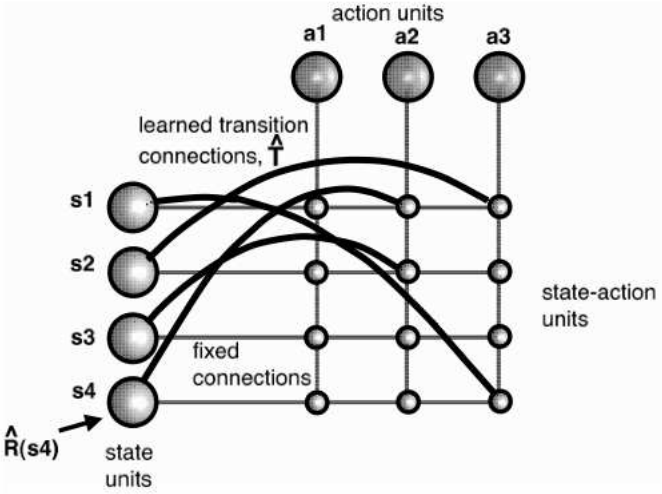


Figure 2.11: The explicit internal model of Smith's reinforcement learning-based approach. The state units represent each time step following the presentation of an environmental stimulus, such as the unconditioned or conditioned stimuli. These states are associated with a number of possible actions, represented by action units. In the context of a conditioned avoidance paradigm, these will be avoid, escape etc. As for other models of this type, the difference between expected and received reward is calculated and used to direct learning to both estimate the reward associated with each state, and the probability that a particular state-action combination will transition to another particular state. Having developed a sufficiently accurate representation of the environment and its contingencies, the model determines action selection by determining which action will result in the greatest future reward from the current state. Taken from Smith et al. 2005 Neural Comput. [10].

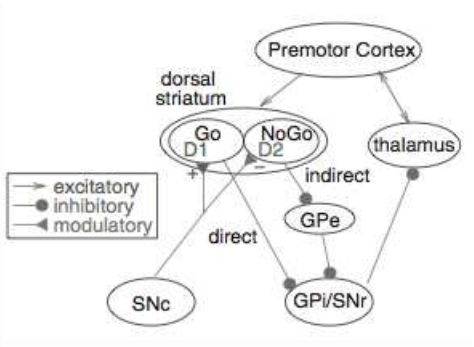


Figure 2.12: The simplest incarnation of Frank's model. Its key feature is the implementation of direct and indirect pathways through the BG. These are defined by subpopulations of striatal neurons that dominantly express D1 or D2 receptors respectively. These then project to different sections of the globus pallidus: the net effect is that the direct pathway disinhibits the relevant cortico-subcortical loop (the Go pathway), whereas the indirect pathway causes a competitive increase of suppressive activity, preventing cortical modulation (the NoGo pathway). The substantia nigra pars compacta (SNc) releases dopamine to potentiate the Go pathway via D1 receptors, and dampen the NoGo pathway via D2 receptors. From Frank and Claus 2006 Psychol. Rev. [11].

This principle of aberrant salience has been taken beyond the realm of simple behavioural decisions, and into cognitive processing: the ability to evaluate information and select goals. Frank and colleagues has created a series of models of cortico-BG function [248, 249, 11, 250, 251], elaborating significantly on intrinsic BG architecture (figure 2.12). This provides a different means to gate cortical activity to that of Braver et al. [6], and is more along the lines of Monchi et al's model [238]. Frank et al. proposed that the detection of motivationally-relevant phenomena, the maintenance of their representative attractor within working memory and the selection of appropriate action are all gated by this subcortical mechanism, which is itself profoundly modulated by dopamine. Within striatum, the functional role a neuron plays is defined by the dopamine receptor it more dominantly expressed: D1 potentiates firing, whereas D2 inhibits it. This is apart from the cortical understanding and modelling of dopamine modulation described in other models. Frank and colleagues report that increased striatal dopamine (as is observed in schizophrenia) leads to more liberal gating of cortical activity - which would arguably augment attention to irrelevance.

2.2.6 Summary

The efforts to computationally model schizophrenia have for the most part focused on four themes: stability of cortical representations; the ease of updating those representations when appropriate; how clearly those representations stand out against background noise; and whether those representations are motivationally-relevant, goal-directed and capable of guiding salient action selection. Some models have focused on behavioural expression and symptoms: their neurobiological detail tends to be brief. Others examined detailed, compartmentalised neural models, examining channel and receptor-mediated conductances for pyramidals vs interneurons, and dendrites vs soma: these provide rich comment on neural dynamics. Others still have produced attractor networks with simplified conductances: these go some way to bridging the divide. In general, it appears that cortical representations are stabilised when both recurrent excitatory connections and competitive inhibitory circuits are amplified: this is augmented by tonic dopamine acting via D1 receptors. Destabilisation and updating seem to be mediated more by the sudden elevations of dopamine provided by phasic release, acting via D2 receptors. This may occur purely intracortically, but given dopamine's involvement, it seems more likely that a complex interaction will be occurring between cortex and subcortical structures such as the BG and dopaminergic midbrain. This has been explored by Tanaka, Monchi, Frank and others, and suggests that the neurobiology of schizophrenia may be examinable via a set of relatively well-delineated nuclei, performing relatively simple tasks of motivationally-guided selection.

2.3 Frontrostriatal architecture

Plentiful evidence has been gathered of neurobiological dysfunction in schizophrenia. However the means by which this can all be integrated is less clear. Some aspects have been elaborated on within the computational implementations described in the previous section. In general, the prevailing hypothesis is of a pathology of glutamatergic/dopaminergic transmission, with some groups favouring an intracortical approach, and others taking a broader view of cortico-subcortical interactions. From the perspective of fMRI, the prefrontal cortex does not present an ideal region for hypothesis testing, given the limits of spatial and temporal resolution. However, subcortical structures such as the BG present a series of relatively well-delineated nuclei, making them more amenable to specific hypothesis testing. Here, we will therefore proceed along the lines that schizophrenia is a frontostriatal pathology.

In this section, normal frontostriatal and BG function will be explored, with a particular emphasis on how this modulates behaviour. Such relationships are frequently described using computational models. Using these theoretical frameworks, the potential relationships between frontostriatal dysfunction and the symptoms of schizophrenia will be elaborated.

2.3.1 Basal ganglia anatomy and physiology

The BG are a collection of subcortical nuclei that are structurally relatively well-delineated and functionally highly integrated [252, 253, 254]. They consist of the striatum, pallidus, subthalamic nucleus and substantia nigra (figure 2.13).

Striatum The primary input node of the BG, receiving a massive innervation from both the thalamus [255], and the majority of cerebral cortex. This is itself broken into the caudate nucleus (located dorsoanteriorly, targeted predominantly by prefrontal cortex) and putamen (positioned more ventrally, being a primary target of motor-related afferents). These structures share a characteristic intrinsic architecture, comprising a series of opioid receptor-dominated striosomes, surrounded by an acetylcholinesterase-enriched matrix [256, 257]. The functional significance of this is believed to mediate adaptive behaviour [258] and will not be elaborated on here.

90-95% of cells within the striatum are GABAergic MSNs [259], principally project to the pallidus, but also collateralise extensively with multiple varieties of inhibitory interneuron [260]. At a broader scale, the striatum exhibits a marked topography by virtue of its afferent connections: this projects through the BG, and is on the whole preserved through thalamus and back to cortex, forming looped channels [27]. For those projections originating from motor cortex, there is a marked somatotopy within striatum [261]. MSN conductance characteristics means that they're either in a "down" state, where afferent potentials rarely result in an action potential; or an "up" state, where excitability is much higher, albeit in a manner that requires significant co-excitation [252]. The requirement for multiple afferent potentials imbues MSNs with an integrative function. MSNs have been found to differentially express D1 and D2 receptors, with each being dominant for one of the two [262]. This affects how MSN activity is modulated by dopamine: D1 acts to potentiate excitatory potentials, wheres D2 diminishes their effect [262, 263]. D1/D2 distribution is also tied to the MSNs efferent targets: those expressing D1 project directly to the BG's output nuclei (internal globus pallidus and substantia nigra pars

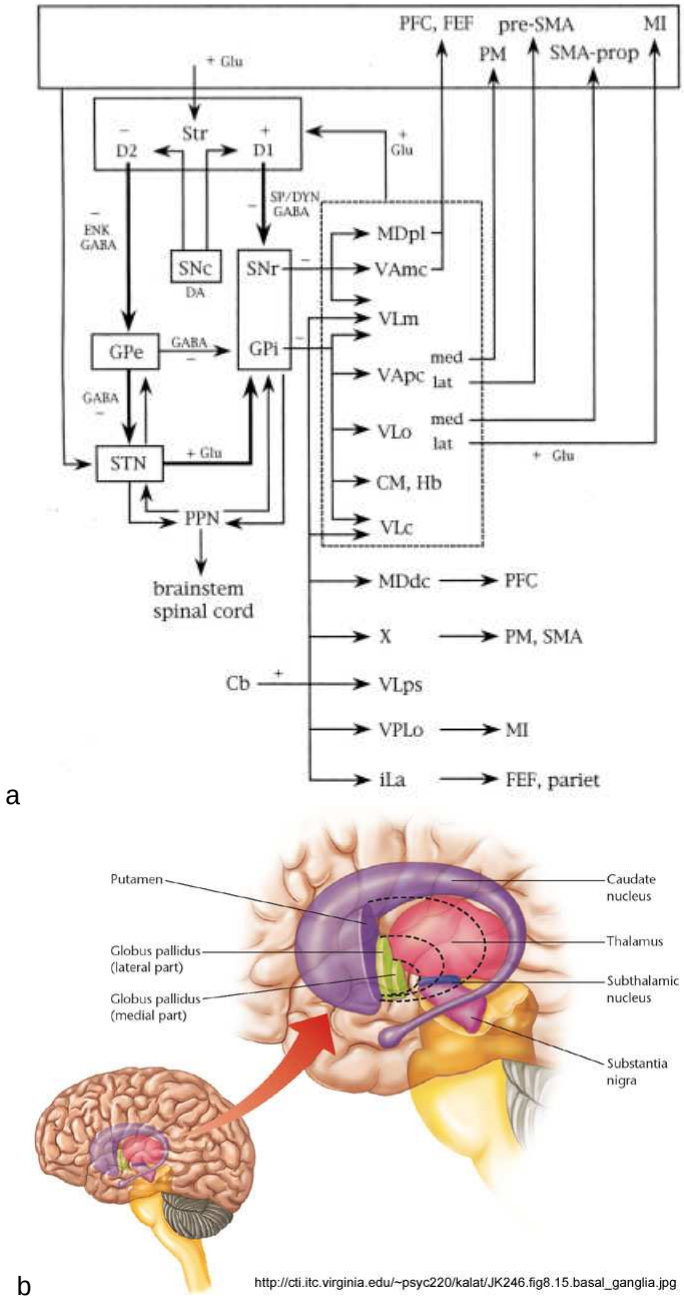


Figure 2.13: (a) Schematic of the BG and its projections back to cortex (via thalamus). ENK: encephalin; Cb: cerebellar nuclei; CM: centromedian thalamic nucleus; DA: dopamine; DYN: dynorphin; FEF: frontal eye field; GPi: globus pallidus (internal segment); GPe: globus pallidus (external segment); Glu: glutamate; Hb: habenular nucleus; iLa: intralaminar thalamic nuclei; lat: lateral part; med: medial part; M1: motor cortex; MDdc: densocellular part of mediodorsal nucleus; MDpl: paralaminar part of mediodorsal nucleus; pariet: parietal lobe; prefrontal cortex: PFC; PM: premotor cortex; PPN: pedunculopontine nucleus; pre-SMA: presupplementary motor cortex; SMA-prop: supplementary motor area (proper); SNc: substantia nigra pars compacta; SNr: substantia nigra pars reticulata; STN: subthalamic nucleus; Str: striatum; VAmc: magnocelluar part of ventral anterior nucleus; VApc: parvicellular part of ventral anterior nucleus; VLc: caudal part of ventral lateral nucleus; VLo: oral part of ventral lateral nucleus; VLps: pars postrema of ventral lateral nucleus; VPLo: oral part of ventral posterolateral nucleus; X: area X. Taken from Nakano 2000 Brain Dev. [12]. (b) A visualisation of the BG nuclei within the brain.

reticulata), whereas D2 neurons project along an ”indirect pathway” via the GPe [262, 252].

Subthalamic nucleus (STN) This is the BG’s second point of entry, receiving projections predominantly from motor and premotor and motor cortex [264, 265], but also inferior frontal regions [266], having an appropriate topography [261]. It demonstrates a coarser somatotopy than that of striatum, but this partial segregation remains present in projections to both pallidus and substantia nigra, which is tonically excitatory in nature [253].

Globus pallidus This nucleus is divided into two regions: the external (or lateral, GPe) and internal (or medial, GPi) segments. Both send GABAergic projections to their respective target nuclei: the internal segment is one of the primary output nuclei of the BG, projecting to the thalamus in accordance with topography; the external segment projects to both the STN and GPi, providing a modulatory way station along the indirect pathway [253, 252]. Thanks to the tonic and somewhat diffuse excitatory input received from STN, pallidal neurons are themselves tonically active: it is the more focal inhibition provided by striatum that disinhibits a topographically-confined region of pallidus, ”removing the brakes” from downstream thalamus.

Substantia nigra Like the globus pallidus, the substantia nigra is also divided into two segments: substantia nigra pars reticulata (SNr), which is the BG’s second output nucleus, and more involved in frontal loops that GPi; and substantia nigra pars compacta, which is a key dopaminergic region of the brain, primarily innervating the striatum [252].

Cortico-subcortical channels Although there is some degree of overlap and integration as each cortical signal is propagated through the BG [267], somatotopic representations are essentially maintained at the point of exit in GPi/SNr [261]. This is believed to result in a series of separate channels - each tonically inhibiting its associated output pathway, and each capable of being selectively disinhibited in accordance with voluntary action [24, 268, 269]. Yelnik described the BG as being analogous to a piano [270]: complex action plans would be implemented as chords, with sequential movements being the equivalent of a harmonic progression. The way in which this is modulated by dopamine within the striatum - the differential potentiation of the direct pathway, and inhibition of the indirect pathway - provides fertile ground for theories concerning its computational purpose, and potential role in various pathologies.

2.3.2 Basal ganglia function

Traditionally, the BG have been viewed as a crucial modulator of motor production: this is in part due to the clear links between its damage/destruction and conditions such as Parkinson's and Huntington's disease [271]. However, given its extensive connectivity with cortex, its unsurprising that its apparent influence is extending deeper into more cognitive roles [272], and the profound influence of dopamine highlights its importance as a centre of learning mediation [273]. Above all, it's becoming increasingly clear that the BG are essential in mediating goal-directed action [274].

Motor control As motor cortex sends its primary projection towards the spinal cord, a proportional number of collaterals diverge off towards the striatum, to form a modulatory loop via thalamus. It has been a long-held view that this fine-tunes and finesses the primary signal, enabling better control of parameters such as force, rate of change of force, and coordinated contraction of multiple muscle groups [275, 276]. This is underwritten by the common demand of accurate sequencing, a function that is related to another of the BG's emerging roles: rhythm perception [277] and the explicit appreciation of time [278].

Learning This principle of remodelling behaviour - primarily in response to error-derived feedback - has been shown to be of crucial importance with respect to how dopamine modulates BG activity [279]. Schultz, Dayan and Montague observed that phasic dopamine release bore a strong resemblance to prediction error [245] - a signal described by temporal difference learning models as part of reinforcement learning theory [244]. This is made most clear during tasks of classical conditioning, where a positive or negative cue-response relationship is gradually learned through experience [280]. This influential school of thought has driven much of the subsequent work concerning dopamine within the striatum: not just from the perspective of action acquisition, but also for subsequent action selection and execution [281, 282]. In particular, studies involving tasks of instrumental conditioning during fMRI have found that the limbic regions of the striatum - the nucleus accumbens - bears a correlation with anticipated reward value [283, 284].

Goal-directed action The ultimate objective of detecting salient phenomena - be they cues, responses or outcomes - is to better guide future action towards achieving particular goals. In addition to motor dysfunction, Parkinson's disease can also be associated with marked difficulties in pursuing goals, for example difficulty in controlling attention [285] and in skill acquisition

[286]. Indeed it has been proposed that the determination and continued pursuit of goals best articulates BG function [287]: this is apparent in the types of tasks that are especially demanding of the BG, such as categorisation [288], set-shifting [289], and preventing impulsivity [290]. In each case, it is the theoretical capacity to perform action selection that gives the BG its prominent role [268, 274].

Action selection The phrase ”action selection” requires a clear definition of what an action is: for the most part, this is taken to mean a motor ”programme”, a series of coordinated movements, the full collection of which would be an organism’s motor repertoire [268, 291]. Deciding which action to take in a given situation is a problem requiring multiple sources of information to solve [16]. Incomplete information gives rise to uncertainty, the resolution of which engages a distributed cortical network (for recent reviews, see [17, 18, 19]). It remains unclear how they jointly act to produce motor output, though computational approaches [20, 21, 22] suggest that centralized systems such as the BG provide an efficient solution to the “action selection” problem [23]. In the absence of a simple reflexive relationship between current behavioral context and a course of action, available options must be evaluated to allow the execution of any single behaviour. Efficient selection therefore depends upon the suppression of competing alternatives. Biologically, the BG seem optimally configured to arbitrate motor decisions [24, 25]. Acting as a key way-station within corticothalamic loops [26], the BG exert a suppressive influence via partially-segregated "channels" [269]. Given sufficient cortical input, the BG’s influence is theorised to become focally disinhibitory, removing suppression and allowing the execution of a single course of action.

2.3.3 Describing frontostriatal selection: computational modelling

Models that represent the interactions between cortex and elements of the BG offer a means to look at each of these themes in a structured way. They can potentially benefit from two key properties: to capture several clinically important elements, such as ”hypofrontality” and dysfunctional subcortical dopamine release; and to address neural function at a scale that is accessible to functional neuroimaging.

Many modelling implementations have focussed on the computational power of dopamine’s influence over striatal neurons: more directly during the detection and learning of environmental contingencies, but also as to how that is then used to select the most rewarding action. To

this end, such reinforcement learning-based models use the softmax activation function - a biologically-plausible means of selecting the most valuable action [292].

Others have considered the looped, channeled architecture as operating at a more systemic level: for example, ”cognitive” planning circuits, versus procedural learning streams [293]. In this context, the BG’s proposed role is to arbitrate between competing sources of evidence. Nakahara and colleagues implement this within a reinforcement learning context: for a task of visuomotor sequencing, two systems compete to encode actions both in terms of the spatial locations of the target objects, and the joint-angle coordinates of the required movements. The BG’s role is to mediate when conflict arises between the two, albeit in an abstract mathematical implementation [294]. Haruno and Kawato build on the spiralling nature of the looped architecture argued for by Haber [295, 254], modelling ”heterarchical reinforcement learning” [296]. The need for arbitration is especially pressing when the sensory information guiding a particular course of action is ambiguous, and requires some accumulation over time. Shah and Barto have considered this using a connectionist-style approach that doesn’t faithfully replicated BG architecture, but does represent a competitive process between a ”planner” controller (analogous to prefrontal cortex), which only operates when input is resolved, and a ”value-based” controller (analogous to the BG), which can select in a weighted manner at intermediate time points [297].

These approaches have tended to be focused on the adaptive processes of learning rewarding actions: models exploring action selection in a more isolated context have characterised the BG’s parallel loops as being analogous to specific action plans - a more direct competition between possible responses, regardless of the source of the sensory evidence. Consequently, they also claim a closer relationship to the underlying anatomy, with ”units” taken to represent specific BG nuclei. Within this class, there are the models of Monchi et al. [238], Amos [237] and Houk et al. [298] (see section 2.2). A much more biophysical approach was adopted by Lo and Wang [299]: here the role of the BG was to set the threshold at which accumulating evidence could finally force the selection of the indicated action. However, the ”BG” only involved a single striatal and single internal pallidal unit per action, with competition being enacted as direct mutual inhibition - this does call into question the role of the STN and GPe.

Rubchinsky and colleagues go some way to addressing this, by exploring the part the GPe and STN could play in suppressing unwanted actions [300]. With one channel active, they find that the GPe-STN complex acts to suppress a gradually ramping second input, until its evidence overwhelms that of the first channel. However, their model has focal subthalamic projections, which is not supported by neuroanatomical data. It tended to fail if the STN’s projections were

implemented in a biologically-supported diffuse manner.

Given the BG's characteristic inter-nuclear connectivity, it is likely that more fruitful approaches will attempt to incorporate a little more of this complexity. Brown and colleagues implemented the relatively well-understood saccadic system, incorporating significant cortical complexity [301]. Their Telencephalic Laminar Objective Selector model (TELOS) creates motor plans and their associated movement parameters within different layers of frontal eye field cortex. Selection is mediated by the direct pathway through the BG, whereas the indirect path acts to suppress less well-supported alternative action plans.

However, this architecture is rather specific to performing saccades: Frank and colleagues have created a more generalised model, that successfully incorporates a lot of the BG's inter-nuclear detail. They too express the facilitative direct and suppressive indirect pathways, complete with the appropriate dopaminergic D1 vs D2 modulation [302, 303, 304]. As well as selecting motor action plans, including fine-grade motor control [305] this system is seen as a suitable means to mediate the kind of cortical gating described by Braver in 1999 [6], enabling the updating of working memory. Although much of the BG's pathways are described in these models, the STN remains somewhat on the fringes, being reduced to a relay along the indirect path - or more recently as a ”global no-go” pathway, that suppresses all action by means of its diffuse projection to GPi.

One series of models that gives the STN a more nuanced and centralised role is that of Gurney, Prescott and Redgrave [13, 28, 306, 307, 308]. Their general model incorporates many of features that are of particular relevance to schizophrenia: the explicit description of the role D1 versus D2 dopamine receptors play in action selection, and how the whole BG system regulates the global suppression exerted on competing actions. It does so using a neurally-grounded connectionist model, producing predictions of activity that are accessible at the resolution achievable with fMRI. The Gurney model will therefore form the basis of the computational aspects of this thesis.

2.4 The Gurney model of basal ganglia-mediated action selection

Though created by Gurney, Prescott and Redgrave, for the sake of brevity this model shall henceforth be referred to as the Gurney model. It is grounded upon empirical anatomical and electrophysiological findings, qualifying it as a useful test bed for in vivo studies. It is an artificial

neural network consisting of the major BG nuclei, and is capable of selecting between different "channels" (analogous to actions) by selectively inhibiting its output nucleus (globus pallidus) in response to increases in afferent "drive", analogous to cortical activity. A given action is associated with one unit (representing a population of neurons) within each nucleus, arranged as a series of segregated channels across the architecture of the BG. Cortical activity drives a competition process across the network via striatopallidal and reciprocal subthalamopallidal interactions, the result being the selection of the most cortically-active channel through focal inhibition of associated pallidal activity.

2.4.1 Defining an action

In agreement with Mink [24], the Gurney model describes how the BG is ideally suited to acting as a centralised selector of action, drawing from a repertoire of motor plans. The drive behind each potential action is termed its "salience". This is described as being a summary of the intrinsic and extrinsic causal factors denoting the importance or urgency of an action [25] - its propensity to be selected for execution [13]. The examples they provide consist of higher-order behaviours such as drinking vs talking, with survival being the key factor defining urgency. They do however suggest that selections between higher-level behaviours are in fact selections between combinations of lower-level selections. Their proposed hierarchy consists of (1) behavioural strategy; (2) action - which would be for example chew or whistle; and (3) movement - a component of an action, such as "close jaw". An action is anything that can be represented by a population of BG neurons: they leave it to others to define what level of action the BG mediates, be it an elementary motor act or comprehensive behavioural strategy. During this work, an action will be defined as any goal-directed motor response.

It is arguable that salience is just one factor that contributes to the overall appropriateness of a particular action. The use of the term "propensity" implies that further factors that must be considered during a decision-making process include the associated magnitude of any positive/negative outcome, and the likelihood of that happening. There is evidence that such properties - summarised by the economic term "utility" - are coded for in parietal regions. Within a given context, some actions will be associated with an increased probability of obtaining a positive outcome - a property that is analogous to certain actions being considered more "correct" than others. Therefore it is arguable that within the context of a simple stimulus-response-type task, the drive to perform the correct action should be proportionally greater than that of all other options, in an effort to maximise the likelihood of a positive outcome. For optimal

performance/gain, an agent must select the most appropriate action.

2.4.2 Defining the selection mechanism

The Gurney model reconceptualises the BG's potential functional role as being one of signal selection [309, 13]. It features a neural network capable of selecting between different "channels" (analogous to actions) by selectively inhibiting its output nucleus (globus pallidus) in response to increases in afferent cortical activity. Through disinhibition, the system implements off-centre on-surround transforms, with high inputs becoming low outputs.

Here, a set of n inputs X were transformed into a selected set S, a non-selected set $\overline{S}$, and an intermediate set Y_0, on the basis of two selection thresholds θ_1 and θ_2, such that:

$$y \in Y_0 \Rightarrow \theta_1 < y < \theta_2$$

As the BG was believed to select through disinhibition, it would effectively be implementing small-signal selection, that is, the mapping of inputs x to outputs y would observe the order preserving relation:

$$x_i \leq x_j \Rightarrow y_i > y_j$$

where i and j represent different channels. The degree to which this selection mechanism implemented a competitive process was summarised by two properties: decisiveness and promiscuity. Decisiveness was defined as:

$$D(x) = 1 - \frac{|Y_0|}{|X|}$$

When $D(x) = 1$, this denotes a "clean" selection where no signals are left in the intermediate group, and $D(x) = 0$ occurs when selection has not occurred. Promiscuity was defined as:

$$\phi(x) = \frac{|S|}{|X|}$$

This described how many channels can be selected at one time: if $\phi(x) = \frac{1}{n}$, only one channel is selected, and the system is described as implementing "hard switching". If $\phi(x) > \frac{1}{n}$, multiple channels can be simultaneously selected in a system that is said to be operating "soft switching".

In the general case, they describe the activation a_i of a node representing a single channel with respect to its competitors:

$$a_i = -w^- x_i + w^+ \sum_{j \neq i}^{n} (x_j)$$

Where w^- and w^+ are the inhibitory and excitatory weights of inputs. In order for this to implement off-centre, on-surround disinhibitory selection, for channels p and k, $a_p - a_k = (w^+ + w^-)(x_k - x_p)$, so if $(x_k > x_p)$ then $a_k < a_p$, i.e. order is preserved for small-signal selection. To determine the balance between the excitatory and inhibitory weights (to allow for a useful, decisive and promiscuous system), they set $w^+ = \delta w^-$, and determine that in order to stop activation from just increasing with an increasing number of channels to saturation point, δ should be $O(\frac{1}{n})$. Also, if the activation function transforming a_i to node output y_i is $y_i = m(a - \epsilon)$, there needs to be a non-zero tonic input else everything will be selected in the quiescent state - i.e., $\epsilon < 0$.

2.4.3 The basal ganglia model

The general architecture of the Gurney model bears some resemblance to that of Frank and colleagues, with the exception of the role they bestow upon the STN (figure 2.14). This provides a diffuse and tonic excitatory input into both segments of pallidus, which serves to ”apply the brakes” onto what they argue is a disinhibitory selection mechanism (figure 2.15). This elaboration also draws the GPe more to the fore - the indirect pathway is no longer a general suppression mechanism, but something that is able to tune the system in accordance with overall demands. As the indirect pathway has a stronger relationship with striatal neurons expressing D2 versus D1 receptors, this also provides a means by which the effects of anti-psychotic medication could be understood and explored. A more detailed version of the model and a reference for its weights and parameters is provided in figure 2.16.

Striatum As implied by neurophysiological evidence and the suggestions of others, they envision the striatum as containing many inhibitory recurrent networks, each containing a small number of competing channels. Some competitions are resolved here, with the winners being passed forward to the rest of the proposed selection mechanism. However, they state that their model will only be compatible if the net output is not ”winner takes all”, that is, its salience is not maximally ramped up by its losing neighbours. This would be in accordance with recent electrophysiological evidence [260], that suggests that MSN collaterals within striatum are not able to mediate such lateral inhibition.

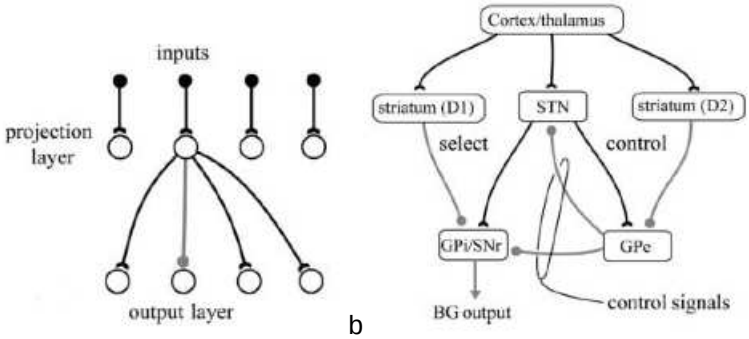


Figure 2.14: The principles underlying the Gurney BG model of action selection. (a) Disinhibitory selection via an off-centre on-surround schema. (b) The anatomy of Gurney's connectionist model. The direct-D1 and indirect-D2 pathways are implemented, and the STN is given a prominent role in "applying the brakes" to the disinhibitory system. The unique feature of the Gurney model is this role for STN, and the adaptive control that can be exerted through STN-GPe interactions. Images taken from Gurney et al. 2001 [13].

Due to the prevalence of dopamine synapses on the shafts of MSNs, they argue it has a multiplicative effect on the input from cortex into striatum. In the model, w_s is the weight denoting this connection, which is therefore modified to become:

$$w_s(1 + \lambda_g)$$

for the "select" (previously direct), D1-mediated pathway, and

$$w_s(1 - \lambda_e)$$

for the "control" (previously indirect), D2-mediated pathway, in accordance with the view that D1 and D2 receptors respectively potentiate and suppress the ability of MSNs to move from a down to an up state. Cortical input is deemed to arise from a number of sources, and together these are integrated within striatum to produce a salience c for each channel.

STN The model assumes that the STN is able to approximate salience to some degree. However, it lacks the intrinsic architecture of the striatum. They therefore considered adding a channel-dependent noise term, but given the summed nature of its projections to its output nuclei, it was assumed that this became irrelevant for large numbers of channels, and so they didn't include it in the model. A similar view was taken for cortical inputs from channels that had not achieved an up state within striatum - these will still provide input to STN, but instead of approximating this with a noise term, they added a positive ϵ' term to its activation function. The model calculates each STN channel individually, and this is summed in formulae receiving STN input, representing tonic diffuse excitation.

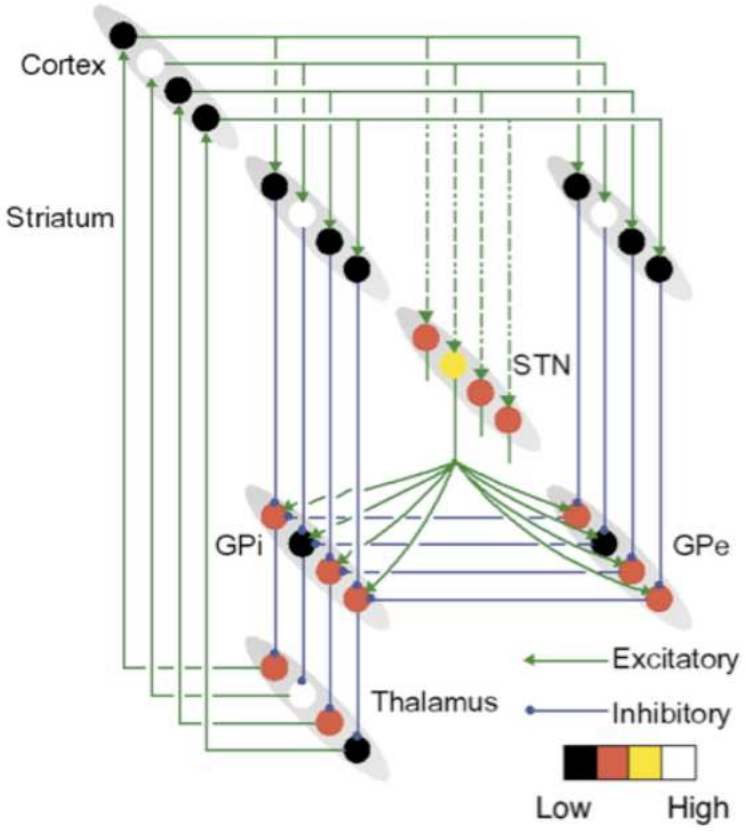


Figure 2.15: A diagramatic representation of the Gurney model of BG-mediated action selection [13, 14], showing four cortico-BG-thalamic loops. Circles within anatomical regions represent neural populations associated with a particular channel and hence action, with tone indicating activity. GPi: internal globus pallidus; GPe: external globus pallidus; STN: subthalamic nucleus. Gurney and colleagues argue that this is ideally mediated by BG neuroanatomy: the inhibitory striatopallidal projection is focal and remains relatively confined within channels, whereas the excitatory subthalamopallidal projection is more diffuse. The pallidus acts to suppress thalamocortical projections, and is kept tonically active by diffuse excitatory subcortical inputs (here only one diffuse projection is shown for clarity). Cortical activity associated with a particular action disinhibits a closed loop, thereby selecting that action.The figure shows the effects of decreased corticosubthalamic influence (dotted lines) and hence Selectivity on BG activation.

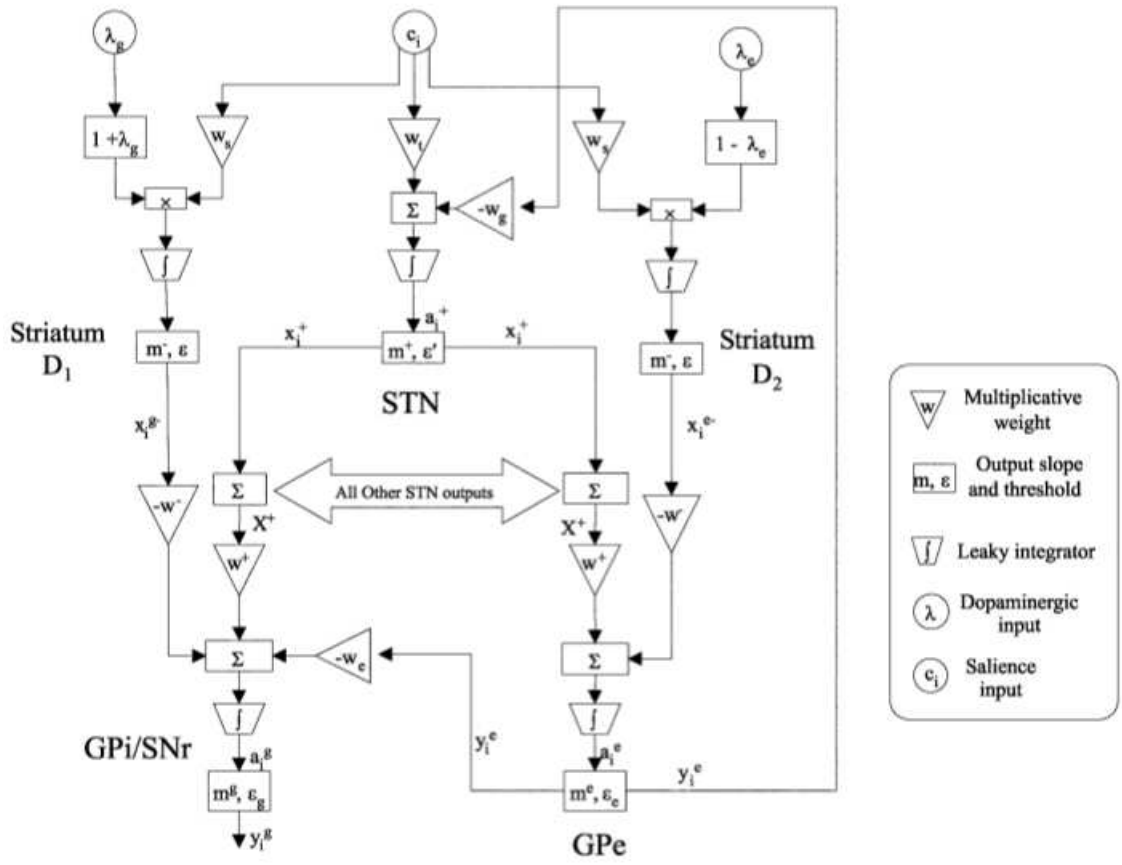


Figure 2.16: A diagram of the units, weights and activation functions describing the Gurney model of BG-mediated action selection. Image from Gurney et al. 2001 [13].

$$X^+ = m^+ \sum_{i=1}^{n} (w_t c_i + \epsilon' - w_g y_i^e) H(\tilde{a_i^+})$$

They determine that in the quiescent state:

$$\hat{X^+} = \frac{n(\epsilon' - w_g \epsilon_e)}{1 + n w_g \delta w^-}$$

and go onto prove logically that S^* (the number of active STN channels) can never be empty so long as $\epsilon' > w_g \epsilon_e$, providing a tonic excitation of pallidus.

GPe They set $w^+ = \delta w^-$, and determine that the equilibrium activation is

$$a_i^e = w^-(\delta X^+ - x_i^{e-}) + \epsilon_e$$

Inserting the striatal control pathway for x_i^{e-} gives:

$$a_i^e = w^-\{\delta X^+ - m^-[(1 - \lambda_e) w_s c_i - \epsilon] H_i^{\uparrow}(\lambda_e)\} + \epsilon_e$$

GPi The STN's input into GPi is identical to that of GPe, in line with anatomical evidence of collateralisation. They also assume that STN-pallidus synaptic efficacies for both internal and external segments are the same. The activation equilibrium, incorporating diffuse STN excitation, selection pathway and GPe point inhibition is:

$$a_i^g = w^-(\delta X^+ - m^-[w_s(1 + \lambda_g) c_i - \epsilon] H_i^{\uparrow}(-\lambda_g)) - w_e y_i^e + \epsilon_g$$

Though analysis, Gurney determined that the GPi would always exert its suppressive activity so long as $w_e < 1$.

2.4.4 Selection characteristics

Gurney set several criteria that their model had to meet in order for it to be considered a selection mechanism. First, sets of channels were defined:

- P - set of potentially active channels
- P_g^* - set of active (non-zero) channels output from striatum on the select path
- P_e^* - set of active channels output along the striatum's control path
- P^* - set of active channels on both pathways.
- $\wp_0$ - non-select GPe channels, i.e. $P - P_e^*$ (tonically active)

- $\wp$ - selected (part of P_e^*), and active GPe, $y_i^e > 0$
- $\wp^*$ - selected and non-active GPe, $y_i^e = 0$
- S^* - set of channels for which $H(\tilde{a_i^+})$ =1 (the STN's stepwise function) i.e. active at STN
- n^s - number of elements in S^* i.e. selected at STN

Order preservation In order to be a disinhibitory selector, the model needed to transform high salience cortical inputs into low GPi outputs (small-signal selection). Through analysis [28], they find that signal ordering is preserved if $c_j > c_i \Rightarrow y_i^g - y_j^g \geq 0$, and by extension $c_j > c_i \Rightarrow a_i^g - a_j^g \geq 0$. They investigate several cases according to the three possible GPe and four striatal channel sets, and determine that $y_i^g - y_j^g \geq 0$ if and only if both y outputs are 0, or neither channel is in the selected P^* set.

Capacity scaling This is dependent on the balance between excitation and inhibition of the GPi, to allow the selected set to be non-empty, and to prevent all channels saturating at 1. This could have been solved in two ways: by conditionalising the inputs, making mean x a function of the number of channels (which lacks biological plausibility), or by setting the ratio between excitatory and inhibitory inputs to be constrained by the number of channels.

In the analysis this was investigated by determining the difference between two competing channels Δ^i with respect to tonic levels:

$$\Delta^i = \delta(k_1 X^+ - k_2 \hat{X^+}) - L(\lambda_g, \lambda_e)c_i + k_3 \tag{2.1}$$

where k_1 gathers terms to do with active output, and k_2 gathers terms concerned with tonic activity, so $k_1 > k_2$ i.e. $\Delta^i > 0$.

The effect of GPe If GPe's inhibition of STN vanished, all channels would be selected at the STN. X^+ (both active and tonic) would be reduced to a form dependent on n, and given the STN's dominance in Δ^i, would demand a constraint be placed on δ. In other words, with no GPe, capacity scaling is only achieved by placing a O(1/n) constraint on δ. This is biologically dubious.

Considering STN tonic activation:

$$X^+ \leq \frac{n^s c_0}{1 + \delta w_g w^- n^s \varphi_s^m}$$

where

$$\varphi_s^m = \frac{|S^* \cap \wp \cap \wp_0|}{n^s}$$

where c_0 is tonic cortical activity, and n^s is the number of channels active within striatum. In theory, n^s can increase, but φ_s^m is bounded thanks to its division by n^s so X might not have an upper bound. However, they show that given a sufficiently large X, there will be no non-active GPe channels. Therefore $\varphi_s^m = 1$; i.e. it can be reduced to $\varphi_s^m = \frac{S^*}{n^s}$. Logically, they then show that X is bound from above by its own input into GPe, which in turn suppresses STN. It then follows that $(k_1 X^+ - k_2 \hat{X^+})$ is upper bounded. Therefore $\delta(k_1 X^+ - k_2 \hat{X^+})$ is bounded without having to manipulate δ. GPe is therefore innately able to maintain capacity scaling as $n \to \infty$, purely by virtue of its connectivity, and without any "hard coding" requirements.

Competitiveness This is the degree to which activity on one channel causes suppressive excitation of its competitors. Here they quantify the competitive aspects of the model by examining:

$$\frac{d\tilde{a_i^g}}{dc_j}$$

Two main conclusions are drawn:

Effects of increasing salience on competitors If $i \neq j$, and i is active at the STN, $da_i^g/d\tilde{c_j} > 0$, i.e. GPi activity increases for i competitors as the salience of j increases. If i is not active at the STN, $da_i^g/d\tilde{c_j} > 0$ remains at 0. An increase in salience in any channel at the STN will increase its excitation to other active channels at the GPi. Given a sufficiently large competing salience, the currently winning channel can be forced too high to be considered selected, allowing the competitor to take over.

Effects of salience within a single channel If i is not active at the STN, increases in the salience of i causes $da_i^g/d\tilde{c_j} < 0$, i.e. its GPi output decreases, unless it is not selected striatally in which case $da_i^g/d\tilde{c_j} = 0$. Likewise, if i is active at the STN, increasing c_i also makes $da_i^g/d\tilde{c_j} < 0$, unless it is not striatally active, in which case $da_i^g/d\tilde{c_j} > 0$, i.e. STN provides more excitation, without the corresponding striatal inhibition. This only stands analytically if $\delta < 1$, and $w_t = w_s$.

Selectivity Selectivity is said to be "hard" if only one channel can be selected at any one time, or "soft" if this can be multiple. From the general equation 2.1

$$\Delta^i = \delta(k_1 X^+ - k_2 \hat{X^+}) - L(\lambda_g, \lambda_e)c_i + k_3$$

(where $\Delta^i = \tilde{a_i^g} - \hat{a^g}$), if the salience of channel k stays fixed, and the others rise, X^+will rise, making it harder to reduce activity on k, i.e. the system is more selective. They determine that the level of STN suppression for a given increase in salience is dependent on how salience affects the $(k_1 X^+ - k_2 \hat{X^+})$ part of Δ^i, so is proportional to changes in X^+. By categorising channels according to their selected status within STN and GPe, they approximate

$$\frac{\delta dX^+}{dc_i} = \frac{\chi_s(i)[w_t + w_g w^-(1-\lambda_e)w_s\chi_q(i)]}{1 + \delta w_g w^- n^s \varphi_s^m}$$

where $\chi_s(i)$ is the characteristic function for membership of S^* ($\chi_s(i) = 1$ iff $i \in S^*$), and $\chi_q(i)$ the membership function for $\wp$ ($\chi_q(i) = 1$ iff $i \in \wp$). After applying various constraints to the weights (argued on the basis that if the activation of their neurons is to be maintained at levels no greater than 1, this places constraints on the input activity, and consequently the weights, particularly the single inhibitory ones of striatum and GPe on GPi (compared to diffuse STN) must also be constrained i.e. w_g and w^- must multiply to 1.

If $w_g w^i = 1$, then $\frac{\delta dX^+}{dc_i} > \frac{w_t}{n}$. In the absence of a GPe -> STN pathway, it is equal to this. i.e. suppression becomes more rapid with respect to increasing salience in the presence of the GPe.

2.4.5 Dopaminergic modulation

To determine the effects dopamine would have on GPi activation, Gurney differentiated $a_i^g$with respect to λ_g and λ_e. They found that so long as a channel was active within the striatum, the action of dopamine on the select path via D1 receptors was to decrease its respective GPi activity. If there was activity within the STN, control-path dopamine acted to increasingly depress the GPi as a whole, i.e. increased promiscuity / decreased Selectivity. These effects were also found to be lost when the input from GPe to GPi was removed - the control pathway lived up to its name.

2.4.6 Relating the Gurney model to schizophrenia

Although there is some variation, the symptoms of schizophrenia are generally clustered into positive (hallucinations, delusions), negative (poverty of action, anhedonia) and disorganised (thought disorder) categories [310]. The division concerning negative and disorganised states has an added relevance when one considers them in the context of selection mechanisms.

A system that was too rigid, requiring high level of salience to select a single channel, would theoretically produce behaviour and cognitions resembling the deficit or negative state. Conversely, systems that were excessively promiscuous would allow inappropriate simultaneous selections, and be unable to maintain single appropriate selection: analogous to the disorganised state. Given its clear, summarised representation of how dopamine modulates BG function and activity - particularly via D2 receptors - the Gurney model may be able to provide insights into how such states can emerge for people with schizophrenia.

D2 receptors are a key target of most efficacious antipsychotic medications, although the advent of atypical drugs and their complex pharmacological profiles implies an inevitably more colourful landscape. Sadly, negative symptoms (which have been cited as a core and early feature of schizophrenia [311]) frequently do not respond to medication [312] raising a question mark as to whether dopamine dysfunction is the key underlying neurobiological issue. Although there is evidence for differences in the BG associated with schizophrenia [313, 30, 314], it may be the case that irregularities in subcortical dopamine transmission are secondary to a root dysfunction.

Chapter 3

Preliminary methods

3.1 Computational modelling-derived predictions

This study used the computational model previously described by Gurney and colleagues [13, 28], and elements of its foundation are described in section 2.4. To summarise its more relevant features; the model is grounded upon empirical anatomical and electrophysiological findings, qualifying it as a useful test bed for in vivo studies. It is an artificial neural network representing the major BG nuclei, and is capable of selecting between different "channels" (analogous to actions) by selectively inhibiting its output nucleus (GP) in response to increases in afferent "drive", or cortical activity. A given action is associated with one unit (representing a population of neurons) within each nucleus, arranged as a series of segregated channels across the architecture of the BG. When a decision must be made between two actions, representative cortical activity drives a competitive process across the network via striatopallidal and reciprocal subthalamopallidal interactions. This results in the selection of the most cortically-active channel through focal inhibition of associated pallidal activity. Gurney et al. define two important characteristics pertaining to selection mechanisms: Selectivity and Competitiveness.

3.1.1 Selectivity

This refers to the proportion of actions that can be selected at once: a system that permits only one action is said to implement hard switching, whereas a more "promiscuous" system operates soft switching. The strictness with which losing channels are suppressed is a result

of tonic inhibition exerted by the pallidus on all channels, which is dynamically scaled by the STN's excitatory projection. Selectivity has therefore been defined as a function of the rate of change of subthalamic activity with respect to the system's inputs, which can be approximated as

$$\frac{dX^+}{dc_i} \approx \frac{w_s}{w_g w^- n}$$

Where X^+ is total subthalamic activity, c_i is the input for action i, n is the number of possible actions, and the w terms describe synaptic strengths between cortex and striatum (w_s), GPe and STN (w_g), and striatum and GPe (w^-) (Gurney et al., 2001b). To assess the effects of explicitly altering Selectivity on model behaviour, the parameter w_g was systematically varied between -1.0 and -0.1 in steps of 0.1, with Competition being held constant with $c_1 = 0.7$ and $c_2 = 0.3$, with Selectivity being inversely related to the strength of pallidosubthalamic connectivity.

3.1.2 Competitiveness

This is the degree to which activity on one channel modulates the suppression of another [28]. This was determined by simulating Competitions of increasing intensity between two channels. At either extreme, a low-level Competition involved cortical values of $c_i = 0.9$ and $c_j = 0.1$, whereas a high-level Competition used values of $c_i = 0.5$ and $c_j = 0.5$. These values can be seen as being analogous to the expected value of each action (the magnitude of an associated reward, multiplied by the probability of receiving it [315]), therefore expected utility was constant for each Competition (the summed expected value of all possible actions). Twenty channels were included in the simulations, and all other parameters were matched to those used by Gurney and colleagues in 2001.

3.1.3 Predicting pallidal activation

3.1.3.1 The main effect of Selectivity

Keeping the level of Competition between channels constant, it was found that decreasing the weighting of pallidosubthalamic connectivity, and hence increasing Selectivity, had the effect of raising the total activity of the pallidus (figure 3.1). As the coupling between pallidus and STN decreased, so focal selection-related activity within pallidus had an elevated effect over STN [300]. By virtue of STN's diffuse projection back onto the pallidus, these focal changes were

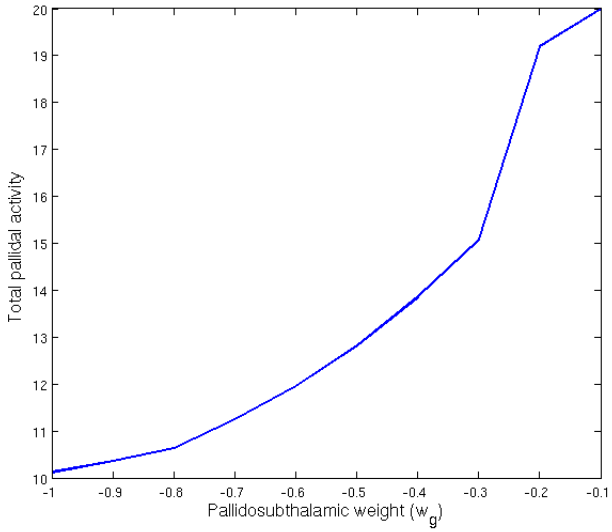


Figure 3.1: Simulated predictions of the BG-mediated model of action selection. Simulated selection-induced pallidus activation with respect to Selectivity, summed across all channels, in arbitrary units of activity. This was varied according to the strength of influence between pallidus and STN (w_g).

transmitted over all simulated channels, resulting in the increase of total pallidal activation. These simulations underwrote the first hypothesis:

Hypothesis 1 *If a disinhibitory selection mechanism as described by the Gurney model was indeed operating within the BG, average pallidal activity would become elevated as the demand for Selectivity increased.*

3.1.3.2 The main effect of Competitiveness

It was also of interest to determine what effect Competitiveness had on pallidal activation, separated from that of Selectivity. The model's Selectivity was held constant with $w_g = 0.6$, and Competitiveness varied as described above. According to the model [28], the change in total pallidal activation $d\sum_k a_k^g$with respect to Competition $c_j - c_i$ is:

$$\frac{d\sum_k a_k^g}{d(c_j - c_i)} = w^- w_s (c_j - c_i)\left[(1 + \lambda_g) + w_e(1 - \lambda_e)\right]$$

Where w^- is striatopallidal connectivity, w_s is corticostriatal connectivity, w_e is external-to-internal pallidal connectivity, and λ_g and λ_e are tonic dopamine levels for the striatal SEL and

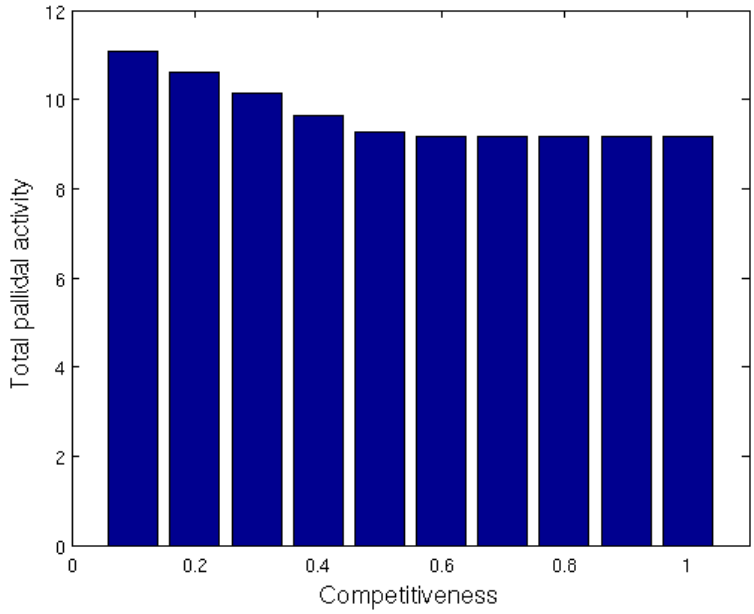


Figure 3.2: The effect of increasing Competition on simulated pallidal activity, summed over all channels of both the internal and external segments. A Competition value of 0.1 translates to $c_1 = 0.9$, $c_2 = 0.1$, whereas Competition = 1.0 represents $c_1 = c_2 = 0.5$.

CONT pathways respectively. In this situation where all factors except the saliences driving the decision were constant, a second hypothesis can be made:

Hypothesis 2 *Total pallidal activity would be expected to decrease with increasing Competition, i.e. as* $c_j - c_i \to 0$.

The simulation concurred with this (figure 3.2).

3.1.3.3 The interaction between Selectivity and Competitiveness

To consider the interaction between Selectivity and Competitiveness, two corticosubthalamic weights (w_g) were taken to represent high (0.9) and low (0.6) Selectivity, and varied the Competition between two channels according to their respective expected values. When pallidal activity was displayed in terms of a High > Low Selectivity contrast over the full Competition range, it was found that inhibitory-related activity dipped at intermediate values (figure 3.3). This provided the third and somewhat counter-intuitive hypothesis:

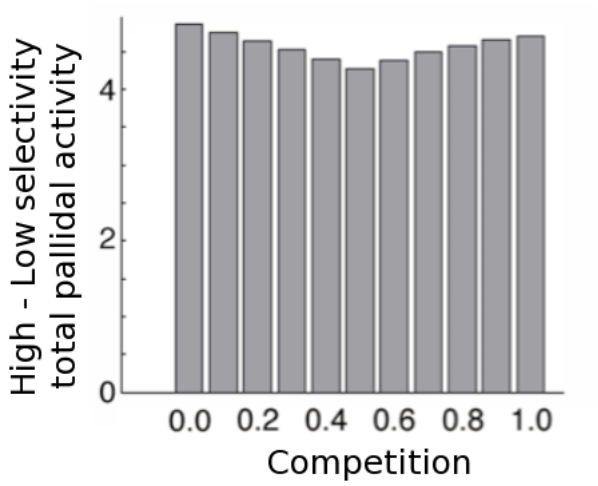


Figure 3.3: Simulated predictions of the BG-mediated model of action selection. A simulation of Selectivity's interaction across the range of Competitiveness values. Here, a Competition value of 0.0 means $c_1 = 1.0$ and $c_2 = 0.0$, whereas a Competition value of 1.0 means $c_1 = c_2 = 0.5$.

Hypothesis 3 *The High > Low Selectivity contrast within pallidus would be diminished at intermediate Competition levels.*

3.1.4 Predicting striatal activation

3.1.4.1 The main effect of Selectivity

Selectivity was modulated according to GPe-STN connectivity strength (w_g), and as expected, this had no effect on summed striatal activation.

3.1.4.2 The main effect of Competitiveness

The principle determinant of striatal activity was expected to be the degree of Competition between active channels. From the description of the model [28], Competition between two channels i and j will affect total striatal activity $\sum_k x_k^{g-}$ (considering the SEL pathway) according to:

$$\frac{d\sum_k x_k^{g-}}{d(c_i - c_j)} = w_s(1 + \lambda_g)(c_i - c_j)$$

Where w_s is corticostriatal weight, and λ_g the level of tonic dopamine. It would therefore

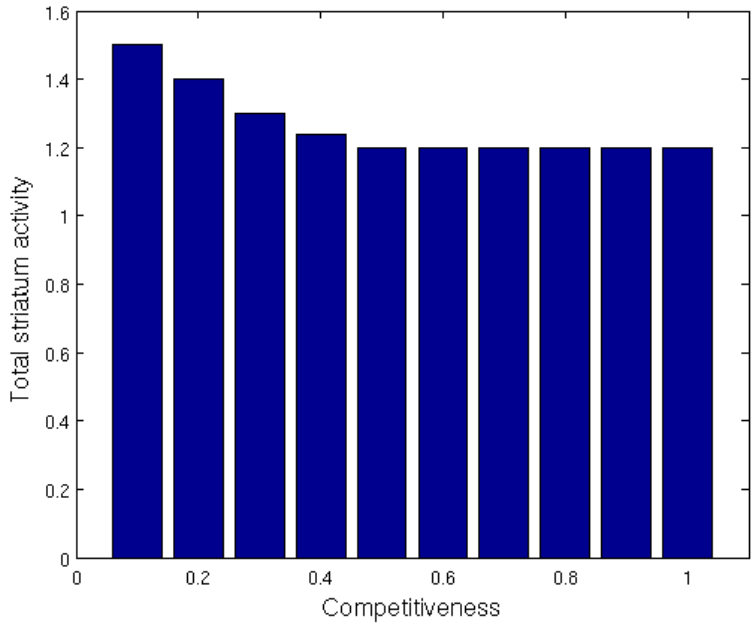


Figure 3.4: The effect of increasing Competition on simulated striatal activity, summed over all channels of both the SEL (D1-direct-select) and CONT (D2-indirect-control) pathways. A Competition value of 0.1 translates to $c_1 = 0.9$, $c_2 = 0.1$, whereas Competition = 1.0 represents $c_1 = c_2 = 0.5$.

be expected that a lower-level Competition (where the differences in competing salience was greatest) would elicit the greatest changes in striatal activation. This was indeed born out it simulations (figure 3.4). As for the pallidus, striatal activity was highest when Competition was at its lowest, i.e. when the decision was most clear-cut. The response to increasing Competition saturated at intermediate levels. This produced a fourth hypothesis:

Hypothesis 4 *Striatal activity would peak at low levels of Competition, with intermediate and high levels being broadly equivalent.*

3.1.4.3 The interaction between Selectivity and Competitiveness

Selectivity was a property modelled to exert its influence solely on the pallidus, therefore no Selectivity-Competitiveness interaction was present in the striatum.

3.2 Paradigm Development

The next step was to ascertain how closely these model predictions agreed with in vivo data using event-related fMRI. This necessitated the design and piloting of a novel decision-making task, that manipulated the properties of Selectivity and Competition. The task needed to satisfy a number of constraints:

1. To maximise BG engagement.
2. To vary Selectivity and Competition in a way that permitted clear inference of their effects.
3. Stimuli and motor responses must be balanced across conditions.
4. To be implementable within the MR environment.
5. As the intention was to use this task during a second study involving patients with schizophrenia, it needed to have limited working memory requirements, and be robust to the underlying perceptual abnormalities that have been observed in patient groups. Before fMRI scanning was initiated, a brief behavioural pilot involving six healthy volunteers and two people with schizophrenia was carried out, to establish whether the paradigm and its associated procedures (see sections 3.3 and 3.4) were going to be practical.

3.2.1 Considering existing paradigms

3.2.1.1 Response inhibition

A good source of possible task designs to consider were those manipulating response inhibition, which have been found to induce BG activation [29, 316, 32]. A simple example would be the GO/NOGO paradigm, which leads the participant into a pattern of repeated responses, but with the occasional oddball trial where a response is supposed to be withheld [29, 317]. Although simple and previously successful, this type of task will not be suitable here: the binary choice of withholding or providing a response provides no means to manipulate Selectivity, and a means to parametrically implement Competition is not clear. In addition, events where responses are successfully withheld are motorically imbalanced compared to trials where a movement has occurred, clouding subsequent inference.

A more complex implementation of the basic GO/NOGO task is the Stop Signal task, with a good example being that of Aron and Poldrack [32, 318]. Here, the participant is again

cued to provide a response, but this time a second cue is occasionally presented very shortly afterwards, signalling that the participant should try to withhold this already initiated response. By varying the delay between the two cues, you can estimate the participant's stop signal reaction time (SSRT), with long SSRTs indicating a strong capacity to inhibit responses. This has the strength of being able to parametrically vary the difficulty of the task, though again it suffers in that the result of a successful inhibition is the complete absence of a response. The nature of the competition is also not quite appropriate: this is not a direct assessment and choice between two explicit alternatives - an established prepotent selection is simply being overridden at short notice.

An alternative that goes some way to establishing a direct competition are the Stroop [319], Simon [320] and Eriksen flanker [321] tasks (for examples of their use in fMRI, see [322, 323, 324]). These tasks manipulate conflict monitoring by using task-relevant distractors whose influence must be suppressed if the participant is to perform appropriately. During the Stroop task, the participant is shown the name of a colour, but it may be printed in a discordant ink - the task is to name the ink colour. The Simon task exploits the phenomenon that responses are faster if they share their relative spatial location with their cueing stimulus, for example, if the cue for the left-handed responses appears on the left side of the screen. The Eriksen flanker task displays a row of arrows pointing left or right - the subject must respond according to the direction of the central arrow, though interference is introduced by the flanking arrows pointing in a congruous or incongruous direction. In all these tasks, the competition is indirect, and categorically driven by distractor-lead interference rather than direct competition, therefore it is unclear how these would be represented within the channel architecture of the Gurney model.

3.2.1.2 Decision-making

Tasks that involve a more clear-cut competition between two salient responses are those examining decision-making. The means by which it becomes difficult to choose between the two can vary dramatically, ranging from problems of perceptual ambiguity [325], to tasks where beneficial outcomes are probabilistic, and effective action depends on an assessment of risk and value [326].

The decision-making process is frequently conceptualised as consisting of several stages [327], not necessarily discrete and serial, but presumably essential:

- Sensory input: the reception and transformation of sensory stimulation relevant to the decision. Depending on the paradigm, this may be the frequency of flutter of a tactile

stimulus [328], the dominant flow direction of a visual field of moving dots [329], or the appreciation of a more categorical cue.

- Sensory accumulation: As visual input is assessed by primary sensory areas, this will accumulate over time to support or refute a particular course of action, depending on the context of the task. When considered in primates using electrophysiological methods, areas described as evidence accumulators show a gradual increase of activity, that remains elevated until a decision is made [327]. The rate at which this occurs is reduced for more difficult decisions, and the region/pattern of activation does not necessarily correlate with the final response made, suggesting this is still emphasising the sensory aspect of the process, and not directly relating to a motor response quite yet.
- Decision transformation: Previous studies have also identified areas that are very similar to the evidence accumulators, except they are one step closer to the motor response itself, in that they apparently perform a subtraction between the activity levels ”voting” for particular sensory appraisals, allowing one to emerge as the categorically ”selected” action [330, 331, 332].
- Action selection: This final element involves the final settling on a particular course of action within the given context, and initiates its execution. It is of course the principle concern of this work.

Other proposed concurrent processes are those of uncertainty/difficulty appreciation [333, 334], which may drive an additional recruitment of attentional resources [335]; and a means to monitor performance and impose corrective signals on other regions involved in the decision-making process.

A couple of exemplar tasks demonstrate the parametric flexibility of perceptual decision-making tasks: Shadlen and Newsome’s electrophysiological investigations in non-human primates presented their participants with collections of randomly moving dots, though with an overall directional bias. The monkeys were trained to move a joystick according to their judgement of that dominant direction [336, 329, 337, 327]. Both the clarity of that direction could be varied, and the range of possible responses was continuous. The face vs house tasks of Heekeren et al. presented humans with images that were cross-faded blends of human faces and house facades. Participants were asked to categorise them appropriately during fMRI [325, 330, 335]. As well as allowing for a parametric manipulation of this direct and explicit competition, the neural substrates encoding houses and faces were clearly differentiable even using the spatial resolution of fMRI.

A key elaboration of the decision-making process involves the consideration of possible rewarding outcomes, and the assessment of a stimulus' subjective value [326, 292, 338]. Such tasks usually involve the learning of cue-action-outcome relationships for which there may be a variation of risk though probabilistic outcomes [339], altering the quantity and valence of that outcome [19], as well as variations in the time at which that it is delivered [290].

3.2.1.3 Proposed paradigm design

In considering which paradigm framework might be more suitable for this work, it is clear that both perceptual decision-making and the more "neuroeconomic" approach offer a flexible and powerful means to parametrically manipulate competition between actions, through either sensory blending/ambiguity, or the learned probability of receiving a reward when cued to perform an action. Therefore the overriding concern has to be the task's suitability for use with patients. The neuroeconomic tasks are effectively dependent on instrumental conditioning - the successful association of a cue with an action-outcome relationship. It is well documented that associative function is disturbed in patients with schizophrenia [209], which is of course of profound interest, but is not the emphasis of this work. Value is also highly subjective, and more challenging to balance both within and between patient and control groups. Using a paradigm that is dependent on successful associative value learning is therefore unwise.

A perceptual decision-making paradigm will therefore be selected. This will only involve a small degree of cue-response learning, which will be invariable over the implementation of the task, and be thoroughly established during a training session before scanning begins. The decision itself is therefore much purer: all the necessary information will be instantly available in the currently presented stimulus. As in the house vs face tasks described earlier, the presented stimulus will demand a categorical decision - a 2-alternative forced choice paradigm. The nature of the mechanical end effectors used to signal this decision, i.e. the action that each perceptual alternative is associated with, will underlie the means by which Selectivity is manipulated. The next question to address is the sensory means by which Competition will be altered.

3.2.2 Manipulating Competition

Continuous and well-defined perceptual ambiguity must be an intrinsic quality of the cues used within the task. The obvious choice would be to implement Heekeren et al's face vs house task:

however, it is known that the perception of human faces is not an entirely unconfounded process in patients with schizophrenia [340, 341, 342]. Several alternative options are available, although the intended subject populations, dopaminergic interventions and the fMRI environment itself impose certain constraints.

3.2.2.1 Considering sensory modalities

The first stage of cue selection involves choosing an appropriate sensory modality, the prime contenders being auditory and visual. The loud and repetitive noise of MRI is not conducive to using sound as an experimental stimulus, and incorporating ambiguity would be a challenge. If pitch were to be varied, how would one define the absolutes against which a subject is supposed to compare intermediate tones? Should they be remembered, or presented in tandem with the experimental cue? Such cues would be highly vulnerable to individual differences, the nature of which may be impractical to quantify. It is therefore logical and pragmatic to favour visual cues.

Here though, further potential confounds must be confronted. The first is to determine those components of visual processing that are altered in the schizophrenic population. The second is assess what role dopamine has to play in mediating visual function.

3.2.2.2 Elements of low-level visual processing

Generally, this can be described as proceeding along 2 main routes. The magnocellular “where” pathway (M) is primarily responsible for reporting rapid-transient, low-luminance contrast, low spatial frequency stimuli, and is particularly important for motion perception and spatial localisation. The parvocellular “what” pathway (P) is concerned with static higher contrast and higher spatial frequency stimuli, to pick out key details for the purpose of stimulus identification. It also is the sole route for colour processing [343]. Both originate from specific retinal ganglion cells, project to segregated layers of the lateral geniculate nucleus, and synapse cortically within different elements of the multilayered blob-interblob system of V1. From here, the M path takes a predominantly dorsal route to parieto-occipital cortex, whereas the P path heads ventrally to temporoccipital areas.

3.2.2.3 Selecting the least compromised visual element

Schizophrenia’s seemingly all-pervasive influence over the brain has not left low-level visual processing unaffected. The M pathway has proved the most fruitful avenue for malfunction in

this regard, with motion perception [344], spatial localisation [345, 346], low spatial frequency discrimination [347, 348] and low-contrast sensitivity [347, 349] all being adversely affected. M pathway deficits have been replicated in typical, atypical and neuroleptic naive subjects, as well as unaffected siblings [350, 351]. They seem to be particularly striking in patients with predominantly negative symptoms [352, 348], which could impinge on future interest in this patient subgroup. Therefore it seems wise to exclude stimuli that are are heavily dependent on the M pathway.

The evidence for a deficient P pathway is less convincing, indeed there is a lot to say that perception of higher spatial frequencies are unaffected [349, 347]. However, P-type deficits have been observed elsewhere [344, 352]. This ambiguity could prove to be a future downfall. As quantitative assessment of each subject's P pathway performance would prove to be an experiment in itself, it would be prudent to pursue a different visual avenue. In order to avoid any reliance on spatial information, it seems reasonable to focus on the use of colour.

3.2.2.4 Colour and schizophrenia

Of course, colour will only prove useful if its perception remains relatively unaltered in patients. Specific investigation is a little thin on the ground: two papers have focused on it directly [353, 354], which unfortunately offer conflicting findings. Heim and Morgner used both the Lanthony 15 Desaturated Hue and the Farnsworth 15D tests on 50 schizophrenic patients (no details provided). Such tests provide dichotomous evaluation of mild acquired and strong congenital colour deficits respectively. They found that 72% of both men and women were apparently affected (compared to 8% of men and 0.5% of women in the general population). However, these tests are not extensive, and given the unknown nature of the sample's medication and symptomatic states, it is conceivable that such a dramatic finding can be at least partially explained by means not relating to colour perception.

A more comprehensive study by Shuwairi et al. added three more tests to the battery: the City University, Farnsworth-Munsell 100-Hue, and the Lanthony New Color tests, the latter two being more extensive and providing qualitative results. 16 chronic schizophrenics (mean age 45.4, range 30-56) completed the tests. Although their overall error rate was significantly higher than that of controls, there was no skew towards a particular red-green or blue-yellow axis. This suggests that the errors made were a consequence of more general executive deficits [355] rather than anything colour-specific, though others have found no correlation between neuropsychological and visual-perceptual measures [356]. Error rate was not correlated with

either the positive or negative scores of the PANSS (a finding that has been replicated with spatial perception [355]), reinforcing the idea that colour perception is not directly perturbed in schizophrenia. The failure of this more thorough study to replicate Heim and Morgner's findings, as well as there being no correlation with subject-specific factors, lends credence to the idea that colour perception may be only marginally affected in a subset of schizophrenics.

Indeed, although not directly indicative of colour discriminatory ability in itself, Keri et al. have found that the size of a minimal detectable offset between two dots is unaffected if isoluminant coloured versus black and white dots are used [349, 356]. This failure to find a deficit in what was a uniquely colour and P-pathway task is a further positive sign that colour perception within most schizophrenic patients is within normal bounds. Nevertheless, all subjects will have to be screened for normal colour perception on the basis of a clinical test such as the Farnsworth-Munsell 100-Hue, Lanthony's 40-Hue, Roth's 28-Hue, Farnsworth 15D or Lanthony Desaturated 15-Hue (in order of complexity, most complex first). In the case of the schizophrenic sample, our rejection/acceptance ratio would in itself be a result of interest, given its brief coverage in the literature.

3.2.2.5 The role of retinal dopamine in colour perception

The final stage in asserting the validity of colour cues entails exploring dopamine's influence over early visual processing. A good review of this is provided by Witkovsky [357].

Light adaption Within the retina itself, dopamine is emitted by a subgroup of amacrine cells in the inner plexiform layer [358], receiving input primarily from OFF bipolar cells, and in turn output to elements of the rod system. This is suggestive of a role in light/dark adaption, for which there is evidence [359, 360]. D1-mediated dopamine acts to decouple horizontal cells that would otherwise act in concert via gap junctions [361, 362], which is again implicated as a means of mediating responsiveness under varying light conditions [363]. As dopamine is elevated during daylight hours [364], it is conceivable that a drop in its efficacy would result in a switch to rod-driven vision, thus compromising colour perception.

Photoreceptor modulation Members of the D2-type family are known to occur on both rods and cones [365], the action of which is to reduce Ca^{2+} influx, and therefore subsequent glutamate release. However, it has been reported that the most predominant member of the D2 family present in the rat retina is D4 [366]. Within the BG, D4 has been localised presynaptically on

striatonigral and striatopallidal cells within globus pallidus (internum and externum), and substantia nigra reticulatum [367]. Presynaptic action of dopamine within the BG's output nuclei is outside the remit of the Gurney model, and as antipsychotic medication has a lower affinity for D4 receptors [368], it will be assumed that retinal confounds will not be an especial concern during this work.

Centre-surround balance The most obvious means of altering colour perception would be to affect the centre-surround balance of bipolars and their afferent ganglion cells. Given the strong influence of OFF bipolar cell input, the means to implement a centre-surround-driven pattern of dopamine release is in place. However, our concern is not what mediates dopamine emission, but its subsequent effects upon reception at afferent cells. D1-type receptors have been localised to bipolar, horizontal, amacrine and ganglion cells [369, 370], the action of which varies the extent of the surround field and thus alters spatial contrast perception [361, 371]. However, the very low spatial resolution of our proposed cues would preclude such an influence.

Dopamine also modifies ganglionic surround field activity in addition to extent. Dopamine is known to inhibit horizontal cell activity [372], which in turn increases surround field activity [373]. Indeed, D1 antagonist SCH 23390 has been shown to reduce OFF-centre ganglion cell surround activity in light adapted conditions, leaving centre activation unaffected [359]. It seems this is an affect which would be relevant to colour perception, as D1 antagonism is most effective in light adapted conditions. D1 agonism on the other hand is generally more affective in the dark [374], so is less of a concern.

This all indicates that dopamine would affect colour discrimination. However, it is again apparent that spatial rather than chromatic factors are more vulnerable, as M-type ganglions require relatively smaller current injections into their associated horizontal cells to elicit activational changes [373]. Colour therefore still resembles the safer choice, though is clearly not immune to dopamine modulation.

Effects of dopamine receptor manipulation As alluded to earlier, an overall reduction in the efficacy of retinal dopamine could lead to more emphasis being placed on rod-driven vision, at the expense of colour perception. Parkinson's disease (PD) reduces the quantity of dopamine in retinal amacrine cells. The resultant hypodopaminergic state, similar to those undergoing cocaine withdrawal, produce deficits in colour discrimination [375, 376], with some showing a trend towards blue-yellow deficits [377], and others red-green [378]. These deficits are im-

proved after the administration of L-dopa [379]. However, these are extreme cases. Thanks to spontaneous activity and a sustained extracellular presence, retinal dopamine concentrations are normally high [380]. It would theoretically take severe antagonist manipulations to elicit similar effects, which wouldn't be contemplated here. It should also be noted that clinical tests (Farnsworth-Munsell 100-Hue, Lanthony 15-D) results are correlated with motor control performance [381]. This suggests that colour deficiencies of PD, and therefore of dopamine depletion in general, may be overestimated. On the other hand, deficits are still found when motor-dependent clinical tests are abandoned in favour of staircase psychophysical methods [378], so the effect seems reliable.

Exploring the specific effects of receptor agonism/antagonism should prove more illuminating for our purposes. Concerning spatial contrast resolution in schizophrenia, some have found that atypical neuroleptics have no effect, whereas typicals decrease performance, and neuroleptic naive patients actually show improved performance over healthy controls [382]. Indeed, specific dopamine receptor manipulation is widely reported to result in altered performance, particularly in the M pathway [359, 383, 384, 385, 386, 387, 388]. This indicates an effect on surround field extent rather than absolute activation level.

With regards to colour, specific investigations are sparse, with most data coming from non-mammalian studies. D1 agonism via apomorphine in PD patients offers no improvement in colour discrimination [389]. D1 antagonist intravitreal injections have been shown to reduce red-green discrimination in goldfish [390], although the effects of administering antagonists in this way will presumably be more acute than when delivered intravenously. Shuwairi et al. found no correlation between neuroleptic dosage (typical/atypical nature not reported) and colour test error rate [354], and I can find nothing specific concerning D2 agonism and colour perception.

Overall, it seems that dopamine manipulations preferentially alter spatial rather than colour perception.

Dopamine and the lateral geniculate nucleus For the sake of completion, the final stage of early visual processing will be briefly considered. Dopamine has a relatively small influence over LGN anatomically and physiologically [391], that is not responsible for mediating centre-surround effects [392].

3.2.2.6 Summary

A clear picture of the effects dopamine has on colour perception has not yet been painted by the literature. The evidence that exists points towards dopamine being an effective modulator of spatial, particularly M pathway-related functionality. There is no strong evidence that colour perception is altered in schizophrenia, therefore given the ease with which colour cues can be parametrically varied, this will be the means by which perceptual ambiguity and hence Competition is varied in the decision-making task to be developed during this work.

There does however remain a significant theoretical risk of greater variability in colour perception in our proposed patient group, and the lack of literature cannot be taken as a lack of effect. Given that clinical tests are believed to provide only semi-quantitative results [393], the best approach would be to perform psychometric colour tests to assess each participant's perceptual biases, and generate a set of cues that conformed to these parameters. A clinical test would still be used to establish that the participant's colour perception was within normal bounds.

To conclude, colour is the best option available, given that it remains relatively unaffected by both psychotic pathology and dopamine modulation (compared to other visual modalities), and there are effective means in place to quantify individual differences in its perception.

3.2.3 Manipulating Selectivity

Whereas basic action Competition will be derived from sensory ambiguity, Selectivity is more concerned with the action end effectors, and their degree of mutual compatibility. The need to suppress a competing action will vary according to whether or not those actions could be simultaneously performed: if there is no inherent conflict, the selection mechanism could remain promiscuous, allowing dual selection. If however the actions have a degree of mechanical interdependence or even direct opposition, promiscuity will have to be minimised in order for any action to be selected. However, different end effectors will have different implications for the amount of expected BG activation. For a given task, these nuclei have been repeatedly found to produce a signal that is around one third of that found in cortex [394, 395]. This is in part due to a higher level of variance - during finger tapping, the BG will show no activation at all for 10-30% of trials [395]. The apparent weakness of the BG signal may be partly due to the dynamics of the haemodynamic response in this region: Moritz and colleagues found that a better fit was found for functional data if it was modelled with more transitory haemodynamic

response functions (HRFs), rather than as sustained blocks [396]. A briefer, more fleeting response may account for the approximate 66% drop in detectable signal.

Bearing this in mind, the next decision that needs to be made is to identify the most appropriate biomechanical system to be used for this paradigm.

3.2.3.1 Considering end effectors

The oculomotor saccade system The ideal scenario would involve two pairs of responses: one which is incompatible, and one which is mutually independent. One end effector which satisfies these demands in a well-balanced way is the oculomotor saccadic system: by elevating vs depressing the eyes, and moving them left vs right, two mutually incompatible and independent response pairs are available. The muscles subserving such movements share a degree of specificity. For example, the inferior and medial rectus muscles are only involved in the abduction and adduction of the eyeball, whereas the superior and lateral recti play a part in both lateral-medial and elevatory-depressive movements [397]. However, the final end act of moving the eyes left vs right, or up vs down, is certainly incompatible - successful execution of either movement implies strong suppression of the competing alternative. Therefore this would seem to offer an ideal conceptual means to manipulate Selectivity.

Tasks requiring saccadic responses are known to be effective activators of the BG [395, 398, 399]. However the primary output nucleus for oculomotor actions is not the globus pallidus, but the substantia nigra pars reticulata [269], a substantially smaller structure. As the intention is to test hypotheses of activation concerning the output nuclei, it would not be optimal if such tests were dependent on imaging a spatially small region with presumably high inter-subject variability. In addition, saccadic dysfunction is widely reported amongst patients with schizophrenia [400], implying this is an inherently compromised response mechanism, casting a shadow over proposed group comparisons.

The musculoskeletal system Here one could choose to set up decisions between end effectors at a number of scales, for example, short reaching motions involving the whole arm have been shown to produce strong BG signals [401, 314]. However, it was decided that movements around larger joints was ill-advised during scanning, due to the increased risk of motion artifact. Instead, the emphasis was placed on the smaller joints of the hand.

As each Competition would be resolved with a categorical decision, a series of well-defined motor acts were required. Most fMRI tasks use button presses, that is, brief flexions of the fin-

gers. These are of course essentially independent of each other - simultaneous button presses are perfectly possible. This presents a problem if we wish to elicit decisions with a higher level of Selectivity. However, the one action that cannot be performed during, say, index finger flexion, is index finger extension. By including finger extensions within the gamut of possible motor actions, Selectivity could be varied in a way that was entirely counterbalanced across the fingers. Therefore, Selectivity was manipulated by evoking decisions involving finger flexions and extensions. For a given finger, deciding between flexion and extension requires the selection mechanism to have a high level of Selectivity, where dual selection is implicitly made impossible. Lower Selectivity decisions are those where the competing actions are on entirely different fingers: here the relative mechanical independence means dual selection is in principle permissible - the inherent demand for Selectivity is lower. Using flexions and extensions of different digits offers a potential improvement over previous studies of response inhibition, where successful suppression involves the complete absence or withdrawal of a response. This makes subsequent contrasts of interest, and hence inference, somewhat imbalanced. Here you are always able to compare like with like, with the contrast being much more tightly controlled around the manipulation of Selectivity, with no confounding covariates. It is however quite likely that these four movements cannot be considered entirely equal - a strength of this design is that each specific movement is equally distributed across the levels of Selectivity: there is no systematic bias.

Using these motor responses also enables one to detect those moments when the competing alternative was not entirely suppressed:

- In the case of high-Selectivity decisions between flexion and extension of the same finger, a failure of suppression would present itself as a significantly slowed or hesitant movement.
- For low-Selectivity decisions, we may observed significant movement on the alternative finger.

A means to detect such movements will be discussed in section 3.4.

3.2.3.2 Hand restraint

Using finger extensions as well as flexions presented another technical challenge - how could participants reliably and comfortably perform what is a somewhat unusual manoeuvre? To optimise these factors, a hand restraint was developed and designed. This consisted of:

- An acrylic beam that rested on the medial surface of the hand's proximal phalanges. This provided a pivot around which the participant could flex and extend their index and middle fingers.
- Two elasticated bands that wrapped separately around the index and middle digits, securing them to the beam. These three functions: to provide a resistance against which the participant had to work, and hence boost BG activation [402]; to passively return the participant's finger to a resting position after movement execution; and to set that resting position to be half-way between flexion and extension, maximising the range of movement for each.
- A padded acrylic board that was anchored to the posterior surface of the participants hand and forearm by Velcro straps. This provided comfortable support for the participants wrist, minimising extraneous movements, while allowing them to remain relaxed for the duration of the paradigm.

3.2.4 Paradigm design

So far it has been established that this new paradigm will use blended colour cues as stimuli, and single flexions or extensions of the index and middle fingers as motor responses. This was implemented as a decision-making task in the following way: each of the four motor responses: index flexion, index extension, middle flexion and middle extension, was associated with a ”pure” colour cue, as depicted in figure 3.5a. These were chosen to correspond approximately to red, yellow, green and blue, and were equidistant in CIEluv color space (white=D-50, observer=1964, 10'). Decisions of varying Competition were set up by presenting blends of these four pure cues, and asking the participant to perform the associated movement for the pure cue to which it was closest. This was varied over three levels: low, medium and high. Selectivity varied over two levels as a function of the two motor responses that were called into conflict: decisions between, say, index flexion and middle flexion, required a lower level of Selectivity than, say index flexion vs index extension (figure 3.5b).

Trials were presented with E-Prime (http://www.pstnet.com) on a desktop computer. For each trial, a square color stimulus was displayed centrally on a monitor for 150ms, cueing the subject to execute the finger movement corresponding to the square's color. A response period followed for 2650ms, during which a central fixation cross replaced the stimulus, giving an in-

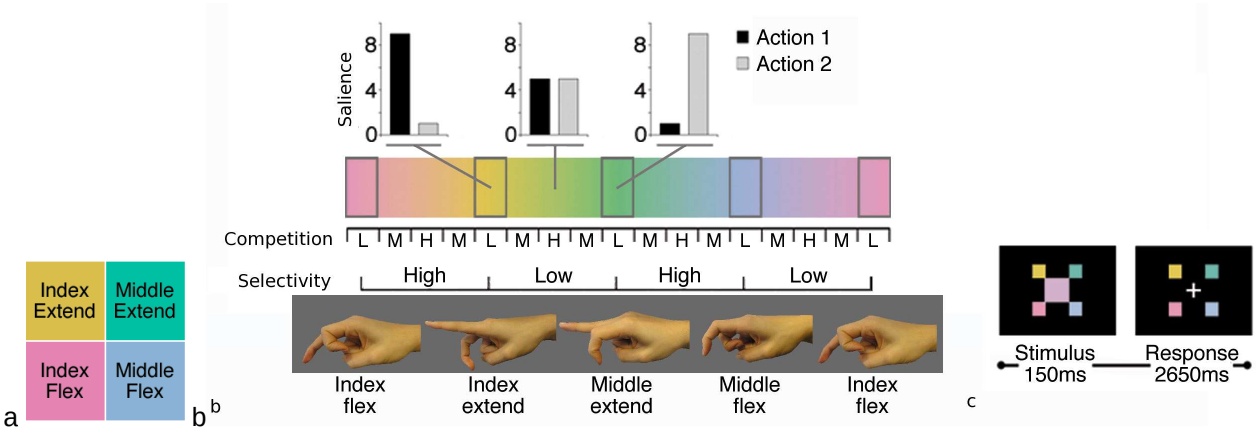


Figure 3.5: The colour-cued decision-making task. (a) Participants were trained to associate four finger movements with four ”pure” colour cues. The factor of Selectivity was determined by the level of mechanical interference between the two responses evoked by each presented stimulus.(b) Here, the pure cues are displayed within grey rectangles. Stimuli were drawn from equidistant points along the blended colour spectrum between two pure cues. Competition was a function of how close each stimulus was to the perceptual midpoint between two cues. The graphs display examples of the saliences of the competing actions. Motor parameters and expected utility (that is, the summed salience of each trial) remained constant throughout. (c) The timecourse of a single trial. Samples of each of the four pure colour cues remained onscreen at all times, to permit the participant to make a comparison, and to indicate the general direct of the correct movement - for example, index extension was on the left side, and up. A stimulus was briefly presented amongst these reference samples, and the participant asked to respond as quickly as possible within a timed window.

terstimulus intervan (ISI) of 2800ms. At all times during the experiment the central cue was flanked by small squares of the primary cues in those corners corresponding to their associated motor responses' direction of movement (for example, the cue associated with index flexion was shown bottom-left, and middle extension top-right). The reference panels' proximal location to the stimuli limited the need to make saccades.

The resultant 3 x 2 factorial design produced six experimental conditions of interest. Two additional control conditions were also included: a low-level visual control condition where checkerboards were presented instead of a blended colour stimuli, for which the subject was asked not to respond; and a higher-level control condition involving those trials for which a pure colour was presented - these induced zero Competition, and hence zero variation in Selectivity demands, whilst still requiring a colour judgement and motor response. 12 trials were presented for each of these conditions, producing 96 trials per run. Each participant was asked to complete four runs within the scanner.

Every motor response was counterbalanced across each of the cells of this 3 x 2 factorial design, removing variability due to the difficulties in performing each given movement. Also, the saliences of each decision summed over the competing actions was always equal to one, meaning that individual trials did not vary systematically in their total motivational value.

Such investigations are contingent on the ability to evoke subtle graduations in Competition between alternative responses. This will be achieved by associating "primary" (red, yellow, green and blue) colour cues with a series of motor responses. By presenting intermediate blends of these colours, and asking the subject to respond according to which primary they believe it is closest to, continuous direct Competition will be evoked for each decision.

3.2.4.1 Optimising basal ganglia activation

Optimised event-related design The observation that BG fMRI signals are particularly fleeting [396] suggests that event-related designs may prove more appropriate than traditional blocked approaches. Although blocked designs have been used successfully in the past [395, 403, 402, 399], event-related designs are also capable of detecting differential BG activation, and have the additional advantage of opening up the flexibility of the paradigm to be used [396, 324, 404, 32]. In this case, it would be feasible to present the six experimental conditions as blocks, though this may lead to a degree of automaticity in the participants, especially for the low and high Competition trials where they may be prone to making rapid, flippant responses. In addition, event-related designs allow you to isolate trials of a more spurious or

unplanned nature - in this case, where the movement expressed was compatible with a failure of suppression, or if no response was provided at all. Such events would otherwise cloud an experimental block, compromising subsequent inferences.

However, this makes optimising the power of detection for each of the contrasts of interest quite challenging. As well as comparing each experimental condition to the visual and motor controls, the main effects of Competition and Selectivity, and the interaction between them, also needed to be detected. This multivariate problem requires a more algorithmic approach: in this case, it was solved using the genetic algorithm developed by Wager and Nichols [405]. This generates random trial orders, interspersed with moments of rest, which are used to construct HRF-convolved design matrices at the resolution of the desired scan volume repitition time (TR). These are subject to the high and low-pass filters one intends to apply to the real data once acquired, and the results are evaluated according to several fitness criteria: how effectively counterbalanced the trials are (equalising the probability that one trial type follows that of any other); the power with which the specified contrast of interests can be estimated (where specified contrasts can be weighted to signify their priority); the efficiency with which the shape of the haemodynamic response can be detected; and how strictly the specified desired number of trials is adhered to. Those matrices with the highest fitness levels are then selected and ”interbred”, undergoing partial crossover of their trial orders. This process continues iteratively until no further optimisation of fitness is produced.

For this paradigm, the four fitness parameters were weighted as:

- Counterbalancing: 0.6
- Contrast estimation efficiency: 1.0
- HRF estimation efficiency: 1.0
- Trial frequency maintenance: 1.0

The simple experimental condition versus control contrasts were each weighted at 1.0, whereas the main effects and interactions were 3.0. An different optimised trial order was produced for every participant.

Real-life contingencies It has been shown that BG activation is enhanced when the actions performed have some motivational significance attached to them [406, 407, 408]. To enhance this phenomenon, participants were informed that at the end of the experiment, three of their responses would be selected at random, and they would receive £5 for every correct answer.

This method was chosen over, say, providing a small amount for every correct answer, as it was assumed the relative motivation of accruing multiple small amounts is not as great. However, for ethical reasons, each participant was in fact rewarded with the same value of £10, regardless of their performance. This was to ensure everyone felt they had successfully performed the task, whilst staying within budget.

Motor parameters As mentioned earlier, a hand restraint provided some physical resistance to the participant's movement, requiring greater deviation in the rate of change of force, which is known to enhance pallidal activity [402].

Self- versus externally-initiated action Whether or not an action is self-initiated versus externally-paced also impacts on BG activity, with self-direction having an enhancing effect [403, 409, 410]. In this paradigm, each decision is directed by the presentation of a visual stimulus, but the decision itself emerges following the participant's internal resolution of the matter. The timing of responses is encouraged to be as fast as possible, but again there is no external metronome-like pacing. Indeed, as jitter will be introduced in between trials to further optimise the detected blood oxygenation level dependent (BOLD) signal, there shall be a minimal sense of the paradigm having a set pace.

fMRI scanning protocol Due to the transient, non-sustained nature of BG HRF dynamics [396], contrast detection power would increase with a decreasing TR, therefore the scanning protocol was optimised to bring this down as far as possible. The field of view was set to 22cm (matrix 64 x 64), and 28 4mm slices acquired per volume - this was just sufficient to cover the majority of cortex (and of course the entire BG), but sacrificed the cerebellum. This allowed the TR to be brought down to 2200ms. Gradient echo-planer images were acquired axially: to reduce the need for reslicing during preprocessing, these were oriented parallel to the AC-PC line. Scanning parameters are elaborated on in section 4.1.4.1.

3.2.4.2 Training

Prior to scanning, each subject completed a training session during which they were taught to associate the visual presentation of four pure color cues with the different single finger movements of the right hand. Stimuli were generated for each subject using the purpose-built psychophysical application (see section 3.3). A reduced number of trials involving only

the pure colour stimuli were presented, though the mode or presentation was identical to that used in the experiment itself, to allow the participant to become familiar with the task and hand restraint. Feedback was provided after every 12 trials, with both color-response association accuracy and movement consistency being assessed until subjects achieved an accuracy of 90%.

3.3 Psychophysical colour cue calibration

Each prospective subject was tested for relative normality of colour vision using a saturated Farnsworth panel D 15 color vision test derived from the full 100-Hue panel set [411]. However, there will still remain subtle differences in colour perception within the normal range across subjects. As the intention was to manipulate Competition in as linear a manner as possible, such differences had to be accounted for. To this end, a small piece of software (named IsoColour) was developed using Java (see appendix). For each participant, this:

- Calibrated the pure colour cues to be isoluminant
- Established the true midpoints between each of the pure colour cues
- Referring to these calibrated midpoints, generated a set of blended colour stimuli that were equidistant in colour space

It took each participant approximately ten minutes to complete the IsoColour procedure.

All colours were considered using the CIE L*u*v* colour space, which has been designed and empirically tested to linearly approximate human colour perception [412]. It is specifically tuned for use with additive colour displays, such as the LCD screens used within MRI scanners. It is particularly well-suited to the task at hand, as it allows independent adjustment of luminance(L*), red/green (u*) and yellow/blue (v*) components. All these components alter perceived chromaticity.

The pure cues were chosen to correspond approximately to red, yellow, green and blue, with equidistant coordinates in CIEL*u*v* color space (white=D-50, observer=1964, 10'). Their values before participant-specific tuning are shown in table 3.1.

Colour	Luminance (L*)	Red/green (u*)	Yellow/blue (v*)	Associated movement
Red	80	103.73	-12.13	Index flexion
Yellow	80	25.75	55.80	Index extension
Green	80	-78.87	11.48	Middle extension
Blue	80	-27.99	-68.30	Middle flexion

Table 3.1: The CIE L*u*v* values chosen to represent the pure colour cues to be directly associated with finger flexions/extensions.

3.3.1 Isoluminance

As luminance affects perceived chromaticity, it was necessary to ensure that the 4 pure cues were isoluminant for each subject. This was determined using the minimum motion technique [413], which exploits the property that only luminance can provide rapid movement information via the M pathway.

In the IsoColour software, this involved presenting the participant with a background of one of the pure cues, and a black fixation cross in the centre of the screen. Four spots of another pure cue (e.g. red background, yellow spots), were then presented outside their foveal field, flashing at 15Hz (figure 3.6). The participant then pressed buttons to gradually adjust the luminance of the four spots, until the perceived flicker was reduced, ideally to the point where it was no longer apparent at all. This is the value for which the background and foreground pure cues can be said to have match luminance.

The yellow, green and blue primaries were adjusted to match that of red. The order of primary-pair presentations was randomised, which each being presented 3 times. These would start from numerical isoluminance, and then be adjusted from each previous best judgement. The final luminance value was averaged across the 3 repetitions.

3.3.2 Midpoint determination

To accommodate individual differences, and attempt to manipulate Competition as linearly as possible, each participant's "zone of confusion" around the midpoint between each pure-pair (e.g., red vs yellow, yellow vs green etc) was calculated. The experimental stimuli could then be created by determining an equidistant spectrum of colour blends between this calculated midpoint, and the pure cues at each end. By generating these perceptually-linear spectra, it was hoped that Competition would have no points of saturation, and that each step change would be equal to the last. The exact midpoints would not be presented as experimental

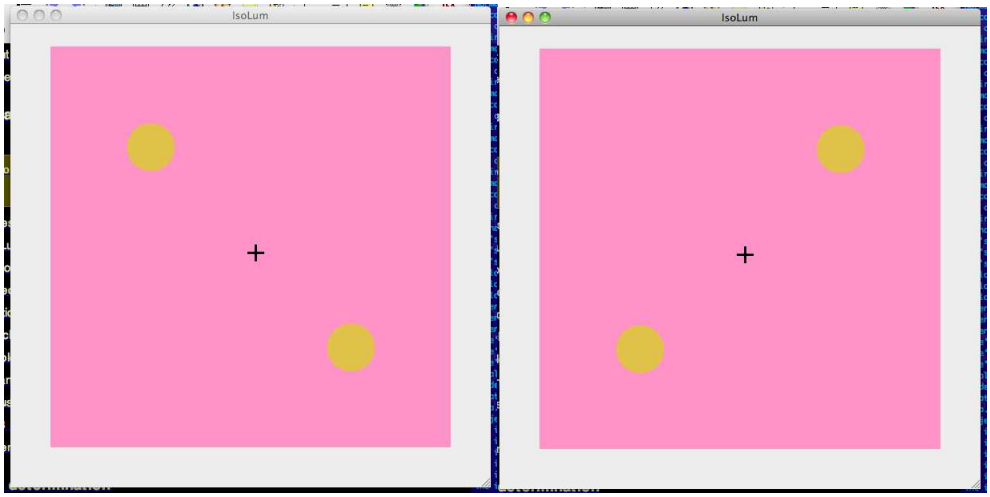


Figure 3.6: Screen shots of the isoluminance calibration stage.

stimuli, meaning that every Competition would technically have a correct answer, which the participant should theoretically be just able to perceive.

To be able to estimate each participant's perceived midpoints in as bias-free and accurate a manner as possible, a Cambridge C-style test [414] was implemented, during which subjects had to indicate the location of the break in a "C". These were presented in a spatially- and luminantly-noisy manner (figure 3.7), forcing the participant to rely entirely on chromatic information. The background was the numerically-defined midpoint, whereas the "C" was initially a pure cue, whose colour values gradually converged on those of the numerical midpoint. The task therefore became increasingly challenging as the "C" blended into the background. This method allowed the accurate location of the chromatic point at which the subject could no longer differentiate the numerical midpoint from the currently presented comparison colour.

An adaptive psychophysical approach was adopted to determine how much the colour parameters of the "C" should change for each subsequent presentation [415]. The specific algorithm chosen was Parameter Estimation by Sequential Testing, or PEST [416], which involved eight randomly interleaved staircases - one for each pure cue -> midpoint spectrum - and used the following heuristic rules to govern step changes in "C" colour parameters:

- At each reversal, i.e. when the participant started to go down the staircase when they had previously been going up (for example, is they incorrectly located the break in the "C"), the step size in colour parameters was halved
- The second step along the staircase was the same size as the first.

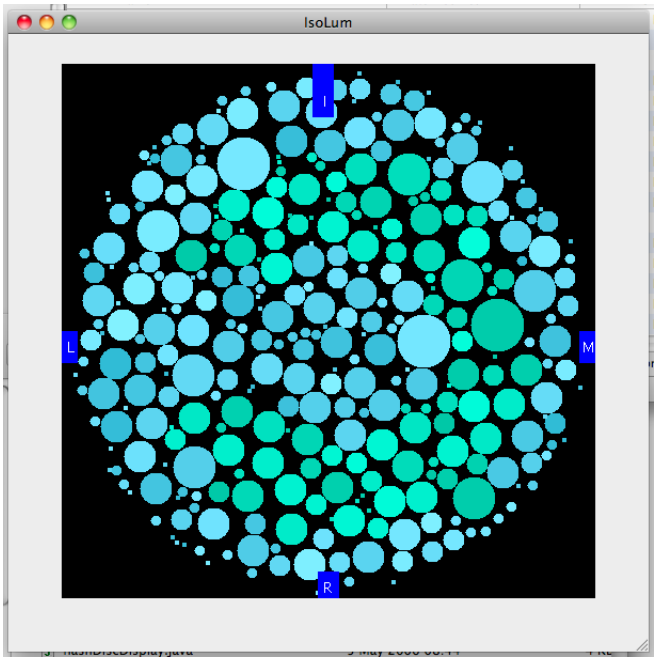


Figure 3.7: Cambridge C-style test sample displayed as part of the procedure to identify the participant's personal colour midpoint between two pure cues, in this case blue and green. The task is to locate the "break" in the C character formed by the contrasting dots - here on the left. The colours are presented within a spatially- and luminantly-noisy environment, forcing the participant to depend on chromatic information alone.

- The fourth and subsequent steps along the staircase are double their predecessors
- If the step before the last reversal was a doubling, don't double the third; however if not, then do

The procedure terminated when the size of the step dropped below 1% of the numerical primary-to-midpoint distance, i.e. in accordance with PEST's Minimum Overshoot and Undershoot Sequential Estimation (MOUSE) mode [416]. The colour parameters for this final "C" were taken to be the perceptual midpoint for the associated spectrum. After all the participant's perceptual midpoints had been determined, they went through a final stage of luminance calibration, using the same flashing minimum motion technique as described earlier. The experimental stimuli were then generated by calculating the CIE L*u*v* parameters for 9 equidistant points between along each pure-to-midpoint spectrum. This meant each of the three experiment levels of Competition contained 24 colour stimuli - three for each of the eight pure-to-midpoint spectra.

3.4 Monitoring finger movements

The need to record finger extensions as well as the usual flexions presented a small technical challenge. As no device was available locally to perform such a task, some alternative hardware was sought. The greatest constraint on the nature of that hardware was the need to be MR-compatible. One class of devices that offered good functional flexibility, real-time accurate data logging and complete MR compatibility was the fibre-optic data glove. The specific product chosen was the 5DT Data Glove 5 MRI (http://www.5dt.com/products/pdataglovemri.html), which also included a software development kit to allow optimal data logging (figure 3.8).

3.4.1 Data acquisition

In order to monitor and record the glove data, a small program was written in C that called various methods provided by the 5DT SDK. The functionality covered was:

- openGlove(): Opening the glove via a USB port
- monitorGlove(): Sampling the index and middle finger data readouts at a frequency of 33Hz, storing this data (together with the system time stamp) in a buffer, and writing to disc every 5 seconds.

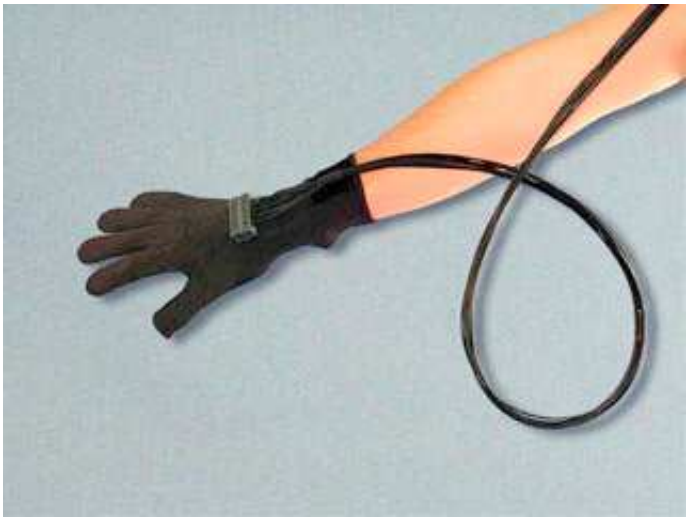


Figure 3.8: The 5DT Data Glove 5 MRI. This includes five sensors recording the relative degree of flexion/extension of each of the digits. These work via optical fibres: the increased attenuation of transmitted light brought about by mechanical deformation is used to determine how bent each finger is. The glove itself is made of Lycra, and the MR-incompatible control box remains in the scanner control room, with communication taking place over a 7m fibre-optic cable passed through the waveguide.

The scanner's RF onset pulse signal was also recorded via a CED MICRO1401 data acquisition unit (http://www.ced.co.uk), and the system time stamp at which this was received written to disc. This allowed for the synchronisation of stimulus, response and scanner data during subsequent analysis.

3.4.1.1 Calibration

The first study involving only healthy volunteers did not provide the participants with any feedback during the scan itself. In this case, the glove monitoring program simply wrote the data to disc, and it underwent no direct interrogation during the experiment itself. However, the glove's dynamic range varied quite significantly according to the size of the participant's hand, and the manner in which the hand was secured within the restraint. To improve the sensitivity and reliability of the glove data, it was therefore necessary conduct a pre-scan check of the glove data stream. 5DT provided a simple monitoring package, displaying the data in real-time for each digit. While this was running, the participant was asked to systematically flex and extend their index and middle fingers, returning to central rest each time, and the results observed by the experimenter. This allowed her to determine if the dynamic range was of a sufficient size to be able to reliably determine each of the participant's movements during subsequent analysis,

or whether the hand restraint would have to be subtly adjusted to allow for greater clarity.

However, as shall be further explained in section 6.1.3, this was not true of the second study: here trial-by-trial feedback was provided to the participant during the scan itself, which required rapid analysis of the glove data, and significant integration with E-prime. A more rigorous assessment of the participant's movement parameters therefore needed to be established before scanning began. This was done via a pre-scan E-prime calibration paradigm, conducted within the scanner. This followed a similar procedure to the original pre-scan check, though this time the averages of each movement were recorded to disc, and displayed on screen, allowing the experimenter to gauge sensitivity/reliability. The recorded values were used by E-Prime to assess the participant's experimental responses, described in the next section.

The following two sections concerning preprocessing and analysis of the glove data will reflect the different ”batch versus real-time” nature of their use.

3.4.2 Preprocessing and analysis

The glove data underwent a number of preprocessing steps before analysis could take place. For both the first and second studies, prior to the actual fMRI analysis, this was done by means of a Matlab script. For the second study only, as real-time feedback was being provided to subjects during scanning, a cruder form of preprocessing and analysis also took place in order to improve the accuracy of movement categorisation. In all cases, the glove data and E-prime logs were brought into temporal alignment by means of their respective time stamps.

3.4.2.1 Second study only: analysis for real-time feedback

Within E-prime, the user-defined script facility was used to implement several methods for the preprocessing and analysis of glove data. First, the mean rest and flexion/extension readings that had been determined during pre-scan calibration were imported from their log file. These set the parameters by which each response was to be judged.

At the start of each trial, two important events occurred: the time stamp of stimulus onset, and the expected response for that trial were written to a text file; and the glove monitoring program was run (described in the previous section). All the participant's movements were recorded over the course of stimulus presentation and the response window, the end of which signalled the glove monitoring program to close, saving the glove data to file. These two files were then opened, and the trial onset time stamp used to seek to the right location in the glove

data file. The method *getBaseline()* was then used to determine a reference baseline position ref_i for each finger i, by averaging the previous 6 samples.

The method *testAccuracy()* was then used to determine what kind of feedback the participant should receive. A particular movement m (flexion or extension) was considered to have been performed if the finger deviation dev_{im} was:

$$dev_{im} > 0.66\,(max_{im} - ref_i)$$

Where max_{im} is the maximal deviation for each finger movement, determined during the pre-scan calibration procedure. In this case, the participant received ”correct” feedback. For cases were no response was detected, or multiple significant movements were detected, they were informed that their movement was ”incorrect”.

To accommodate any extra delay that may have been introduced by this processing of glove data, the intended and actual length of each trial was compared, and the difference subtracted from the pre-stimulus rest period of the subsequent trial.

3.4.2.2 More thorough Matlab preprocessing prior to fMRI analysis

For the actual fMRI analysis, a more detailed examination of the trials took place.

Defining trial types according to the participant's responses To ensure that events were representative of the intended experimental manipulation, glove data were analysed to assess the outcome of each event, with incorrect/unclear responses being classed as nuisance regressors. Trials were then categorised according to the outcome of the trial, that is, whether or not the participant produced a coherent response. Two nuisance regressors were created on this basis:

- A ”failure of suppression” regressor: these were trials that showed evidence that the participant attempted to execute both of the movements evoked by the ambiguous stimulus.
  - For low Selectivity trials, i.e. decisions between movements on different fingers, the suppression process was deemed to have failed if there was significant movement on both fingers. This would clearly be confounded by the structural mechanical interdependence between the fingers, therefore this movement also had to be sufficiently decorrelated. To determine what ”sufficiently decorrelated” would be, for each trial, correlation between the two fingers was calculated, and normalised over the run. Values great than 1.5 standard deviations were considered decorrelated.

  - For high Selectivity trials, i.e. decisions between movements on the same finger, the finger in question had to show a significant movement in the opposite direction before finally completing the contrary movement. This time ”significant movement” was calculated by first normalising all the deviations recorded for each finger movement. Movements above 1.5 standard deviations were considered significant/deliberate.

- A general ”incorrect” regressor: this included trials where no response was provided; where it was simply incorrect (e.g. a movement on the wrong finger altogether for a high Selectivity trial); or the movements were so incoherent that no further categorisation could be attempted.

Glove data processing As occurred during the E-Prime analysis, the trial onset time was used to isolate the correct section of the glove data time stream. The post-stimulus data had a mean of the previous 5 samples subtracted from it to determine the relative movement deviation.

Upon reviewing the glove data streams, it became apparent that in some circumstances the initial stages of a flexion would produce a signal that resembled extension, before turning and being expressed as a proper flexion. This seemed to be a factor of how the restraint's elasticated bands were impinging on the optical fibre's anchor points within the fabric of the glove. It was therefore not entirely straightforward to isolate the true peak of movement deviation. In order to determine this more reliably, two peak values were attempted to be identified for each of the flexion and extension directions. In order to determine that each of these local maxima was indeed a unique maxima and not part of the build-up to a larger peak, these values were tracked back to baseline in both directions. Once a peak had been identified in this way, it was removed from the timecourse, to allow for the clear detection of a possible second peak. The larger of the two was taken to be the true response for that direction, and similarly the larger of the flexion vs extension deviations was also taken to be the intended movement.

These artifact signals were relatively small in magnitude compared to deliberate movements. In order to be able to distinguish these noisy glove signals from a genuine ”failure of suppression”, a deviation of:

$$x_{im} > 0.33\,(max_{im} - ref_i)$$

was considered to be deliberate, and not an artifact.

3.5 Effective connectivity analysis: Granger causality mapping

Localising areas whose activity is correlated with experimental factors is the first step in understanding how the brain systemically mediates a behavioural process. Perceptual decision-making is expected to engage a broad network of cortical and subcortical structures. However, as the levels of Selectivity or Competition increase, it is not clear how each of these areas will respond, or how the relationships between these regions will alter accordingly. It is anticipated that the experiments will reveal a dynamical shift in the influence particular brain regions exert over the process, mediated by the BG. However, the repeated architectural motif of multiple closed- and open-looped cortico-BG-thalamocortical channels makes understanding the directionality of these influences rather difficult to capture. Methods that fall under the blanket term of ”functional connectivity” will not help in addressing this issue: broadly, these quantify the degree of simple correlation between multiple brain regions [417, 418]. Instead, an effective connectivity approach was adopted: these methods make an attempt at determining the direction of influence, and encompass several methodologies that can be said to fall into two broad categories: model-driven and data-driven.

Model-driven approaches include structural equation modelling [419, 420], dynamic causal modelling [421], and psychophysiological interaction [422, 423]. In each case, a limited number of regions of interest must be defined *a priori*, with the direction of influence between each explicitly defined. These links can be modulated by the task itself, with parametric designs allowing for a clearer assessment of those effects. However, if the precise network cannot be identified beforehand, or there is a strong suspicion of as-yet-undefined but important neural nodes playing a part in the network that has been defined, then the model-driven approach may fall down. In these cases, a more flexible approach must be taken.

Granger causality is an example of data-driven effective connectivity [424]. Clive Granger was an economist who devised a means of assessing ”causality” using multiple data timecourses, both by establishing the unique account that one timecourse could provide for another, and by using temporal precedence to establish the direction of that influence. This metric, used in a functional neuroimaging context, works on the principle that a functional time series extracted from a given brain region can be modelled as a projection of both its own history, and the history of time series extracted from other brain regions [425, 426, 427, 428, 429]. A region’s influence over another is a function of how its inclusion in an autoregressive model contributes to a reduction in variance. As the model operates over an order of p previous time points,

temporal precedence is used to inform the direction of that contribution.

In this study, regions of interest would be defined as those areas of the striatum and pallidus that showed a significant modulation by either Selectivity or Competition. In order to explore how cortex's contribution to the BG varied with task demand, a voxelwise bivariate measure of Granger causality would be used, seeded by BG regions of interest.

However, before data could be considered suitable for use in a Granger causality analysis, a number of preprocessing steps would have to take place.

3.5.1 Preprocessing

A requirement of connectivity methods dependent on temporal analysis is that each slice within a single volume is taken to have been acquired at the same moment. However, the TRs involved in fMRI means this is clearly not the case: an important first step of data preprocessing was therefore to perform slice time correction. This is described in section 4.1.4.3. There, the subsequent normalisation and smoothing procedures are also described. After this point, the data underwent further preprocessing stages, beyond those applied for the other more standard fMRI analyses.

3.5.1.1 Reducing physiological and movement-related noise

fMRI data is subject to multiple sources of noise, some of which may have a periodic nature, and so confound both standard fMRI model estimation and the autoregressive techniques of Granger causality. The combination of high-pass filtering and whitening [430] is a popular method for attempting to reduce these temporally-correlated influences. However, whitening would not be an appropriate step to employ before Granger causality analysis, as it could act to remove the very autocorrelations we are looking for. Instead, a different approach for removing coloured noise was used. Physiological and movement factors are two strong sources of temporally-structured noise [431]. Therefore, an attempt was made to assess and remove their effects from the data by modelling them more explicitly. For each participant. timecourses were extracted from 6mm spherical regions of interest located in (a) corpus callosum (Montreal Neurological Institute (MNI) coordinates 0 27 7) to assess white-matter-related signals, and (b) the left lateral ventricle to assess fluctuations in CSF (MNI coordinates -23 -38 17). These were taken to be a measure - or at least a strong correlate - of fluctuations in the BOLD signal that

were due to vascular and respiratory fluctuations rather than the experimental task itself [432]. These were then used as regressors in a simple first-level design matrix, together with the participant's scan-by-scan motion parameters. The models were estimated, with a high-pass filter of 120s, and the residuals taken forward into the subsequent analysis stages.

3.5.1.2 Comparing task conditions

In order to test the effects of increasing Selectivity and Competition, it would be necessary to directly compare these conditions. To allow this, the psychophysiological interaction functions available in the Statistical Parametric Mapping (SPM, (Wellcome Department of Cognitive Neurology, London, UK, `http://www.fil.ion.ucl.ac.uk/spm`) package were utilised to tease out the fMRI signal timecourses associated with each condition.

For each BG region of experimental interest, a volume of interest (VOI) was defined, and a mean timecourse extracted for all voxels within a 3mm sphere. An adapted version of Gitelman's spm_peb_ppi() function [423] was then used to deconvolve the BOLD signal x from the HRF, to produce an approximation of the underlying neural activity x_n:

$$x = HRF \times x_n$$

Using the constraint that x_n has a white spectrum, Gitelman expands this neural signal using a discrete cosine basis set x_b:

$$x_n = x_b B$$

where B is a vector of parameter estimates. A least squares algorithm is then used to estimate x:

$$x = HRF(k,:) \times x_b B$$

where k is a vector representing the temporal resolution of the scans. Once the neural timecourse x_n was calculated, versions specific to each experimental condition of interest were created by multiplying it with the appropriate event onset vector. These were finally reconvolved with the HRF to produce condition-specific BOLD timecourses. Before being used in a Granger causality analysis, they were detrended of linear and quadratic drift, and differenced (each point in the timecourse subtracted from the last) to improve stationarity, which is an important assumption underlying Granger causality measures.

3.5.2 Voxelwise bivariate analysis

This first mode of analysis would take a more open, whole-brain view of the data. It would take as input seeds those areas of the striatum and pallidus that had shown Selectivity- or Competition-dependent activation, and use them to perform bivariate assessments of Granger causality for every other voxel in the brain. It was decided that attempting to determine the *unique* contribution each voxel made to the BG (and vice versa) would be too massively multivariate a problem to tackle, especially given the relatively limited number of data points there were available to work with. Therefore a bivariate approach was taken: each pairwise voxel-to-BG assessment was performed independently of the others. This still remained as a substantial computation: to further limit the number of process cycles the analysis required, the brain was downsampled to a resolution approximating its final resolution following standard fMRI analysis: 10mm isotropic. The analysis procedure was then optimised for parallel processing, with a single participant's data set being divided into z-slices - these were then processed using Eddie, the Edinburgh Compute and Data Facilities (ECDF) high-performance Linux cluster (https://www.wiki.ed.ac.uk/display/ecdfwiki/Home).

The pairwise Granger causality-derived measure of linear feedback [433, 434] between timecourse X and timecourse Y was defined as:

$$F_{X\to Y} = log\left(\frac{|var\left[Y_t|Y_{t-order}\right]|}{|var(Y_t|W_{t-order}|}\right)$$

$$F_{Y\to X} = log\left(\frac{|var\left[X_t|X_{t-order}\right]|}{|var(X_t|W_{t-order}|}\right)$$

where $Y_t|Y_{t-order}$ signifies the regression of Y_t by $Y_{t-order}$, $W = X + Y$, and $order$ is the number of time steps over which the autoregression is to occur, which was set to 1 for all analyses, as it has been found that higher-order autoregressions fail to provide an improved account for fMRI data [426]. $|var()|$ is a measure of the cross-covariance of the residuals for each vector autoregression model, as the determinant of a matrix is its generalised variance, i.e. a function of the difference between its variance and covariance.

By excluding X from the numerator, the equation for $F_{X\to Y}$ asks, "what contribution is X making in accounting for the variance of Y during autoregression?". Each pairwise comparison was inserted into a whole-brain map of linear feedback, either flowing from cortex-to-BG seed, or BG seed-to-cortex, with each experimental condition having a separate map. These maps were then contrasted by simple subtraction. In order to determine the statistical significance of voxels within these difference maps, null distributions for each pairwise comparison were constructed. This was done by bootstrapping: the condition to which each event was assigned

was permuted randomly with replacement 800 times, and null linear feedback measures then calculated and contrasted. These distributions were used to determine p values for each voxel, which were then false discovery rate (FDR)-corrected at a threshold of 0.05 [435].

3.5.3 Inter-region multivariate analysis

The multivariate approach [434] represents an extension of the linear feedback measures described above, and was used to assess how Selectivity and Competition affected directional influences between a limited subset of nodes.

This time, $W = X + Y + Z$, where Z can be multiple additional timecourses. Therefore, in order to determine the unique combination that X makes to Y given Z:

$$\begin{aligned} F_{X \to Y|Z} &= F_{XZ \to Y} - F_{Z \to Y} \\ &= log\left(\frac{|var\left[Y_t|Y_{t-order}\right]|}{|var\left[Y_t|W_{t-order}\right]|}\right) + log\left(\frac{|var\left[Y_t|Y_{t-order}, Z_{t-order}\right]|}{|var\left[Y_t|Y_{t-order}\right]|}\right) \end{aligned}$$

3.5.4 Second-level Granger causality analysis

Assessing the significance of these results at the group level is non-trivial. A popular method conducted during previous fMRI studies is to collapse the extracted timecourses across subjects before performing the Granger analysis [425, 426, 436]. However this risks losing a lot of information concerning the specific interactions seen within individual participants. The approach taken during this work was to keep the data participant-specific for as long as possible, only collating once the individual $F_{X \to Y}$ values and their associated significances had been determined. In order to do this, the method described by Stouffer and colleagues was utilised: this has been used to collate the results of multiple independent studies during meta-analysis [437, 438, 439]. In this case, each participant can be considered an independent "study", and their statistics collated in much the same way.

First, Stouffer's method Z-transforms each first-level p value - as p_i goes from 0 to 1, Z_i will go from -infinity to infinity. These standard normal deviates are then used to form a second level Z value, Z_s:

$$Z_s = \frac{\sum_{i=1}^{k} Z_i}{\sqrt{k}}$$

where k is the number of independent measures, in this case the number of participants. Z_s

has a standard normal distribution if the null hypothesis is true, therefore a standard statistical test can determine if the cumulative evidence points to this being violated.

Chapter 4

Study 1: Action selection in healthy volunteers

A functional neuroimaging study was undertaken to test the hypotheses derived from the computational model, using the specially designed competitive decision-making task. At this validation stage, the participants were healthy volunteers.

4.1 Methods

4.1.1 Participants

Having obtained ethical and NHS R&D approval, 16 healthy participants were recruited from amongst the students and staff of the University (seven female; age range: 24-48y; mean age: 30y). Potential participants were approached either in person or via e-mail, and provided with an information sheet describing the nature and purpose of the experiment. The potential risks of participating in an MRI study were clearly explained. Written consent was obtained from each person beforehand. All participants were successfully scanned to completion, with nobody reporting discomfort or claustrophobia.

Inclusion criteria

- Right-handedness
- Aged 18-50

- Native English speaker

Exclusion criteria

- Criteria normally associated with MR imaging
- A family or personal history of mental illness
- Any neurological condition
- Currently taking psychotropic medication, including for recreational use
- Abnormal colour vision
- Any history of eye disease, which may have resulted in acquired colour vision defects

4.1.2 Behavioural procedure

For each participant, the experiment was divided into 3 phases.

4.1.2.1 Screening and colour cue calibration

Colour vision assessment The first phase involved screening participants according to the inclusion/exclusion criteria, Colour vision was assessed objectively using a saturated Farnsworth panel D 15 color vision test derived from the full 100-Hue panel set [411]. This involved arranging a set of 15 randomised colour caps into order, based on finding the one that was closest to the last in hue, starting from an initial red cap. Participants were confirmed as having normal colour vision if they correctly arranged the entire set. This test is able to detect moderate to severe abnormality, and covered the spectra employed within the perceptual decision-making task.

Psychophysical colour cue calibration Participants then completed the psychophysical colour calibration program to generate their own cue set. These were perceived by each participant as isoluminant, with the midpoints being matched to their colour vision.

4.1.2.2 Training

Participants were then taken to the Western General Hospital to perform the experiment itself. A radiographer first confirmed their suitability for scanning. They then completed a brief practice run outside the scanner, using an MR compatible optical fibre data-glove. Only unambiguous cues were presented during training, to allow them to fully learn the cue-response associations. Five ambiguous cues were shown at the end, to convey a little of what the subject was to expect in the scanner. Training took no more than 15 minutes.

4.1.2.3 Scanning

The experiment itself involved the acquisition of structural (taking 20 minutes) and functional images, over a task duration of 28 minutes, resulting in a total scan time of 50 minutes. Before scanning commenced, participants were asked to tape a small capsule of cod liver oil to their right forehead, to permit verification of image orientation during processing. During scanning, participants completed the full perceptual decision-making task, using the customised colour cues to evoke Competition between the different motor responses. To bring real-world contingencies to the task [407], participants were informed that at the end of the experiment, three of their responses would be chosen at random and a £5 reward issued for each correct answer. As we were interested in activity relating to selection only, feedback was not provided during the experiment to remove its potential confounding effects. The means by which glove data was recorded for subsequent analysis is described in section 3.4.

4.1.3 Experimental design

The decision-making test was deployed within a rapid event-related fMRI experiment, during which the colour stimuli were presented in an order that was optimised for the experimental contrasts of interest. This was done using the Genetic Algorithm (GA) toolbox [405], which randomly generated trial lists within certain pre-specified constraints:

- Number of conditions: 8, incorporating 3 levels of Competition, over 2 levels of Selectivity, plus 2 control conditions (checkerboard and unambiguous colour cues).
- Scan length: 431s per run, with each subject completing four runs.
- Number of trials per condition: 12 per run.
- Interstimulus interval: 2.8s.

- Scan TR: 2.2s

These random trial lists were then convolved with a canonical HRF function, and both high- and low-pass filtered to produce an output that could be considered in terms of its usefulness for fMRI experiments. They were assessed for ”fitness” according to features that have been specified as being desirable. Arbitrary values are used to denote their relative importance (which need not sum to 1):

- Strictness of counterbalancing: this was set to 0.6, which was deemed sufficiently high to enable estimation of each condition, but reflected that there was no psychological reasons for wanting to avoid runs of similar trials
- Contrast efficiency: set to 1, indicating that it was important to have sufficient power to detect differences between conditions. The desired contrasts were defined as:
  - High > Low Selectivity
  - High > Intermediate and Intermediate > Low Competition
  - All experimental conditions > Checkerboard and All experimental conditions > Unambiguous control (though these were assigned a lower weighting than the key Selectivity and Competition contrasts)
- HRF estimation efficiency: set to 1, as estimating the shape of the HRF response was important for the intended connectivity analyses.
- Strictness of frequency presentation: set to 2, as it was desired that each condition be presented an equal number of times.

Those lists achieving high levels of fitness were selected for subsequent ”breeding”, which involved blending their trial orders in a systematic way, with the hope of producing a generation of lists with even fitter characteristics. Random mutations would also be occasionally introduced, to prevent the process settling on a local minimum. This was performed with 40 organisms per generation, over 100,000 generations, or until fitness no longer varied by a significant amount. An optimised trial sequence was produced for each participant, and for each run: 64 in all.

Finger movements were fully counterbalanced across conditions to average out their individual effects over the course of the experiment, allowing the specific examination of the experimental manipulations of interest.

4.1.4 Functional imaging

4.1.4.1 Data acquisition

Functional imaging protocol Imaging was performed using a GE 1.5 T Signa scanner (GE Medical, Milwaukee, USA) at the SFC Brain Imaging Research Centre, Edinburgh. Axial gradient echo, echo-planar (EPI) whole brain images oriented parallel to the AC-PC line, to facilitate coregistration with the high resolution structural volume, and subsequent normalisation to the MNI template. Acquisition parameters were acquired with echo time (TE=40ms, TR=2200ms, matrix=64x64 and field of view (FOV)=22cm, providing an in-plane resolution of 3.75 x 3.75mm. Volumes comprised 28 contiguous 4mm slices, acquired in a bottom-up interleaved manner. All 16 subjects completed four runs of 440s each, producing 200 volumes per run. The first four volumes of each run were discarded to avoid T1 saturation effects.

Structural imaging protocol A high resolution structural T1 volume was obtained using a coronal gradient echo sequence with magnetisation preparation (MPRAGE). Acquisition parameters were TI=600 ms, TE=3.4 ms, flip angle=15°, FOV=22, u sing a matrix of 256x192, providing an in-plane resolution of 0.86 x 1.12mm. The volume had 128 slices of 1.7 mm thickness. An oil capsule was attached to the right forehead of each participant to allow verification of volume orientation during preprocessing and analysis.

4.1.4.2 Data reconstruction and quality assurance

For each subject, functional data were reconstructed using the GE_convertADW function, supplied as part of the GE2SPM 3.1 toolbox. The structural T1 volume was reconstructed within MRIcro, reoriented to match the functional images, and had its origin reset to the centre of the anterior commissure.

Overall data quality was assessed by manually viewing each subject's complete timecourse as a "movie" using the spm_movie function. Two participants' data contained spikes in the form of single ultra-bright volumes: these single volumes were replaced by interpolation, that is, by averaging the adjacent two volumes using SPM's imcalc function.

Data were then preliminarily-realigned, and the subsequent motion parameters assessed for excessive motion. This was defined as a volume-to-volume translation greater than half of one voxel (1.9mm), or 1 degree of rotation. Three participants showed single rotations between 1 and 1.5 degree: again, the respective affected volumes were again replaced with an

interpolation of their immediate neighbours in time.

Image left-right orientation was verified by referring to the oil capsule on the participants' right forehead. To conclude, all data acquired from the 16 healthy volunteers were able to be used in subsequent analyses.

4.1.4.3 Preprocessing

All preprocessing and analysis was carried out using SPM2. To preserve data quality during reslicing, the default interpolation method was set to 7th order B-spline to improve accuracy. The density of voxel sampling was increased to maximum (quality setting of 1.0).

Slice time correction As the intention was to later perform connectivity analysis, the functional data were initially slice time corrected. This step adjusts each slice of an individual fMRI volume to appear as though it was acquired at the same time, which in this case was the moment the central slice was acquired. With a TR of 2.2s, this is of course not the case, which could affect analyses dependent on the temporal relationships between voxels, which in an umbrella Granger causality analysis firmly falls beneath. Slice time correction achieves this by sinc interpolation: Fourier transforming the data into frequency space, and adding a constant to the phase of every frequency. This has the effect of shifting the signal in time to match that of the nominated reference slice. Any additional interpolation of the data is not ideal, as it is at best an approximation, and acts as a general source of degradation. However, in this case it was deemed necessary, as the *a priori* interest in cortico-BG connectivity meant that the areas to be assessed would be significantly remote from each other along the z axis, and therefore could be over half a TR apart in time due to the interleaved acquisition. Within Granger causality analysis, this could alter the degree of variance that one voxel accounts for in the activity observed in another, producing erroneous or spurious results. It was believed that the data would be robust to the negative consequences of interpolation, as many repetitions of each trial were conducted, and efforts had been made to optimise the efficiency of HRF estimation via the order and timing of trial presentation. This would ensure a broader sampling of the HRF for a given trial condition, allowing for more accurate interpolation. The middle slice was selected as the reference, to minimise any artifact introduced by interpolation within this study's key region of interest: the BG.

One key issue to consider was whether slice time correction should occur before or after spatial realignment. For subjects showing excessive motion, slice time correction can result in data

being interpolated between entirely different regions. However, if performed after realignment, slice time correction can move regions into different time point, which is especially significant for interleaved acquisition, as neighbouring voxels could be pushed half a TR away. This point would apply to the data acquired for this study, and as the participants did not show excessive motion, it was deemed more appropriate to perform slice time correction before realignment. One additional point to note was that the default sampled time bin was changed from 1 to 14, to ensure that event onsets were correct during first level model specification.

Realignment The slice time corrected data was then spatially realigned, by calculating a mean image, and registering all other functional volumes to it. The data were not resliced at this stage, to minimise degradation, although a mean image was generated for use during coregistration.

Coregistration The mean functional image (coregistration source) was then coregistered to the respective structural image (coregistration target). The resultant transformation parameters were applied to every other functional volume by imposing the mean functional image's space onto all others in the time series, using spm_get_space.m. Again, the functional data were not resliced at this point.

Normalisation and smoothing To provide better quality normalisation parameters, the structural (as opposed to the mean functional) image was warped to the MNI T1 template. These parameters were then applied to the now similarly-registered functional data, which were then resliced at a resolution of 2 x 2 x 2mm, and smoothed with an 8mm FWHM (Full Width Half Maximum) Gaussian filter. 8mm was chosen as this was a compromise between targetting the extent of activation intended to be observed within the BG, (a relatively constrained collection of spatial structures) according to matched filter theorem; and the desire to reduce the multiple comparisons problem by imposing broader spatial structure upon the data.

4.1.4.4 Analysis: Perceptual decision-making

There were two main arcs to the analysis for the first study. The first was to interrogate the perceptual decision-making process: to establish a path of how the brain transformed an ambiguous sensory stimulus into an associated motor response, with particular emphasis on understanding what role the BG played in that process. The second arc was to establish how sufficient an account of the experimental data was provided by the computational model.

Functional volumes from all sessions were treated as a time series, and experimental effects estimated using a multi-session design matrix per subject that included separate session mean terms. Effects of interest were modelled as events whose onsets were locked to stimulus presentation.

Trials were also categorised in part according to the response provided by the participant: trials could be ”failures of suppression” or more generally incorrect (see 3.4.2.2). For analyses where the emphasis was on stimulus-driven effects, a trial was classified according to the ambiguity of the cue, regardless of whether or not the participant selected the correct or incorrect opposite response. However, for later analyses where the interest was in what the participant actually selected, trials were classified according to the hue level of the chosen cue associated with the chosen response. This will be elaborated on in the subsections below. Event onsets were convolved with a canonical haemodynamic response function and its temporal derivative. The inclusion of the temporal derivative was intended to allow the model to provide a better account for the data by representing the variance associated with delays in HRF onset. This would reduce overall residual error, and so increase the power of subsequent T tests between conditions of interest. However, the temporal derivative was not itself included in these contrasts: it is somewhat unintuitive as to what the temporal derivative would be representing in terms of task-elicited activation. Finally, the six realignment translations parameter vectors were also included as nuisance regressors, in an attempt to capture motion-related artifacts that were not themselves removed from the data by simple realignment. In addition, a more liberal threshold was applied to first level statistics, by changing SPM’s default.stats.fmri.ufp parameter from 0.001 to 0.01, allowing each participant to contribute more voxels to the second-level analyses.

The data was interrogated from a number of stances, representing the stages of a perceptual decision [335]:

Colour hue level The encoding of the sensory input, which in this case is colour hue level, in a general manner, i.e. not related to a specific hue. The first stage of perceptual decision-making involves an assessment of the stimuli guiding the decision. In the context of this paradigm, movements are prompted by specific, trained colour associations. When participants are faced with blends of the four primary association colours, the relative levels of the two competitors must be appreciated before evaluation and action selection can proceed. The first stage of the analysis therefore was to determine if visual regions were encoding variance of colour hue. The data was therefore considered in terms of the level of each hue

Ambiguity level	Unambiguous	Low	Mid	High	Mid	Low	Unambiguous
Red hue	6	5	4	3	4	5	0
Yellow hue	0	1	2	3	4	5	6

Figure 4.1: An example of how the data were considered in terms of stimulus colour hue between the red and yellow primary cues. The ”Ambiguity level” row states the manipulated experimental condition. The ”Red hue” and ”Yellow hue” rows grade the stimuli according to the level of their respective hues.

displayed, breaking the experimental ambiguity levels out across the colour-response associations. Figure 4.1 demonstrates how this was done between the red and yellow primary cues.A second-level design matrix was constructed (figure 4.7) that gathered the respective ”Stimulus x > Checkerboard” t tests into a large ANOVA (dependent groups, corrected for nonsphericity). The effect of hue variance was assessed in an F test that parametrically examined each primary colour separately, between the hue levels of 3-4-5-6-5-4-3.

Hue levels 2 and 1 were not included, to better orthogonalise the F-test components. The parameterisation was done individually for each colour in order to isolate the hue-specific effects, but all hues were included in the F-test to improve sensitivity, as it may be that not all subjects show clear graduations for all hues.

To improve the sensitivity with which significant colour-related activation could be detected, an a priori anatomical region of interest (ROI) was defined by reviewing previous fMRI and primate electrophysiological experiments of colour processing [440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453]. This isolated several regions especially involved in colour processing, which were bilateral inferior occipital cortex, and bilateral lingual, calcarine and fusiform gyri. These were all defined according to the Wake Forest University (WFU) PickAtlas [454, 455], using the Anatomical Automatic Labelling (AAL) descriptions of these anatomical features.

Accumulation of sensory evidence During the analysis of this experiment, sensory accumulation was considered to be the inverse of ambiguity/Competition - there is a large amount of clear sensory evidence for low-Competition trials, which will lend itself to a rapid rise of neural activity in the relevant anatomical regions. Previous review studies [325, 335] indicated that an appropriate anatomical ROI would contain bilateral middle and superior occipital cortex, inferior and middle temporal cortex, and the intraparietal sulcus.

Integration to a decision variable In this study, the decision variable (as it is known) was also considered as the inverse of ambiguity - it was assayed using the same contrast as for ”accumulation of sensory evidence”. However, the implicated regions are different, therefore a difference anatomical ROI was constructed, again based on reviews of fMRI findings for ambiguous perceptual decision-making tasks [325, 335]. It included the superior and middle frontal gyri; orbitofrontal cortex; precentral and supplementary motor areas; the precuneus, supramarginal and angular gyri; and superior parietal cortex.

Selected response analysis This encompasses brain activation that correlated with the level of sensory evidence supporting the response actually chosen, i.e. transforming the sensory evidence to an integrated decision. It is important to note that, up to this point, the data was considered in terms of the stimulus properties, and its temporal onset. This entails a number of issues which may confound clear analysis and interpretation. As shown by the behavioural analysis, reaction time (RT) significantly covaried with stimulus ambiguity, the inverse of which (”value”) was tested in the sensory accumulation and division integration analyses. In addition, although colour cues were tailored to each participant’s perceptual range, winning stimulus value will not reflect the judgement made by the participant if they incorrectly chose the losing alternative. Therefore a different approach was adopted at this stage: to consider trials in terms of the moment of response execution, and the decision taken at that time. Trials were classified according to the hue level of the response actually chosen, with first-level design matrix regressors being synced to the RT rather than the stimulus onset. A similar parametric T test to those used previously examined the effects of increasing *chosen* stimulus value, applying the decision integration ROI after whole brain analysis.

Competitiveness This refers to how the ambiguity of the presented cue was encoded, i.e. the Competitiveness of the decision being resolved. Ambiguity was examined by constructing whole-brain single-subject t-tests of Medium > Low, High > Low and High > Medium (combining both suppression levels), which were then taken into a second level random effects ANOVA, correcting for nonsphericity across conditions. The resultant maps were thresholded at $p < 0.1$, FWE-corrected for multiple comparisons. This low threshold was chosen to allow for the detection of comparatively weak effects at the whole brain level.

To improve sensitivity for detecting ambiguity-related effects, small volume correction was performed within an a priori mask of previously implicated cortical regions. These were derived from a review of functional imaging studies concerning ambiguity and risk [339], and consisted

of the following regions defined according to the AAL atlas within the WFU PickAtlas [454, 455]: bilateral anterior and middle cingulate; superior, middle and inferior frontal cortex; insula; inferior and superior parietal cortex; rectus; supplementary motor area; left precuneus and right angular cortex; right precentral and postcentral cortex; bilateral caudate and amygdala [456]; left thalamus.

As the key focus here was the BG, an AAL-derived mask formed by a union of bilateral caudate and putamen was also used.

Selectivity This aspect of BG suppression was examined using the contrast of High > Low Selectivity. To test the computational model's specific predictions of the effects of Selectivity (and its interaction with Competition) within the pallidus, voxel-wise analyses were constrained within an AAL-derived anatomical mask of left pallidus (as the responses were performed with the right hand. Single-subject whole brain contrast maps for High > Low Selectivity (combining all ambiguity levels) were assessed within a second-level random effects two-tailed t-test.

Selectivity's interaction with ambiguity/Competition was determined by calculating separate single-subject t-tests of High > Low suppression for each Competition level.

Activation x reaction time correlations It was determined how task conditions altered the relationship between pallidal activity and behavioral response by producing averaged event-related timecourses for each condition (MarsBar toolkit for SPM2 [457], extracting the HRF time-to-peak, and correlating this with condition-specific average reaction time.

Random effects analysis For all the above-described analyses, having estimated these single-subject models, contrasts of interest were taken forward into random effects analyses (either t-tests for simple contrasts, or ANOVAs with appropriate correction for nonsphericity for parametric analyses/interactions), to determine significant experimental effect at the population level. Whole brain effects were established by thresholding group maps at p=0.001 uncorrected. The surviving voxels were considered significant if their height exceeded a significance level of $p < 0.05$ FWE-corrected for multiple comparisons, either for the brain as a whole, or within the relevant anatomical ROI.

4.1.4.5 Analysis: Comparing simulation to fMRI data

Simulations of the computational model were conducted, which allows one to explore the system's theoretic behaviour and dynamics. However, in this case the goal was to determine how well it accounted for the experimental fMRI data. Therefore the model's output needed to first be converted into a form that would allow this to happen. The computational model simulated neural activity in a rate-encoded way: values are scaled via a softmax function to fit between one and zero, with higher values indicated faster firing frequencies. One question to be addressed was how this simulated data was to be compared to the data acquired during experimental fMRI. Previous studies that have used computed regressors, for example those examining temporal difference modelling [458], normally convolve these with a canonical HRF before incorporating them into a general linear model (GLM) for experimental data regression. In this way, the model's rate-encoded output can be simply converted into a form analogous to BOLD activation, allowing the model's goodness of fit to be directly assessed.

To determine whether the computational model could provide a good account of experimental data collected during human fMRI, striatal and pallidal regressors were derived from a further set of simulations. For each participant, their series of experimental stimuli were entered into the model as combinations of cortical inputs, with c_i and c_j values ranging similarly as above (section 3.1) according to stimulus ambiguity. Trials with High and Low Selectivity had their associated w_g values set to -0.6 and -0.9 respectively. The output regressors were computed for each trial by separately averaging across all channels within the striatum and pallidus after system stabilisation. These regressors therefore reflected the summed activity of each nucleus, allowing a better match to the spatial resolution of functional imaging data.

During first-level design matrix specification, event onsets were defined by stimulus onsets, which were parametrically-modulated by the striatum and pallidus regressors. Within SPM, there were convolved with the canonical HRF and its temporal derivative, similar to earlier analyses. Motion regressors were also included as nuisance variables. Separate t-contrasts for the canonically-convolved striatum and pallidus regressors (versus the implicit baseline) were taken into second-level random effects t-tests to determine if the model provided a good account for BG activation across the sample. As the questions posed by these analyses were very clearly delineated anatomically [459], results were examined at an uncorrected threshold of $p < 0.05$ within AAL-defined left pallidus, and left putamen.

4.1.4.6 Connectivity analyses

Given the high degree of parallelisation seen within the BG, the next step was to determine how the pallidus interacted with the rest of the brain under varying task conditions. To avoid excluding potential areas of influence, an approach was adopted that did not require the specification of a constrained model. The areas of peak pallidal activation were used to define seeds that were used in subsequent vector autoregressive modelling, deriving directional measures of Granger causality with the rest of the brain. This metric, used in a functional neuroimaging context, works on the principle that a functional time series extracted from a given brain region can be modelled as a projection of both its own history, and the history of time series extracted from other brain regions [425, 426, 436, 428, 429]. A region's influence over another is a function of how its inclusion in the model contributes to a reduction in variance. As the model operates over an order of p previous time points, temporal precedence is used to inform the direction of that contribution.

The brain was considered at a resolution approximately matching its resel dimensions for the current data - 10mm isotropic. The analysis was decomposed into groups of 10 resels, and performed in parallel on a 128 node cluster machine. For each subject, timecourses were extracted, averaged across the resel, and high-pass filtered (cut-off 120s). To reduce the influence of extraneous physiological and movement-related factors, the data were regressed by timecourses extracted from the corpus callosum, left lateral ventricle, and the scan-by-scan realignment parameters. The residuals were then differenced to improve stationarity. Only those periods during which experimental events of interest were occurring were considered, by determining their psychophysiological interactions, and using forward modelling to produce timecourses of interest [423]. Bivariate Granger causality difference maps were then calculated between pallidal seed regions and the rest of the brain, regressing over the current and 1 previous time point [426]. Null distributions were generated for each resel by permuting the seed timecourse randomly with replacement 800 times, allowing the difference maps to be thresholded at $p < 0.05$ FDR-corrected [435]. The corrected maps of different seeds were then compared, with null distributions generated by randomly allocating resels to one seed or the other, thresholding the results at $p < 0.05$.

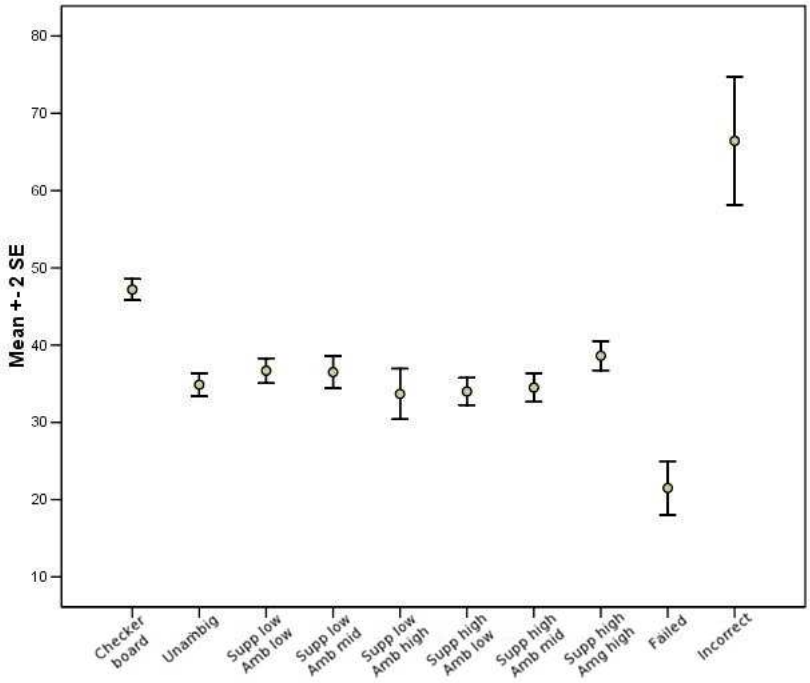


Figure 4.2: Total trials representing each experimental condition. Error bars display standard error.

4.2 Results

4.2.1 Behavioural results

The glove data was analysed to produce trial-by-trial assessments of reaction time and response accuracy. As well as establishing the effects of Selectivity and Competition, it was important to confirm the validity of subsequent functional analyses by checking that conditions were not intrinsically biased according to movement type or reaction time.

4.2.1.1 Examining possible bias

Experimental conditions were equally represented The intended contrasts of interest could be affected if a particular condition was poorly represented, i.e. if participants got an inordinately large number of those trials incorrect, resulting in their exclusion to the nuisance trial regressor. Figure 4.2 shows the number of trials associated with each condition.

The opportunist nature of failed trials will make them inevitably less frequent, however the

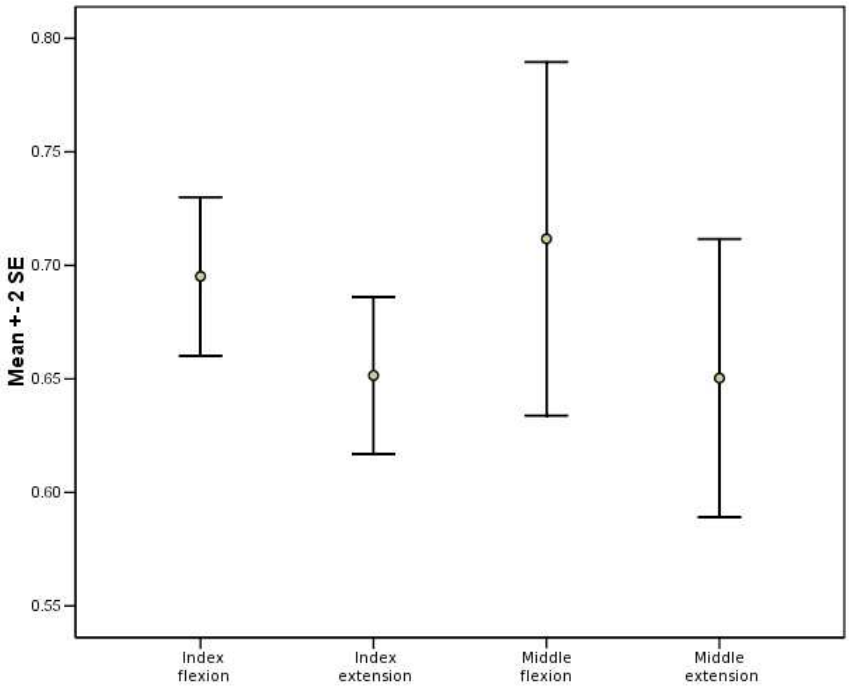


Figure 4.3: Reaction times for the four finger movements. Error bars display standard deviation. Only unambiguous and low ambiguity/Competition trials were considered to allow a clearer assessment of the relationship between movement and RT.

primary conditions should be well-matched. A repeated measures GLM was constructed with movement type (4 levels: index flexion, index extension, middle flexion and middle extension), ambiguity (3 levels) and suppression (2 levels) as within-subject factors. Sphericity was not violated for any factor or interaction. Neither ambiguity nor suppression had significant effects on trial count (ambiguity: $F(2)=0.429$, $p=0.655$, suppression: $F(1)=0.006$, $p=0.941$), therefore bias was not present in this regard.

The four movements do not have different reaction times If participants experienced difficulty in performing one movement in particular - for example middle extension - this could affect their decision to select it, and would certainly alter their reaction time. Significant efforts were made during training to ensure that participants were comfortable with the task. However, a formal assessment of group reaction time would still be prudent. Figure 4.3 shows the RTs for the four movements, considering unambiguous and low ambiguity trials only (to diminish the effects of the experimental factors, whilst still providing sufficient trials for valid analysis).

A repeat measures GLM with a 4-level factor of movement type was constructed. The effect of

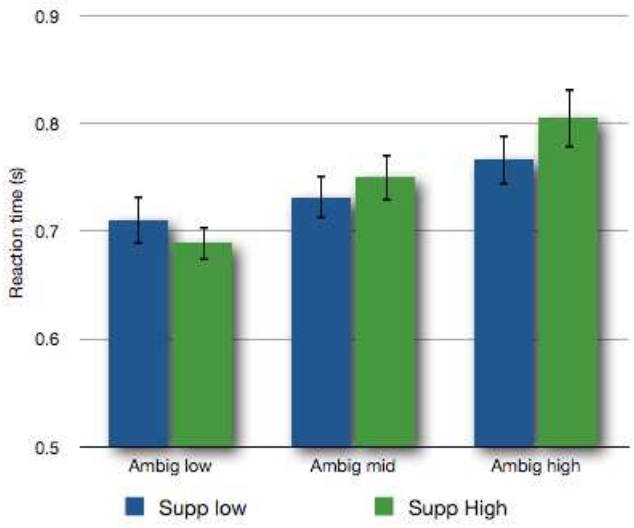


Figure 4.4: The effects of Selectivity and ambiguity/Competition on reaction time. Error bars represent standard error.

movement was not found to be significant ($F(3) = 1.565$, $p = 0.211$), supporting the view that the responses did not inherently differ in their ease of execution.

4.2.1.2 Assessing experimental effects on reaction time

A 3 x 2 GLM with the within-subject factors of Competition and Selectivity was used to analyse the effects the key experimental variables had on RT (Figure 4.4). Assumptions of sphericity were not violated (Mauchley W > 0.696 for both factors and the interaction).

Factor	F	df	p
Competition	22.037	2	0.000
Selectivity	1.357	1	0.262
Interaction	2.502	2	0.099

Only Competition proved significant, having a positive relationship with RT. The effects of Selectivity within each individual Competition level were examined using pairwise t-tests. Again, no effect was observed to a significant level.

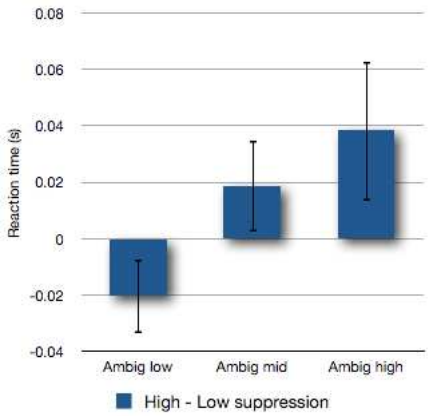


Figure 4.5: The interaction between Selectivity and ambiguity/Competition on reaction time. The graph displays High - Low Selectivity RTs for each ambiguity level. Error bars show standard error.

Competition level	Mean Selectivity (high - low) RT difference (s)	Standard deviation (s)	t	df	p (2-tailed)
Low	0.020	0.052	1.569	15	0.137
Mid	-0.019	0.065	-1.159	15	0.265
High	-0.038	0.099	-1.547	15	0.143

However, the interaction between Competition and Selectivity emerges as a weak trend, as demonstrated in figure 4.5. This suggests that the influence Selectivity has on RT becomes more pronounced as Competition increases.

4.2.1.3 Assessing experimental effects on accuracy

Another 3 x 2 repeated measures GLM was used to assess the influence of Competition and Selectivity on response accuracy. Assumptions of sphericity were close to being violated for Competition (Mauchley's W = 0.657, p = 0.053), with the epsilon for Huynh-Feldt correction being the highest (0.808), which was applied to the statistics reported here.

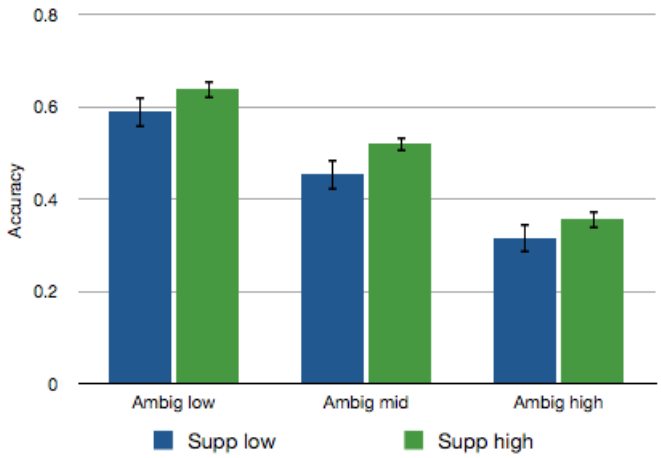


Figure 4.6: The effects of Selectivity and ambiguity/Competition on response accuracy. Error bars show standard error.

Factor	F	df	p
Competition	145.020	2	0.000
Selectivity	3.294	1	0.090
Interaction	0.457	2	0.638

As expected, Competition had a significant effect on accuracy, with Selectivity only providing a weak trend. Pairwise comparisons of Selectivity levels within each stage of ambiguity revealed no significant effects.

Competition level	Mean Selectivity (high - low) accuracy difference	Standard deviation	t	df	p (2-tailed)
Low	0.048	0.108	1.788	15	0.094
Mid	0.066	0.148	1.795	15	0.093
High	0.039	0.135	1.156	15	0.093

4.2.2 Functional imaging results

4.2.2.1 Colour hue level

This analysis was restricted to early visual cortical areas, as defined by the anatomical ROI for colour processing (see 4.1.4.4). Figure 4.7 displays the SPM for the colour hue variance

F-test, with the threshold set at $p < 0.05$ FDR-corrected. An area in the right lingual gyrus was found to be significant, with the contrasts estimates for increasing hue level indicating a positive relationship (figure 4.8).

This is evidence that an early visual area is encoding the ”raw” hue levels of the stimuli, which must then be drawn upon to guide decision-making within the context of the experimental paradigm.

4.2.2.2 Accumulation of sensory evidence

The colour hue variance analysis looked at stimulus-driven changes in brain activation in a deliberately non-specific way, that is, in a manner detached from the decision aspect of the task. Following this information along on its transformation into the appropriate motor response, the next step was to assess the sensory input along the dimensions defined as relevant by the decision-making task. Unlike other perceptual decision-making paradigms, this task did not use stimuli encoded by spatially-separable brain regions (e.g. faces vs houses [335]). Instead, decisions were prompted by the single visual property of colour. Given the nature of the EPI fMRI protocol used for this experiment, it was not feasible to attempt to directly differentiate the two competing sensory signals. Therefore, to evaluate how colour-encoding areas were representing the stimuli, trials were considered in terms of the dominant value, i.e. the value of the correct response, as demonstrated in the table below. In this context, value is therefore the inverse of ambiguity.

Ambiguity	Unambiguous	Low	Mid	High
Colour 1 hue level	6	5	4	3
Colour 2 hue level	0	1	2	3
Correct relative value	4	3	2	1

The first level contrasts of each value level vs the checkerboard visual control condition were taken into a random effects ANOVA, with four dependent groups defined by those value levels. The effect of increasing value was tested using a parametric T test. Figure 4.9 displays whole-brain effects for increasing stimulus value. One area survives a cluster threshold of $P < 0.05$: the left superior occipital/inferior parietal cortex (figure 4.10).

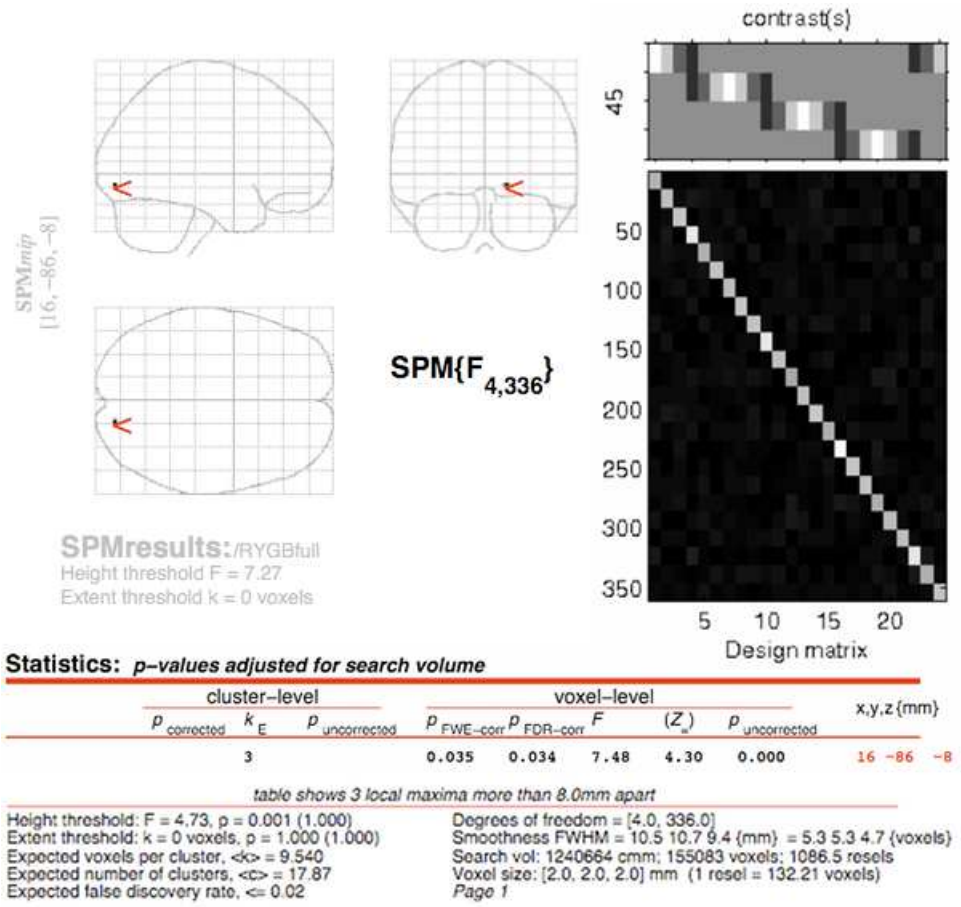


Figure 4.7: The second-level analysis of colour hue variance. Each column of the design matrix corresponds to a subset of blends, listed in colour order, transitioning between red, yellow, green, blue, and finally back to red, in the manner described in figure 4.1. Colour hue variance was then assessed in an F test that separately and parametrically varied each colour's hue levels. SPM threshold: p = 0.05 FDR, corrected for the calcarine, lingual, fusiform and inferior occipital gyri (defined by the AAL atlas). A small cluster with peak 16 -68 -8 MNI is present in the right lingual gyrus, F(4,336) = 7.48, p = 0.034 FDR.

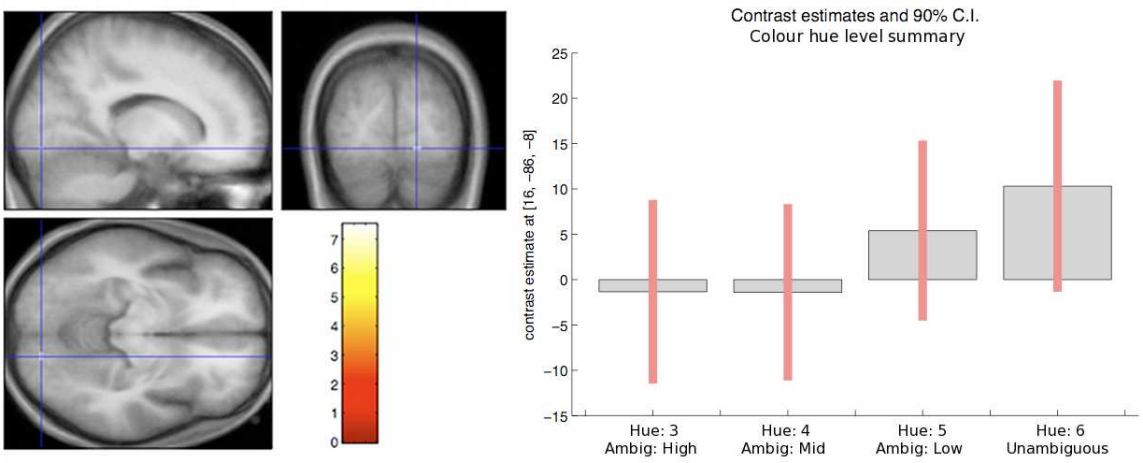


Figure 4.8: Localising the visual cortical region that responded to colour variance: right lingual gyrus. A cluster with its peak at 16 -68 -8 MNI has a significance of p = 0.034 FDR, corrected for multiple comparisons within the *a priori* colour processing ROI. Contrast estimates show that this activation increases with increasing hue level.

A priori ROI analysis Previous studies have brought to light a number of replicable areas that subserve the role of representing visual stimuli in terms of their perceptual value (see 4.1.4.4). To maximise this analysis' sensitivity for these effects, the T test examining parametrically increasing stimulus value was then restricted to within the *a priori* ROI (figure 4.11). This revealed a significant bilateral response in a bilateral region bordering occipital and parietal cortices; and activation in middle temporal gyrus (figure 4.12).

Summary These two key areas of superior occipital/inferior parietal cortex and the middle temporal gyrus both demonstrate activation consistent with the accumulation of colour-defined evidence. When plotting the parameter estimate over different value levels, a gradually-increasing relationship is evident. These areas are commonly conceptualised as being of a higher order within the visual system [335], and so it can be expected that they represent stimuli in a more abstracted and task-determined manner.

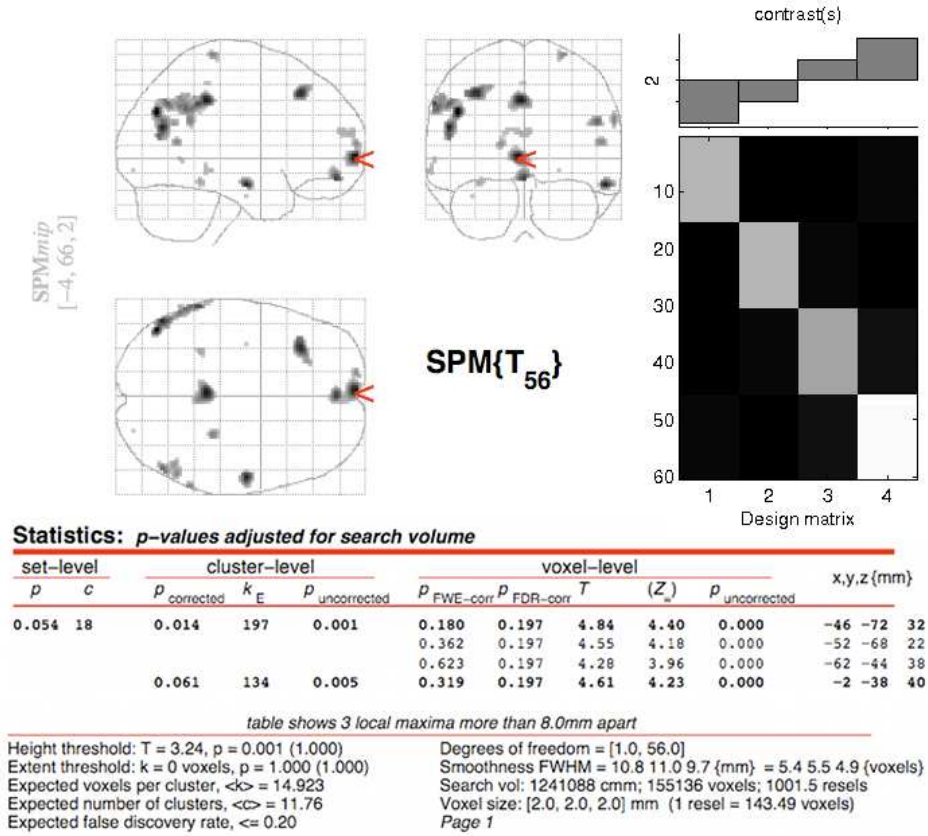


Figure 4.9: The second-level analysis of evidence accumulation. Each column of the design matrix corresponds to a level of "value" of the correct answer for a given ambiguous cue. This is the inverse of ambiguity. The effects of increasing trial value were examined in a parametrically increasing T test. The SPM displayed here was thresholded at $p < 0.001$ uncorrected for the whole brain volume. Clusters having a corrected $p < 0.05$ are reported as significant, with those being $p < 0.1$ reported as interesting trends.

Brain Area	Hemisphere	X	Y	Z	Voxels	t value	Corrected P	Level of significance
Parieto-occipital cortex	Left	-46	-72	32	197	4.84	0.014	Cluster extent (WB)
Parieto-occipital cortex	Left	-52	-68	22	10	4.55	0.046	Voxel height (ROI)
	Right	-50	-68	26	1	4.19	0.046	Voxel height (ROI)
Middle temporal gyrus	Right	56	-13	-18	6	4.55	0.046	Voxel height (ROI)

Table 4.1: A summary of areas responding in a manner compatible with sensory accumulation. WB indicates corrected whole brain significance; ROI indicates region-corrected significance.

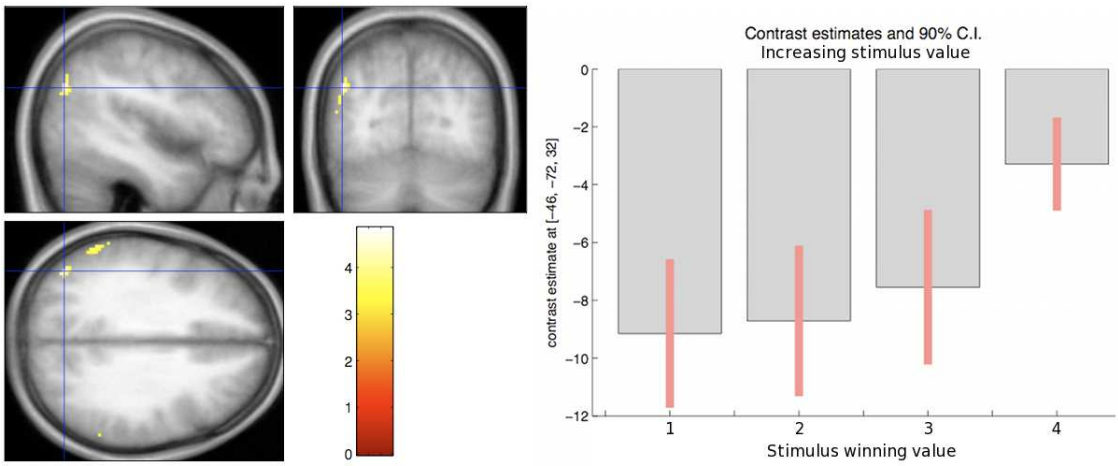
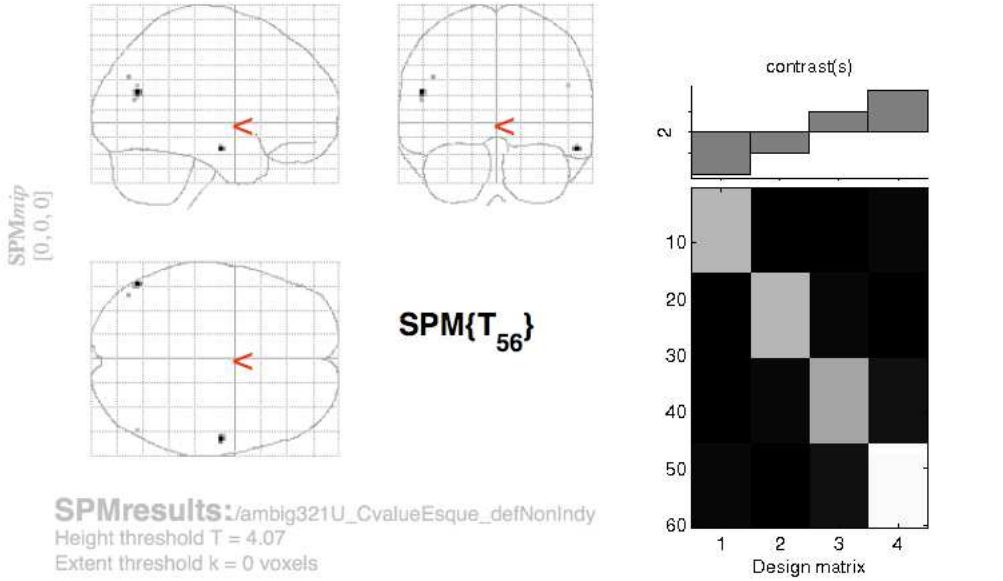


Figure 4.10: Sensory accumulation within superior occipital/inferior parietal cortex, peak at -46 -72 32 MNI. The peak voxel survives a whole-brain corrected threshold of 0.05 FDR. Contrast estimates show that this activation increases with increasing hue level.

4.2.2.3 Integration towards a decision variable

The sensory evidence accumulated in the areas described above must now be integrated and compared to allow a winning response to emerge. This transformation of task-relevant sensory information occurs gradually over the course of each decision, with a response being produced once a threshold or boundary has been reached. As described previously in both electrophysiological and functional neuroimaging studies, areas encoding the integrated decision variable represent in part the difference between competing sensory stimuli. In this sense, activation will be greatest for unambiguous trials, gradually decreasing as the colours comprising the blended stimuli become equally represented. In terms of analysing the data for this experiment, the decision variable will be the inverse of ambiguity - exactly as for the analysis of sensory accumulation. The difference between these analyses was simply defined by the locations within which they were focused.

A priori ROI analysis As for the analysis of evidence accumulation, decision variables are believed to be represented in a number of brain regions. The analysis was then restricted within a mask comprising these areas (figure 4.13). This revealed significant activation in the parieto-occipital cortex (an overlap in the sensory accumulation and decision integration ROIs highlights this area in both analyses); anterior medial frontal cortex, dorsal middle frontal gyrus and medial orbitofrontal cortex.



Statistics: ***p–values adjusted for search volume***

set–level		cluster–level			voxel–level					x,y,z {mm}
p	*c*	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	*T*	$(Z_=)$	$p_{uncorrected}$	
0.000	4	0.061	10	0.230	0.069	0.046	4.55	4.18	0.000	−52 −68 22
		0.092	6	0.351	0.070	0.046	4.55	4.18	0.000	56 −10 −18
		0.181	1	0.728	0.180	0.046	4.19	3.89	0.000	−44 −74 32
		0.181	1	0.728	0.221	0.046	4.11	3.82	0.000	50 −68 26

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 4.07, p = 0.000 (0.240)
Extent threshold: k = 0 voxels, p = 1.000 (0.240)
Expected voxels per cluster, <k> = 7.450
Expected number of clusters, <c> = 0.27
Expected false discovery rate, <= 0.05

Degrees of freedom = [1.0, 56.0]
Smoothness FWHM = 10.8 11.0 9.7 {mm} = 5.4 5.5 4.9 {voxels}
Search vol: 94384 cmm; 11798 voxels; 123.3 resels
Voxel size: [2.0, 2.0, 2.0] mm (1 resel = 143.49 voxels)

Figure 4.11: The analysis of evidence accumulation, restricted within the *a priori* anatomically-defined ROI. The design matrix corresponds to that of the whole-brain analysis. Again, the effects of increasing trial value were examined in a parametrically increasing T test. Results are voxel-level thresholded at $p < 0.05$ FDR corrected for the ROI. Significant activation is found in bilateral superior occipital/inferior parietal cortex, and right middle temporal gyrus.

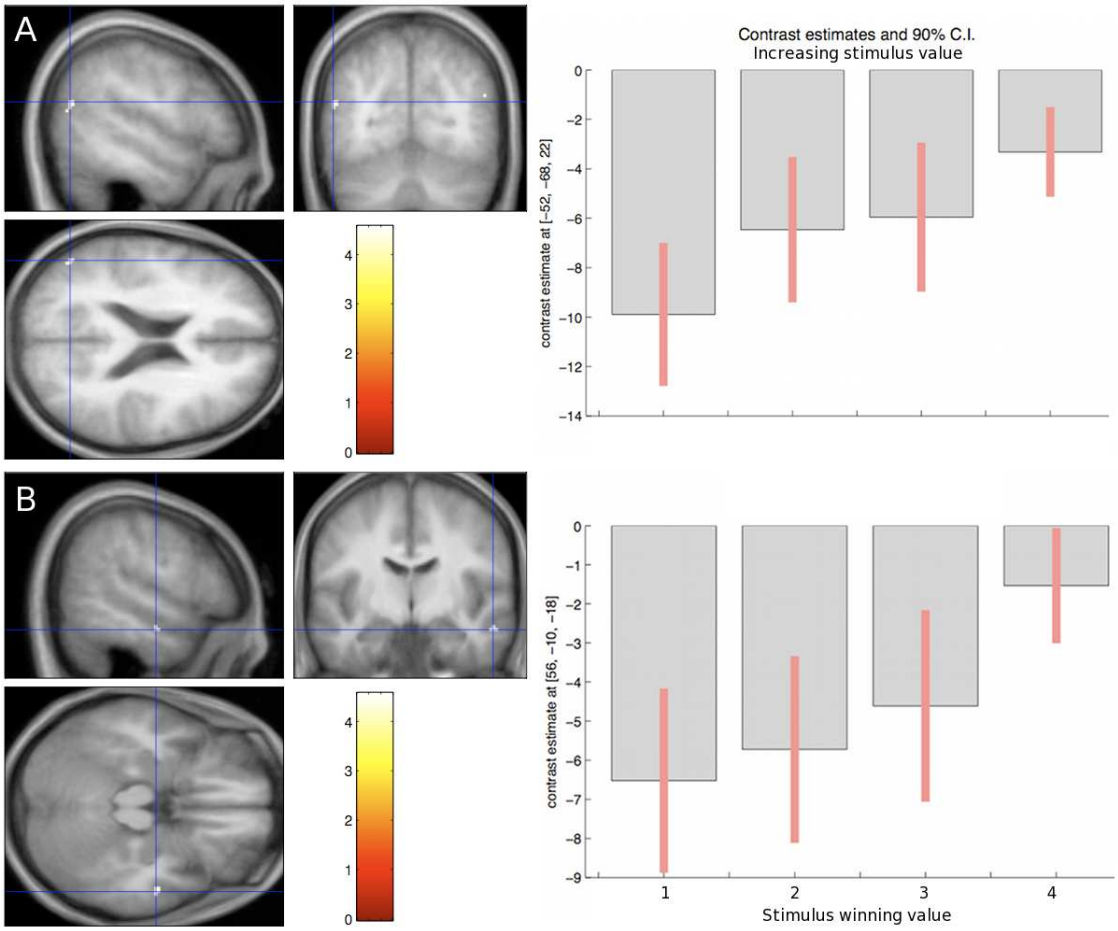
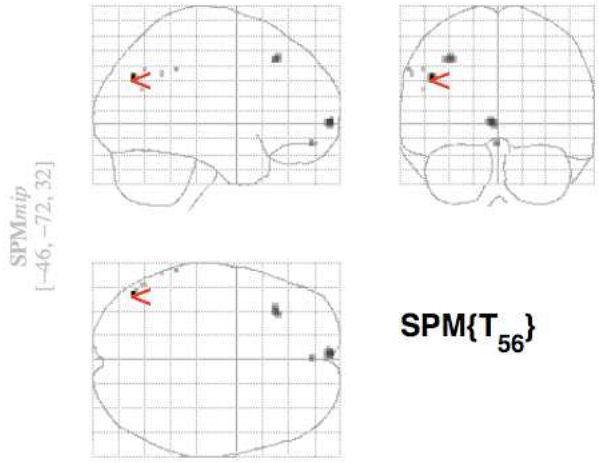
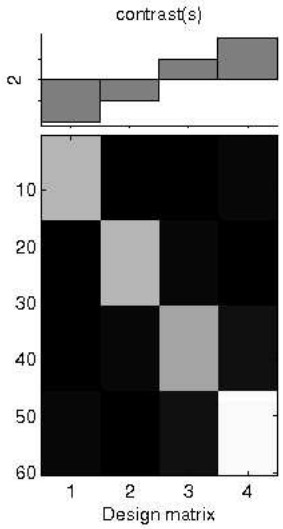


Figure 4.12: Sensory accumulation, isolated within the *a priori* ROI. (A) Bilateral parieto-occipital cortex (-52 -68 22 MNI displayed) p = 0.046 FDR-corrected. (B) Right middle temporal gyrus (58 -10 -18 MNI) p = 0.046 FDR-corrected. Contrast estimates demonstrate increasing activation with winning stimulus value.





Statistics: ***p–values adjusted for search volume***

set–level		cluster–level			voxel–level					x,y,z {mm}		
p	*c*	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	*T*	($Z_=$)	$p_{uncorrected}$			
0.000	8	0.151	10	0.238	0.066	0.049	4.84	4.40	0.000	−46	−72	32
		0.055	23	0.082	0.126	0.049	4.61	4.22	0.000	−4	66	2
		0.051	24	0.076	0.165	0.049	4.50	4.14	0.000	−34	28	46
		0.219	6	0.360	0.246	0.049	4.35	4.02	0.000	0	52	−14

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 4.02, p = 0.000 (0.497)
Extent threshold: k = 0 voxels, p = 1.000 (0.497)
Expected voxels per cluster, <k> = 7.735
Expected number of clusters, <c> = 0.69
Expected false discovery rate, <= 0.05

Degrees of freedom = [1.0, 56.0]
Smoothness FWHM = 10.8 11.0 9.7 {mm} = 5.4 5.5 4.9 {voxels}
Search vol: 310200 cmm; 38775 voxels; 279.6 resels
Voxel size: [2.0, 2.0, 2.0] mm (1 resel = 143.49 voxels)

Figure 4.13: The analysis of decision integration, restricted within the *a priori* anatomically-defined ROI. The design matrix corresponds to that of the whole-brain analysis. Again, the effects of increasing trial value were examined in a parametrically increasing T test. Results are voxel-level thresholded at $p < 0.05$ FDR corrected for the ROI. Significant activation is found in the inferior parietal lobule, medial frontal cortex, middle frontal gyrus and medial orbitofrontal cortex.

Brain Area	Hemisphere	X	Y	Z	Voxels	t value	Corrected P	Level of significance
Parieto-occipital cortex	Left	-46	-72	32	10	4.84	0.049	Voxel height (ROI)
Anterior medial prefrontal cortex	Left	-4	66	2	23	4.61	0.049	Voxel height (ROI)
Dorsal middle frontal gyrus	Left	-34	28	46	24	4.50	0.049	Voxel height (ROI)
Medial orbitofrontal cortex	-	0	52	-14	6	4.35	0.049	Voxel height (ROI)

Table 4.2: Summarising the regions that responded as decision integrators. WB indicates corrected whole brain significance; ROI indicates region-corrected significance.

Brain Area	Hemisphere	X	Y	Z	Voxels	t value	Corrected P	Level of significance
Medial frontal gyrus	Right	14	42	-8	26	< 0.001	0.047	Voxel height (ROI)
Precuneus	-	0	-52	52	145	< 0.001	0.046	Voxel height (ROI)

Table 4.3: Summarising the regions that correlated with the evidence supporting the decision that was actually made. WB indicates corrected whole brain significance; ROI indicates region-corrected significance.

4.2.2.4 Selected response appropriateness

The final stage of considering the data in terms of ambiguous perceptual decision-making was to determine which brain areas correlated with the evidence supporting the responded that was actually chosen, rather than that which the evidence objectively supported. There were no areas that survived whole-brain correction for multiple comparisons. However, two areas (figure 4.15) within the *a priori* anatomical mask were found to be significant (table 4.3).

When considering trials in terms of their stimulus-determined winning ”value”, significant activation is observed in a number of interesting areas. Due to a small partial overlap of the sensory accumulation and decision integration anatomical ROIs, the superior occipital/inferior

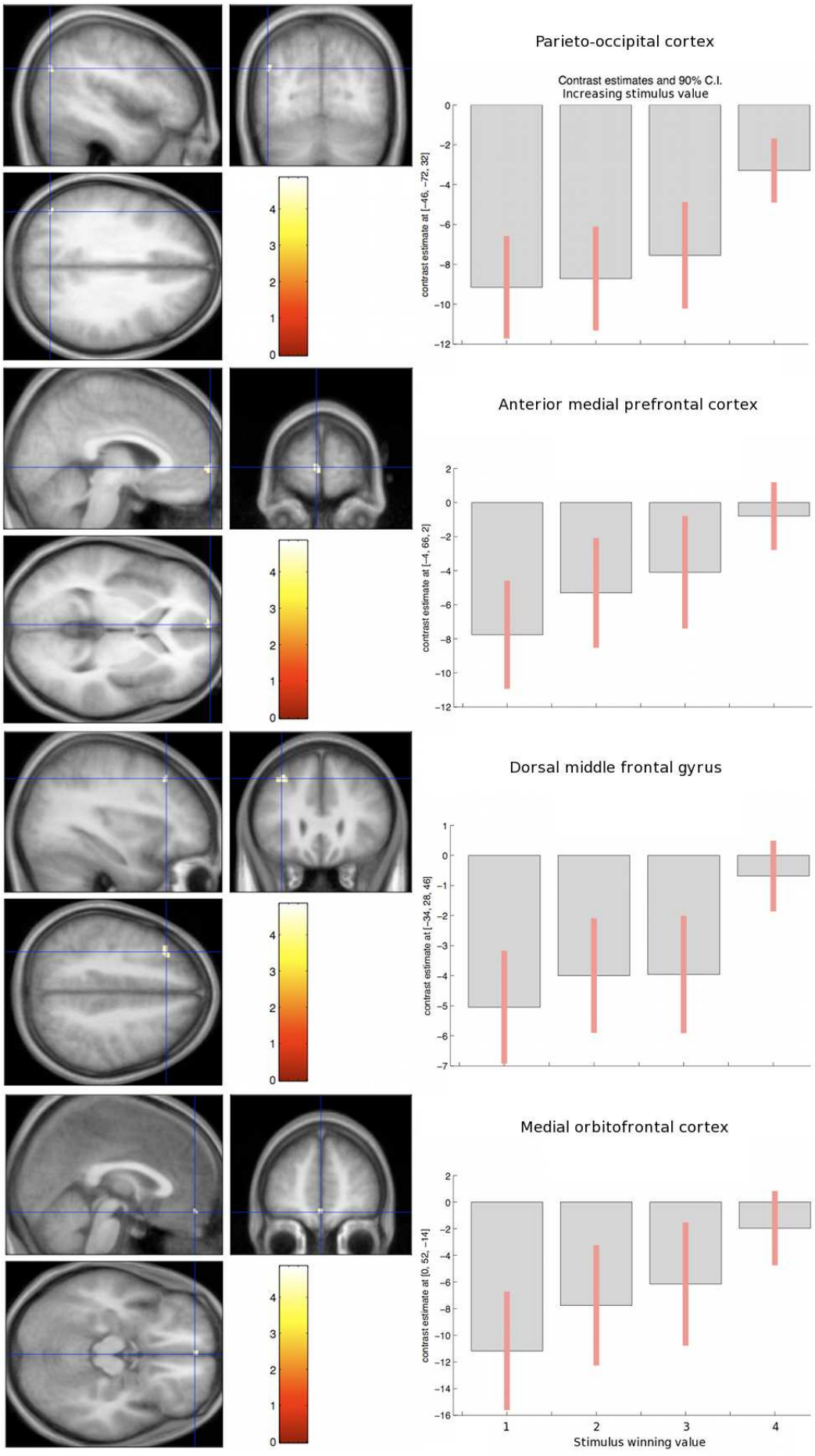


Figure 4.14: Decision integration: significant regions of activation within the *a priori* ROI (voxel-height $P < 0.05$ FDR). Contrast estimates demonstrate increasing activation with winning stim-

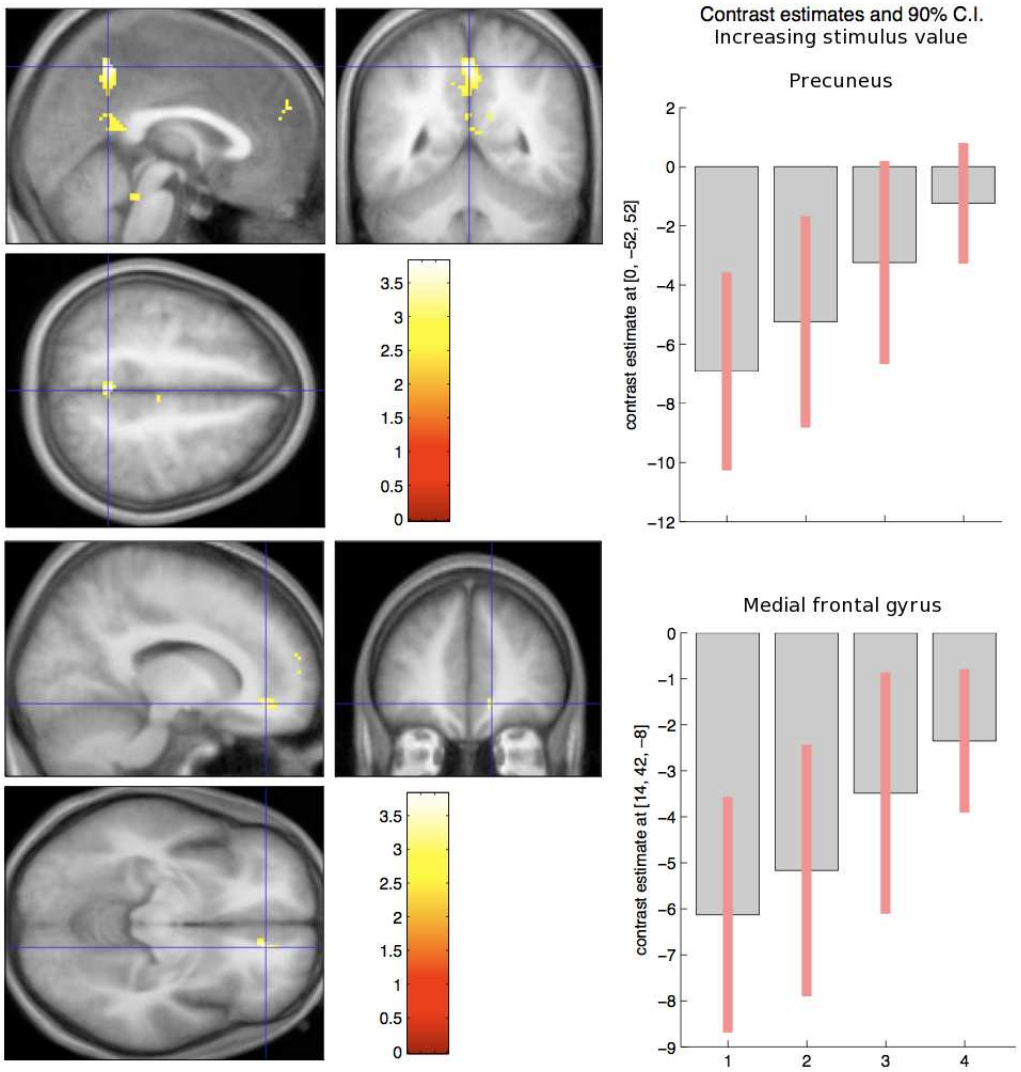


Figure 4.15: Brain regions whose activation correlated with the level of accumulated evidence supporting the actual decision chosen. These were significant after correction for multiple comparisons within the predetermined anatomical ROI.

parietal response seen early is also seen here. Orbitofrontal cortex has long been known to play a crucial role in the evaluation of expected value [19], particularly in tasks involving probabilistically-driven risk or uncertainty. Dorsal middle frontal gyrus has been especially implicated as subserving the later stages of decision-making [335], and the finding here further supports its generalised role across many modalities and tasks.

It is also notable that no striatal activation was observed for these analyses - the computational model anticipates that they essentially represent the summed value of each option. However, this absence could be due to the over-learned nature of the task - the contingencies are well-established by the time the participant is scanned, and the evaluations guiding these decisions could have become downplayed in this highly change-sensitive regions [18].

4.2.2.5 The main effect of Competition

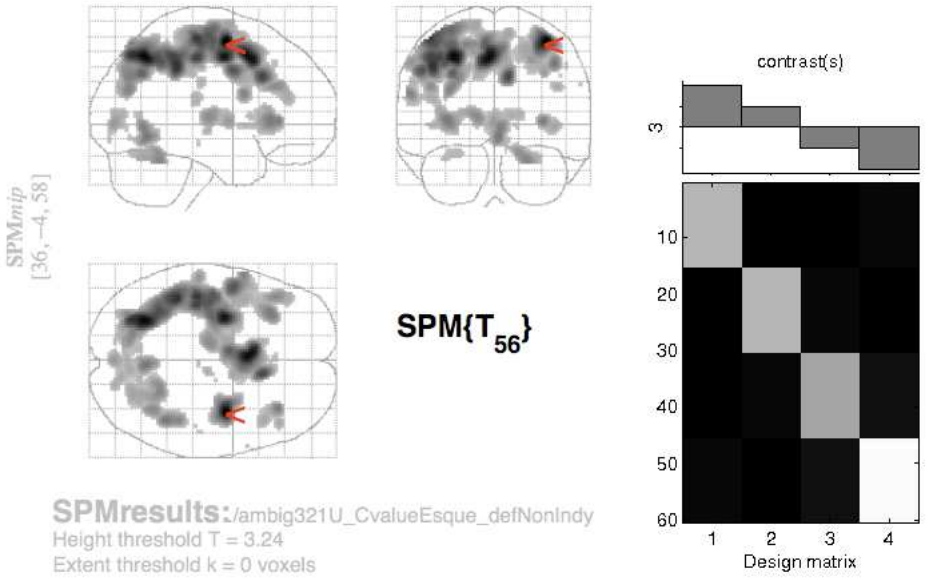
The purpose of this perceptual decision-making paradigm was to attempt to elicit activity within the BG, and determine whether or not it bore a resemblance to the simulated computational model. The experimental paradigm provided the necessary tightly-controlled and, most importantly, the competitive context within which to do this. This degree of Competition was varied as a function of a blended stimulus' distance from a pair of primary cues - i.e. its perceptual ambiguity. The top-down modulatory effects that cortical areas excited by ambiguity impose are also a key component of the general perceptual decision-making mechanism.

Effectively the inverse of winning stimulus value as noted above, this analysis was performed using the same second-level design matrix constructed for the analysis of sensory accumulation and decision integration. The contrast to assess the effects of increasing ambiguity was simply the inverse of the parametric t test used earlier (figure 4.16).

The main effect of Competition was observed within several cortical areas, most notably the inferior and middle frontal gyri (4.17, 4.4).

4.2.2.6 The main effect of Selectivity

In the High Selectivity condition, stimuli elicited a decision that forced participants to choose between two mutually exclusive movements. In the Low Selectivity condition, choices were



Statistics: ***p–values adjusted for search volume***

set–level		cluster–level			voxel–level					x,y,z {mm
p	*c*	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	*T*	$(Z_=)$	$p_{uncorrected}$	
0.002	23	0.000	557	0.000	0.000	0.015	7.65	6.30	0.000	36 −4
		0.000	7253	0.000	0.000	0.015	7.30	6.10	0.000	−28 −62
					0.000	0.015	7.24	6.06	0.000	−4 12
					0.000	0.015	7.10	5.97	0.000	−30 −6
		0.000	754	0.000	0.007	0.015	5.91	5.19	0.000	−50 0
					0.145	0.015	4.92	4.46	0.000	−60 8
					0.248	0.015	4.71	4.31	0.000	−44 30
		0.001	307	0.000	0.070	0.015	5.17	4.66	0.000	4 −60
					0.986	0.015	3.70	3.48	0.000	14 −66
		0.001	357	0.000	0.073	0.015	5.16	4.64	0.000	−32 26
					0.985	0.015	3.70	3.49	0.000	−42 10
		0.012	202	0.001	0.088	0.015	5.09	4.60	0.000	34 32
					0.814	0.015	4.07	3.79	0.000	42 22
		0.003	262	0.000	0.173	0.015	4.85	4.41	0.000	22 −50 −
					0.894	0.015	3.96	3.70	0.000	28 −66 −
		0.001	302	0.000	0.242	0.015	4.72	4.31	0.000	−12 −16
					0.486	0.015	4.41	4.07	0.000	−28 −4
					0.817	0.015	4.07	3.79	0.000	−22 −14
		0.006	237	0.000	0.470	0.015	4.43	4.09	0.000	0 −82
					0.920	0.015	3.91	3.66	0.000	12 −78
					1.000	0.015	3.39	3.22	0.001	0 −86 −

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 3.24, p = 0.001 (1.000)
Extent threshold: k = 0 voxels, p = 1.000 (1.000)
Expected voxels per cluster, <k> = 14.923
Expected number of clusters, <c> = 11.76
Expected false discovery rate, <= 0.01

Degrees of freedom = [1.0, 56.0]
Smoothness FWHM = 10.8 11.0 9.7 {mm} = 5.4 5.5 4.9 {vc
Search vol: 1241088 cmm; 155136 voxels; 1001.5 resels
Voxel size: [2.0, 2.0, 2.0] mm (1 resel = 143.49 voxels)
Page 1

Figure 4.16: The second-level analysis of Competition/ambiguity. Each column of the design matrix corresponds to a level of ”value” of the correct answer for a given ambiguous cue. This is the inverse of ambiguity. The effects of increasing ambiguity were examined in a parametrically decreasing T test. The SPM displayed here was thresholded at $p < 0.001$ uncorrected for the whole brain volume. Clusters having a corrected $p < 0.05$ are reported as significant, with those being $p < 0.1$ shown as interesting trends.

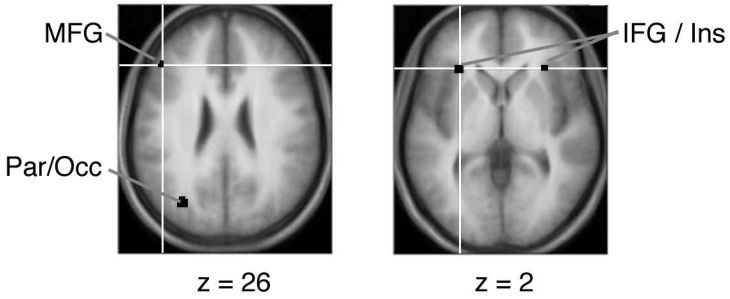


Figure 4.17: Cortical activation and Competition/ambiguity. Voxels survived a familywise error (FWE) correction for multiple comparisons at a whole-brain level of $p < 0.05$. Details of peak responses are shown in 4.4. IFG: inferior frontal gyrus; Ins: insula; MFG: middle frontal gyrus; Par/Occ: parietal-occipital border.

	x	y	z	t	p (uncorrected)	p (FWE-corrected)
High > Low ambiguity						
L temporo-occipital	-32	-72	26	7.02	< 0.001	0.001
L inferior frontal gyrus/anterior insula	-30	28	2	6.00	< 0.001	0.011
R inferior parietal lobule	46	-46	40	5.92	< 0.001	0.014
R inferior frontal gyrus/anterior insula	34	28	0	5.91	< 0.001	0.014
L middle frontal gyrus	-46	30	26	5.58	< 0.001	0.037
Medium > Low ambiguity						
L superior temporal gyrus	-44	-44	-12	5.94	< 0.001	0.013

Table 4.4: MNI coordinates for cortical regions responding to Competition/ambiguity.

between movements for which there was no mechanical reason to prevent simultaneous execution. It was hypothesised that the High condition would recruit suppressive circuits to a greater degree than for Low. Motor responses were counterbalanced between the two levels, precluding experimental effects that might be attributable to somatotopical interactions or low-level variations in motor performance [402].

Whole brain analysis The first-level t-test contrasts for the main effect of Selectivity (High > Low) were entered into a second-level random effects t-test. The resultant SPM was thresholded at p = 0.001 uncorrected for the whole brain volume.

Suppression significantly activated the globus pallidus Voxel-wise analysis revealed a strong effect of suppression within left pallidus (t=7.84, p=0.001 corrected 4.18). This was localised to a ventrolateral posterior region thought to subserve sensorimotor function [460].

No other areas were activated by suppression at the whole brain level When considering significant activity at a whole-brain level, no other regions were activated even at the liberal corrected threshold of p<0.1. This may seem contrary to expectation, as several cortical regions including anterior cingulate (ACC) [461, 462], dorsal premotor [463], presupplementary motor and inferior frontal [32, 318] cortices have been suggested to play a role in selection or suppression. Small volume corrections were therefore performed within these regions to identify if our result could be explained due to our focus on the BG, but again found no activation at p<0.1 corrected. The results indicate a strong and specific effect of suppression within globus pallidus, agreeing with the first computational model-derived hypothesis.

A priori regions of interest analyses No region was significantly activated within the more general ROI.

Summary The key result established at this point was the differential effect Selectivity had within globus pallidus (4.18). This finding suggested that the paradigm was indeed capable of eliciting pallidal activation in the manner predicted by the disinhibitory view of BG-mediated action selection.

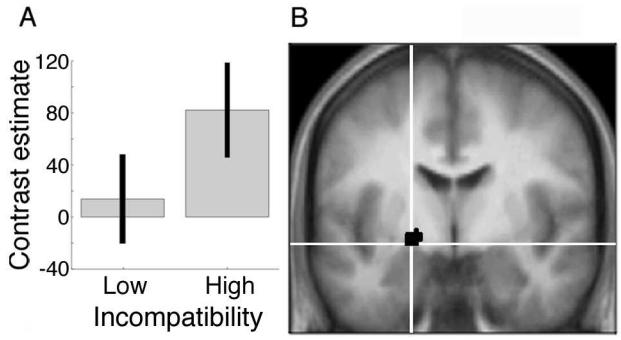


Figure 4.18: The effect of suppression within globus pallidus. (A) Contrast estimates across the two levels of suppression. Bars indicate 90% confidence intervals. (B) Localising the effect of suppression within left pallidus ($p < 0.05$ FWE-corrected within bilateral pallidus). Crosshairs show peak activation at -20 -6 -8, $t = 7.84$, $p = 0.001$ corrected.

4.2.2.7 The Selectivity x Competition interaction

To elaborate on the finding concerning the main effect of Selectivity within pallidus, its interaction with Competition was examined. A random effects analysis was performed within a 3-level ANOVA, having dependent, nonspherous factors. Separate regressors of High > Low Selectivity were used for each Competition level. This was found to broadly match prediction derived from the basis Gurney model (figure 4.19). The next step was to establish how Selectivity-related pallidal activity varied with increasing Competition. Given the prediction of a diminished response at intermediate levels, the first contrast examined was the High > Low Selectivity, constrained to the lowest level of Competition. Within the left pallidal mask, we found one significantly activated region (-16 -8 0, $t = 3.53$, $p = 0.032$ corrected). Contrast estimates were extracted across all three Competition levels, revealing a response that closely resembled the predicted interaction. To determine whether there were any pallidal regions that deviated from this pattern, the Selectivity contrast was isolated to Medium Competition, and a significantly activated area was indeed identified (-20 -4 -4, $t = 3.98$, $p = 0.011$). The extracted contrast values suggested a response that was opposite to our prediction – that the interaction peaked at intermediate uncertainty levels. This provided partial support for the model, although the discovery of a notable exception prompted us to explore the differences between these two pallidal regions.

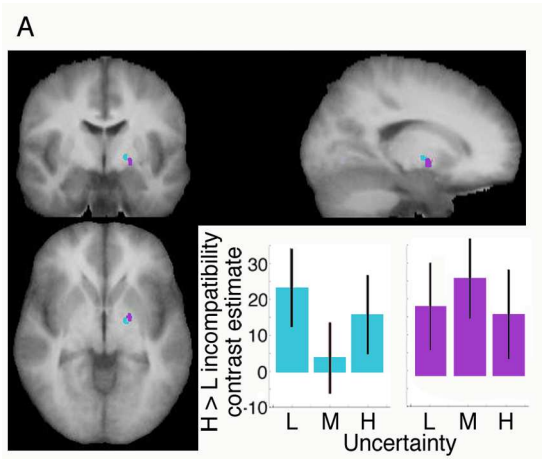


Figure 4.19: The interaction of Selectivity and Competition within separate pallidal subregions. The region in green displays the contrast of (High Selectivity: Low Competition) > (Low Selectivity: Low Competition), with a peak at -16 -8 0, t=3.53, p=0.032 corrected. The associated graph shows the parameter estimates of High > Low Selectivity across each of the Competition levels. This green region displays the pattern of pallidal activation predicted by the model, which encodes responses in terms of their expected values. The purple area shows the contrast of (High Selectivity: Medium Competition) > (Low Selectivity: Medium Competition), peaking at -20 -4 -4, t=3.98, p=0.011 corrected. This region deviates from model expectations, as the interaction is most elevated at intermediate Competition levels.

4.2.2.8 Computational model regressions

As well as performing an analysis that was directed by the explicit experimental manipulation, a more rigorous attempt to test the model itself was also performed. This took each participant's experimental stimuli, and used them to direct a simulation of the model. From this, a regressor was constructed, representing the model's predicted neural activation for the striatum and pallidus.

To allow for more insight to be drawn from this regression, an additional, more general analysis was performed, to establish how the BG were responding to the overall task.This was achieved by constructing a second-level factorial analysis, which included the separate contrasts for each suppress x ambiguity level versus the checkerboard visual control. To isolate areas that maximally responded to the task, an "effects of interest" F test was performed, thresholded at $p < 0.001$ at the whole-brain level, with small volume correction then being applied using the left pallidal and putamen ROIs defined by the AAL atlas, at a FWE level of $p < 0.05$. Parameter estimates for the conditions of interest were then examined, to be used to inform subsequent model modifications.

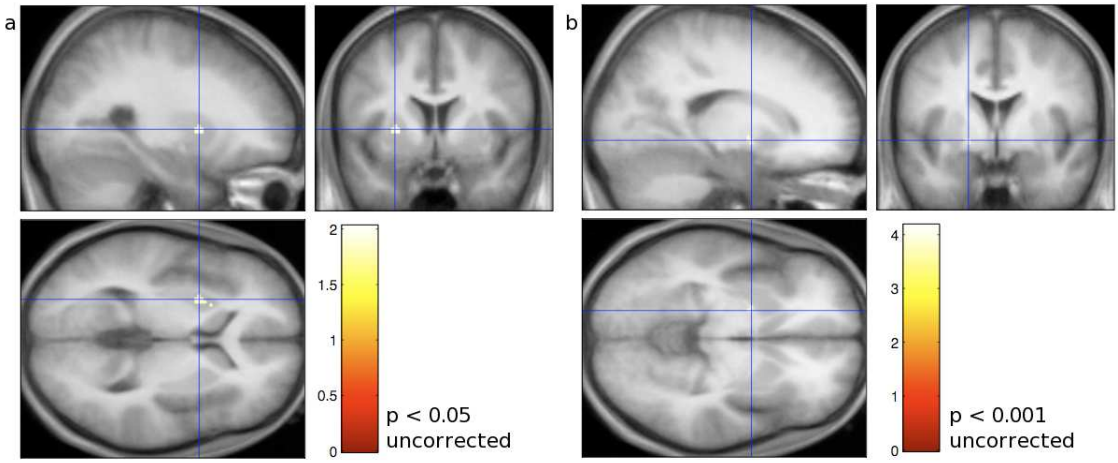


Figure 4.20: Areas of activation that correlated with computational model-derived regressors of predicted striatal and pallidal activity. (a) Left putamen, MNI -26 6 4, uncorrected p = 0.03, FWE-corrected within left AAL putamen p = 0.723. (b) Left pallidus, MNI -18 0 -4, uncorrected p < 0.000, FWE-corrected within left pallidus p = 0.023.

Striatal regressor Unfortunately, the model provided a poor account of striatal activation (figure 4.20a). The uncorrected threshold of p < 0.05 revealed an active area of anterior putamen. However, it had a FWE-corrected significance of p = 0.723. Figure 4.21a demonstrates the actual response in left putamen. One of the basic model assumptions is that suppression does not have an effect within this area, which the data clearly contradicts. In addition, it appears that the overall response increases with respect to increasing Competition, whereas the model predicts that activity should drop off under such conditions. Therefore it is clear that the model must be modified to accommodate these observations: that there is an active suppressive process taking place within the striatum, which is engaged by both Selectivity and the interaction with ambiguity.

Pallidal regressor As indicated by the previous analyses of explicit Selectivity, pallidal activation was well-described by the regressor directly derived from the computational model (figure 4.20b). A significant region was isolated at -18 0 -4, with t(15) = 4.17, p = 0.023, FWE-corrected within AAL left pallidus. However, the account is not complete. The model predicted that pallidus would dip at intermediate levels of Competition: as figure 4.21b shows, this is clear not the case. This is likely to be a downstream effect of the model's poor approximation of striatal activity, and so will itself be the subject of further model modification.

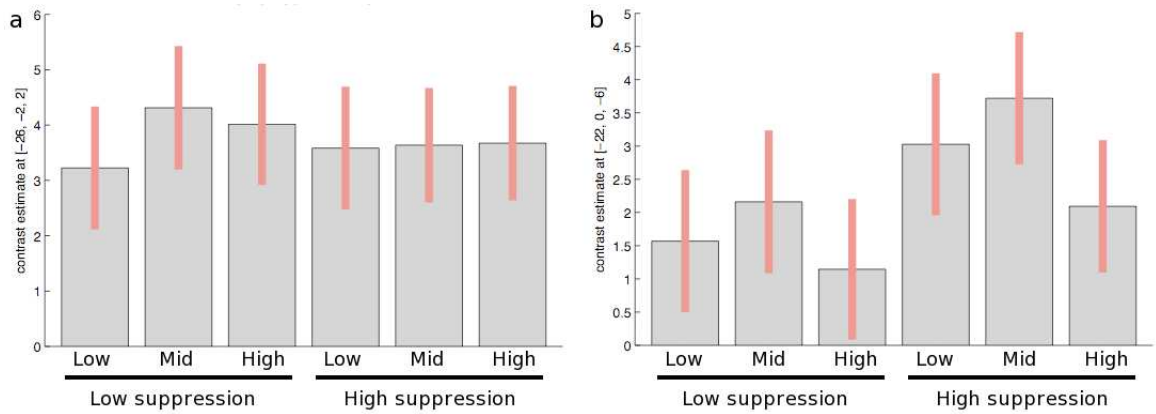


Figure 4.21: Experimentally-determined BG responses to Competition (low, mid and high), and Selectivity (low and high). (a) Left putamen and (b) left pallidus. These values were extracted for each level of Competition x Selectivity, versus the checkerboard control condition. Peak voxels were isolated using an effects of interest contrast within a factorial ANOVA involving all the experimental conditions, thresholded at $p < 0.001$ uncorrected for the whole-brain volume, then small volume corrected within the AAL-defined left putamen or pallidus, to a FWE level of $p < 0.05$.

4.2.2.9 Effective connectivity

The analysis of the Selectivity-Competition interaction within pallidus 4.2.2.7 revealed two zones showing separate responses. To determine how the influences experienced by these so-called model-concordant and model-deviant areas differed, a connectivity toolbox was developed to assess effective connectivity. Using vector autoregression, directional measures of Granger causality were described between these pallidal areas and the rest of the brain. Due to the large scale of this task, and amenability to parallel analysis, it was optimised to run on ECDF's EDDIE cluster. It was therefore possible to demonstrate that the model-deviant area received a significantly greater input from left inferior frontal cortex.

To gain a greater understanding of the cortical architecture driving the model-concordant and model-deviant sections of the pallidus, their peak activation coordinates were used to seed bivariate Granger causality analysis against the rest of the brain. By determining the model-deviant > model-concordant connectivity map contrast, it was possible to identify those areas that made a greater contribution to the model-deviant region. These were the left inferior frontal cortex, insula, putamen and amygdala (figure 4.22). Given our bivariate approach and current anatomical information underlying the pathways involved [464], a likely interpretation is that putamen receives input from the implicated cortical regions, which is then fed forward to the

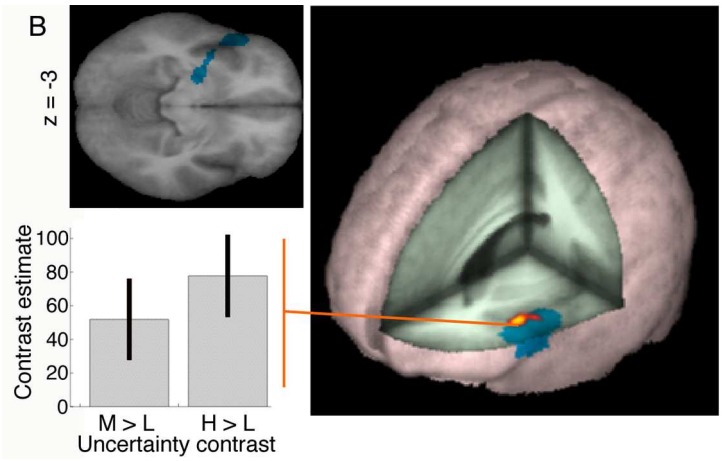


Figure 4.22: Granger causality analysis. On the right, a 3D projection in blue of those areas exerting a significantly stronger influence (P<0.05) over model-deviant (-20 -4 -4), compared to model-concordant (-16 8- 0) pallidus. Overlaid on this in red is an area that significantly responded in a High > Low uncertainty contrast masked within this connectivity-delineated area. The graph on the left displays the parameter estimates extracted for the uncertainty contrasts. Top left displays an axial view of the connectivity map, showing its inclusion of insula and putamen.

pallidus.

Our next goal was to determine the quality of this connectivity by assessing these cortical regions' responses to increasing Competition. The contrast of High > Low Competition was constrained within a mask defined by the significant areas of connectivity: only left inferior frontal cortex/anterior insula survived the corrected threshold. The response to Competition within this areas was non-linearly increasing, perhaps reflecting an optimization allowing for speedier decision-making [465]. Having isolated an additional frontal network that is especially engaged with increasing uncertainty, this finding was used this to inform further model modification.

4.2.2.10 HRF peak/reaction time correlations

As Competition was found to not affect RT, mean RTs were calculated for each Competition level, for each subject. These were then correlated with the time takes for the respective HRF to peak for both the model-concordant and model deviant pallidal regions. Significant positive correlations were found between all region-to-RT comparisons (figure 4.5).

However, the strength of these correlations shifted as Competition increased: they became

Competition	Model concordant pallidus	Model discordant pallidus
Low	0.664 (0.005)	0.567 (0.022)
Medium	0.518 (0.040)	0.641 (0.007)
High	0.539 (0.031)	0.663 (0.005)

Table 4.5: Correlations between the mean RTs for each Competition level, and the time taken for the HRF to peak in the model concordant and discordant pallidus regions. Values in brackets are p values.

weaker in the model-concordant region, whilst becoming stronger in the model-deviant area 4.23. The activation dynamics of the deviant (and hence hypothetically) overlapped pallidal region were therefore becoming more tightly coupled to behaviour as this cortico-basal loop became increasingly recruited to resolve the Competition.

4.3 Conclusions

The key fMRI findings of this first study were:

- The task engaged a hierachy of cortical regions repeatedly implicated in perceptual decision-making tasks. An early visual cortical area (right lingual gyrus) encoded the colour hue intensity of the presented stimulus. Later visual regions (superior occipital/inferior parietal and middle temporal cortex) showed activation consistent with the accumulation of sensory evidence. Finally, several frontal regions (medial, middle frontal and orbitofrontal cortex) showed activation suggestive of information being integrated towards a decision, replicating previous perceptual decision-making tasks [335], further supporting the crucial role middle frontal cortex in particular plays in general decision-making.
- Competition evoked a strong response within inferior frontal gyrus (IFG).
- Selectivity profoundly activated the left pallidus, in agreement with hypothesis 3.1.3.1. Although not achieving significance, parameter estimates extracted from pallidus were supportive of hypothesis 3.1.3.2. This was found to be centred in slightly different regions for trials of low vs high Competition, with a more dorsomedial area showing the activity predicted by hypothesis 3.1.3.3. Granger causality analysis found that this shift also coincided with a greater contribution from the IFG.

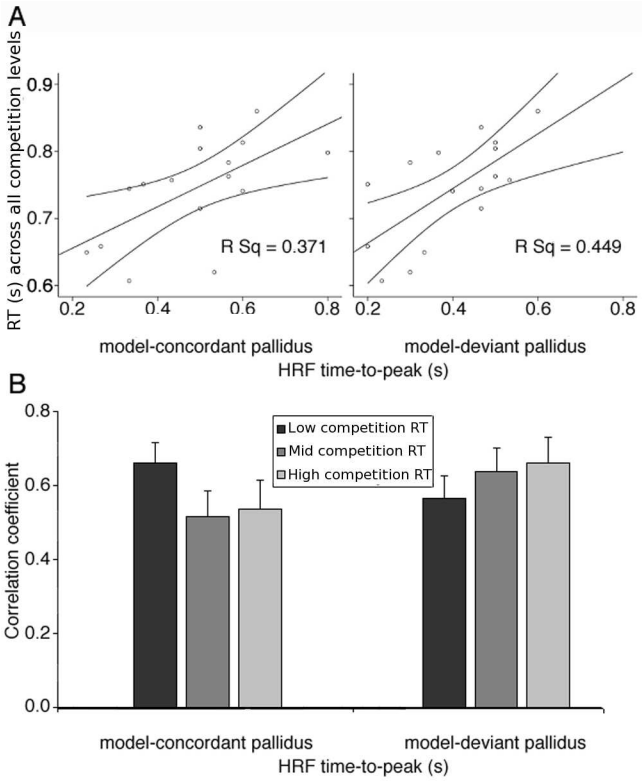


Figure 4.23: Correlating haemodynamic and behavioral response latencies. (A) Scatter graphs displaying summarised correlations between the HRF time-to-peak for each pallidal subregion, and reaction time (averaged across all Competition levels). (B) HRF-RT correlations vary with Competition. In the model-concordant pallidal region, the relationship drops off as Competition increases, but becomes more tightly coupled in model-deviant pallidus. Significance results are shown in 4.5.

- Regressors derived from simulations of the computational model were able to provide a good account for activity seen within pallidus, therefore the model's implementation of a selection mechanism was theoretically consistent with the fMRI responses elicited by this task of action selection. However, the model was less able to account for striatal activity, lending little support to hypothesis 3.1.4.2.

It was also confirmed that the task was not intrinsically biased in terms of the selection and execution of its motor responses, and elicited the expected slowing of RTs with respect to Competition, supporting its further use in subsequent studies.

This study attempted to iteratively use computational modelling and functional neuroimaging to directly test BG-mediated action suppression in terms of Selectivity and Competition. It adds to the understanding of decision-making neurophysiology by providing a parsimonious means of integrating distributed decision variables across cortico-BG channels. The specific response of pallidus to Selectivity supports the underlying theory that the BG mediate selection via action suppression, further reinforcing the known relationship between pallidal dynamics and the point-of-decision [466]. This was taken further by establishing that the interaction between Selectivity and Competition partially agreed with hypothesis 3.1.3.3 derived from the biologically-grounded computational model. A ventral-anterior pallidal area deviated from theoretical expectations, and by employing anatomically-unconstrained effective connectivity analysis, was found to be the recipient of an augmenting uncertainty-encoding frontal loop. In a manner resembling a dual controller [467], there appeared to be a loop contribution specifically to high Competition/ambiguity decisions, originating from the IFG. As Competition increased, the dynamics within the pallidal region to which IFG was projecting developed a stronger correlation with behavioural response, whilst the model-concordant regions of pallidus that lacked this input demonstrated a concomitant decoupling. This therefore demonstrated a Competition-driven activation shift towards a more frontal cortico-BG loop, and explored how this translated into a similarly task-dependent relationship with behavior.

Selectivity-related suppression may be resolved intracortically, in part due to innate oppositional arrangements within the musculoskeletal system. However, the task used here conceptualised these movements as the products of different action plans, called into explicit competition. A differential cortical response to incompatibility was not found, even with the application of small volume correction and liberal thresholds. Although one cannot infer an absence of effect, the results suggest that cortical regions do not directly and specifically mediate inter-competitor suppression. Response conflict is known to engage ACC [461], a structure believed to perform conflict monitoring [468]. At first glance, it is perhaps surprising that a

Selectivity/Competition effect is not seen within ACC. However, evidence suggests that the ACC plays a more evaluative than regulatory role [469], updating motor strategies in order to improve subsequent performance [470]. As our participants were over-trained, and received no trial-by-trial feedback, there is no context within which to perform such a function. Previous investigations concerning response conflict and inhibition have not consistently shown pallidal activation ([471], though see also [32, 472]). This could be down to the choice of task: Go/No-Go, Flanker and Stroop paradigms involve either the suppression of a single motor plan, or the selection of one action in the presence of distractors. These indirect competitions occur over a mixture of perceptual, planning and motor execution levels. By contrasting mutually exclusive and mechanically compatible movements, it has been possible to focus on the central role that the BG and particularly the pallidus play in action selection. These constraints offer an improvement on studies using the imbalanced motor responses of single versus multiple simultaneous finger movements [472].

Here, the finding that IFG responds to uncertainty was replicated [473, 474, 334, 475]. The use of Granger causality analysis demonstrated the specific contribution IFG made to model-deviant pallidus. IFG is believed to act as a forward model generator [476] contributing to cognitive planning by producing internal goal-oriented models [477]. Increased IFG activity therefore reflects the complex goal landscape associated with ambiguous decisions [478, 479, 480, 476]. It is interesting to note that withholding an initiated response, for example during ‘Stop’ tasks, also activates IFG [481, 32]. This may reflect the demand to rapidly update the forward model associated with the consequently-altered goal, thus contributing an augmenting suppressive influence along a “hyperdirect” corticosubthalamic pathway. Indeed, the computational model asserts that Selectivity is dependent on this particular anatomical avenue.

This high-Competition loop is one half of an expanded computational model whose architecture resembles that of a dual controller. These systems are dichotomised into limbic and cognitive components. The BG are well placed to support this architecture, acting beyond the motor domain to select across both cognitive and limbic areas [482, 16]. Daw and colleagues [467] describes these dual limbic and cognitive influences as respectively corresponding to (a) the time-averaged “cached” future reward associated with an action, i.e. a habituated reward representation; and (b) a hierarchical cognitive plan of short-term future response—reward associations. They propose that these systems operate over corticostriatal circuits, the arbitration between which is conducted by ACC, according to their accuracy in predicting reward. However, our results suggest that an explicit switching between two segregated systems is not necessary, and that rather their joint contribution can be utilized within an area of overlap

within pallidus. Competition therefore drives a continuous shift in emphasis between the limbic and cognitive loops. We also see this shift in the relationship between pallidal activation localization and expressed behavior. Although the haemodynamic time-to-peak and behavioral response time will necessarily co-vary, by considering each Competition level individually, an increase in the strength of that coupling was demonstrated, indicating cognitive pallidus' increasingly dominant role over behavioural variance. This further indicates that IFG's assessment of Competition is not post-hoc, but guides action execution when external information is insufficient.

A similar principle is seen in Haruno and Kawato 2006's heterarchical reinforcement-learning model [483], whose predictions have also been borne out experimentally. Rather than selecting for action, this model uses spiralling nigrostriatal connectivity [295] to propagate limbic- and cognitive-derived dopamine reward predictions to motor areas. This guides the successful selection of rewarding stimulus-outcome associations, thus refining action execution. The dichotomy they describe is implemented in a different but compatible manner to that suggested by the data presented here: theirs may have precedence during the learned acquisition of new associations, exploiting the temporal characteristics of phasic dopamine modulation. The framework described here perhaps predominates during selection within established contexts. Indeed, the BG themselves are not the only nuclei implicated in centralised selection [484]. It remains arguable that several cortical and subcortical systems operate in parallel [485], with the BG acting to bias the refinement of detailed action parameters across multiple domains [298].

Empirically-verified computational models can inform us about the underlying mechanisms and potential treatments of neural pathologies (for example, [486, 487, 250]). As a case in point, the symptoms of schizophrenia can be understood as consequences of spurious (thought disorder, disorganization) or diminished (avolition, poverty of thought) selection. This is further supported by the efficacy of striatal dopamine antagonists in reducing psychotic symptoms, which theoretically alters the BG's selection characteristics [28]. Disturbed subcortical activation [30] and connectivity [488, 314] implicate a cortico-subcortical neurochemical cascade in the pathogenesis of schizophrenia [489]. What is particularly salient is the specific contribution of forward model-encoding IFG to the resolution of uncertain decision making. Deficient forward model generation is seen as a core feature of schizophrenia [490], and IFG has shown both abnormal activation [491] and diminished correlation with task performance in patient groups [492]. This is especially marked in patients with pronounced negative symptoms [493], highlighting the potential consequences of dysfunction within the dual-loop model's crucial cog-

nitive arm. The computational model and behavioral paradigm used here could better interpret the efferent functional and behavioral consequences of IFG dysfunction in schizophrenia.

Chapter 5

Computational modelling: a better account for the data

The original computational model provided a good account for pallidal activation with respect to Selectivity, and to some degree for Competition too. However, it did not properly describe striatal activity. A parsimonious and biologically-grounded elaboration was therefore developed to better account for the experimental data.

5.1 Lateral inhibition in the striatum

The striatum clearly did not respond to Competitiveness in the predicted decreasing manner, nor did it prove to be unmodulated by Selectivity. It seemed that inter-channel competition was itself evoking a suppressive process, resulting in an overall response that increased activity with increasing Competition. The means by which this additional level of suppression could be exerted within the striatum is likely to be in the form of lateral inhibition [494]. Gurney et al. acknowledge the presence of such a process, but analytically describe that their implementation considers active populations in an equivalent manner, regardless of whether they are the elicited directly by cortex, or are the product of some degree of intrastriatal ”preprocessing” [28]. However, this data suggests that it is insufficient to consider the striatum to be implementing a covert filtering process, and that an explicit implementation of lateral inhibition is required.

5.1.1 Competitive processing

In simple engineering models of neural networks, winner-take-all (WTA) is a simple means to implement competitive selection using recurrent connections [495]. After learning, the nodes of the network inhibit each other, while exciting themselves: given a series of inputs, the network settles into a state where only the node with the greatest input is active. Other, less severe variations exist where more than one node is able to remain active, in k-winner-take-all or softmax implementations [496]. It has been proposed that a form of WTA takes place within the striatum [497, 498, 499, 482], as striatal MSNs display extensive local axon collateralisation [500], inhibiting neighbouring MSNs [501]. However, electrophysiological evidence suggests this lateral inhibition is not particularly potent [502], or at least is not the dominant form of intrastriatal inhibition [503]. MSN axonal collaterals do not appear to extend beyond the target fields of their innervating corticostriatal projections [500, 504], which would be a necessary feature of functional intrastriatal competition. This has led some to conclude that traditional, lateral inhibition-based winner-take-all behaviour is impossible within striatum [260, 505, 506], and that such connections subserve another function, perhaps modulating dendritic synaptic plasticity.

Although 95% of the striatum's composition has been estimated to be MSN, it does contain three known species of GABAergic interneuron, as well as large aspiny cholinergic interneurons [507, 508, 509]. The most common are parvalbumin-containing GABAergic FSIs. They have dense axonal fields that range beyond the extent of their own dendrites [509], and have been shown to connect to neighbouring FSIs via gap junctions, potentially allowing them act as network-wide inhibitors [510]. FSIs exert a profound inhibitory influence on MSNs for two reasons: they synapse extensively on MSNs' soma and proximal dendrites [511, 503, 512]; and they receive a particularly dense innervation from excitatory corticostriatal sources, often in the form of multiple synapses from the same cortical neuron [513]. As well as being broadly compatible with channel architecture, this means it takes only a small amount of cortical excitation to elicit a strong inhibition of downstream MSNs. This excitation however must be powerful enough to overcome FSIs' innate hyperpolarised membrane potential, in much the same way that MSNs must be pushed into an "up state" [509, 510, 514, 515]. One distinctive feature of FSI-MSN connectivity is that no reciprocal connections from MSNs to FSIs have been observed. Therefore FSI inhibition is described as being "feed forward" [510, 515]. As FSIs receive similar cortical inputs to their neighbouring MSNs, and are capable of acting together via gap junctions, then this arrangement seems well suited to implementing a form of WTA. Other attempts at biologically plausible WTA or max functions have relied on excitatory nodes

inhibiting the dendritic inputs to their neighbours [516], which in some cases was implemented explicitly using inhibitory interneurons [517]. The BG's unique architecture is in some ways too generous in providing potential means to implement these kind of functions. Indeed, Gurney's group has recently explored the role of GABAergic microcircuits, looking within the striatum in isolation [518]. This detailed the formation of cell assemblies with profound dopaminergic modulation. In the model modification described here, FS interneurons will be implemented in a manner more simplistic but generally compatible with that of Humphries et al. [518]. However, the striatum will operate in the full context of the BG.

In order to better describe striatal activity, a new subpopulation of fast spiking inhibitory interneurons will be implemented within the striatum. These will fit into the channel-wise architecture of the existing MSNs, with one FSI unit being associated with an MSN unit in both the D1 receptor-dominated selection and D2 receptor-dominated control pathways. Given that the population ratio of MSNs:FSIs is approximately 18:1 [510, 509], this may not seem immediately appropriate. However, it can be safely assumed that a single channel will be encoded by multiple MSNs, and that the 18:1 proportion will remain sufficient for FSIs to exert their effects, given the profound suppressive action of a single FSI [513].

As FSIs are strongly interconnected via dendrodendritic gap junctions [510], the effect they exert over MSNs will be summed and diffusely projected, in much the same way as STN projects to the two pallidal segments. However, the split between the D1 receptor-dominated selection pathway and the D2 receptor-dominated control pathway will not be equal: it has been found that FS interneurons preferentially target D1-dominated MSNs (interneurons targetting D1 MSNs: 53%; targetting D2 MSNs: 36% [519]). Therefore for the implementation of this modified model, the strength of the D1 projection shall be twice that of the D2 side.

5.1.2 Selectivity processing

Currently in the model, Selectivity demands are varied according to the strength of pallidosubthalamic connectivity. The precise means by which this might be altered is not elaborated on in this work. However, its effects must be transmitted to the striatum if the model is to provide a better fit for the original fMRI data.

One of the strengths of the Gurney model is the way it emphasises the key regulatory role of GPe - a nucleus frequently relegated to "indirect pathway relay" among other understandings of the BG. In the interests of parsimony, the connectivity of GPe was not fully elaborated on. However, it is well known that the STN and GPi are not its only targets of projection [520]. One

pathway that has been argued to be of equal neuroanatomical significance to the STN and GPi projections is that going from GPe to striatum [521, 522]. This is part of a series of projection collaterals that target all three nuclei, therefore remaining compatible with channel architecture [523]. Interestingly, these GPe projection collaterals greatly favour interneurons versus MSNs within striatum [523]. This provides a biologically-grounded means to transmit the dynamic effects that w_g exerts on GPe-STN activity to the striatum, via the very mechanism that could act to impose WTA-type function within this nucleus. This is in part aided by the observation that although the pallidostriatal projection does have a topographical arrangement, it is more diffuse than its striatopallidal counterpart [524]. This is compatible with the modified modelled implementation, in that FS interneuron activity will be summed across all channels before being diffusely transmitted across all MSNs.

5.1.3 Dopaminergic modulation of FS interneurons

Like their MSN counterparts, FS interneurons also express dopamine receptors. D1 receptors facilitate depolarisation, whereas D2 acts to decrease their susceptibility to inhibitory inputs [525] - the net effect of dopamine is therefore excitatory for FS interneurons, leading to a greater suppressions of MSNs. However, given that there are very few inhibitory connections going from MSNs to FS interneurons, it seems the FS interneuron D2 receptors are specifically inhibiting the pallidostriatal projection that targets them. In this modified model, D2 receptors will modulate the input received from GPe units, reducing their efficacy overall. For the sake of parsimony, FS interneuron D1 receptors will not be added to the model - especially as the emphasis here is on the neuropathology of schizophrenia, for which the evidence is very much on the side of D2 receptor abnormalities.

5.2 Dynamic regulation of dopamine

Unlike the original implementation of the model, here background cortical activity is explicitly represented. The motivation for this is to allow dopamine release to be a function of the overall level of cortical activity, with a view to testing hypotheses regarding NMDA hypofunction and subcortical dopamine dysregulation.

Frontal cortex innervates both the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) components of dopaminergic midbrain in non-human primates [526]. This pathway has the characteristics of being more regulatory, in that it is sparse, but terminating onto

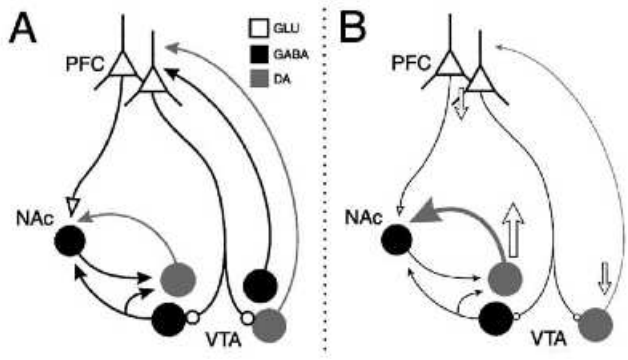


Figure 5.1: The cortico-dopaminergic midbrain projection. Dopaminergic neurons projecting to cortex are directly excited by prefrontal cortex neurons, whereas those projecting to striatum receive their innervation indirectly via an inhibitory interneuron. From Sesack et al. 2002 [15].

neurons richly populated with glutamatergic receptors. However, the nature of these terminal neurons depends on the final destination of the dopamine to be released [15]: dopamine neurons projecting to cortex are directly excited by cortical neurons, whereas dopaminergic neurons targetting subcortical nuclei such as the striatum are inhibited via an intervening GABAergic neuron (figure 5.1). This cortico-midbrain projection will be implemented in the modified version of the model, having a weight significantly lower than than of any other projection originating from cortex, and being inhibitory in nature to accommodate an implicit interneuron. The dopaminergic midbrain itself (denoted the SNc), has a representation of each channel, though its output onto striatal MSNs and FS interneurons is summed across all channels. This summed value will be multiplied by the relevant λ_g (selection pathway D1R efficacy), λ_e (control pathway D2R efficacy) and λ_f (FSI D2R efficacy) values to produce the final dopamine modulation.

The midbrain is subject to a number of afferents. However, one of key relevance in this context is that originating from the GPi/SNr [527, 528]. As well as being one of the BG's main output nuclei, this also acts to inhibit its dopaminergic neighbour, providing a further means to regulate dopamine release in response to demands placed upon the system. In this modified model, a channel-wise inhibitory pathway operates between GPi and SNc. It has been noted that this projection is subject to disinhibition via the GPe [528] - a relationship innately accommodated by the Gurney model's architecture.

5.3 Model implementation

Figure 5.2 displays the modified BG model.

5.3.1 Nucleus definition and initial activation

In a system with n channels in total, the activity of channel i is described. Following the conventions of Gurney et al. 2001 [28], each nucleus is described in turn.

The selection and control striatal MSNs Originally, striatal activity was described by the formulae

$$x_i^{g-} = m^- \left[w_s(1+\lambda_g)c_i - \epsilon\right] H_i^{\uparrow}(-\lambda_g)$$

for MSNs of the D1-modulated selection pathway (g), and

$$x_i^{e-} = m^- \left[w_s(1-\lambda_e)c_i - \epsilon\right] H_i^{\uparrow}(\lambda_e)$$

for the D2-modulated control pathway (e), where w_s is the corticostriatal weight, λ is the tonic dopamine level, and c_i the salience of channel i. m^- is the gradient, and ϵ the threshold of the linear section of the piecewise activation function

$$y = \begin{cases} 0: & a < \epsilon \\ m^-(a-\epsilon): & \epsilon \leq a \leq 1/m^- + \epsilon \\ 1: & a > 1/m^- + \epsilon \end{cases}$$

and $H_i^{\uparrow}(\lambda_e)$ is the Heaviside step function

$$H\left[\frac{c_i - \epsilon}{w_s(1-\lambda_e)}\right]$$

In this modified version, these MSNs receive feedforward inhibition from a functionally-unified set of fast spiking interneurons. In addition, tonic dopamine levels are no longer set numerically in the manner of an additional weight, but are instead explicitly modelled as a function of substantia nigra pars compacta activation, X^d. The striatal MSN activation functions therefore become

$$x_i^g = m^- \left[w_s(1+\lambda_g X^d)c_i - w_{f1}X^f - \epsilon\right] H_i^{\uparrow}(-\lambda_g)$$

and

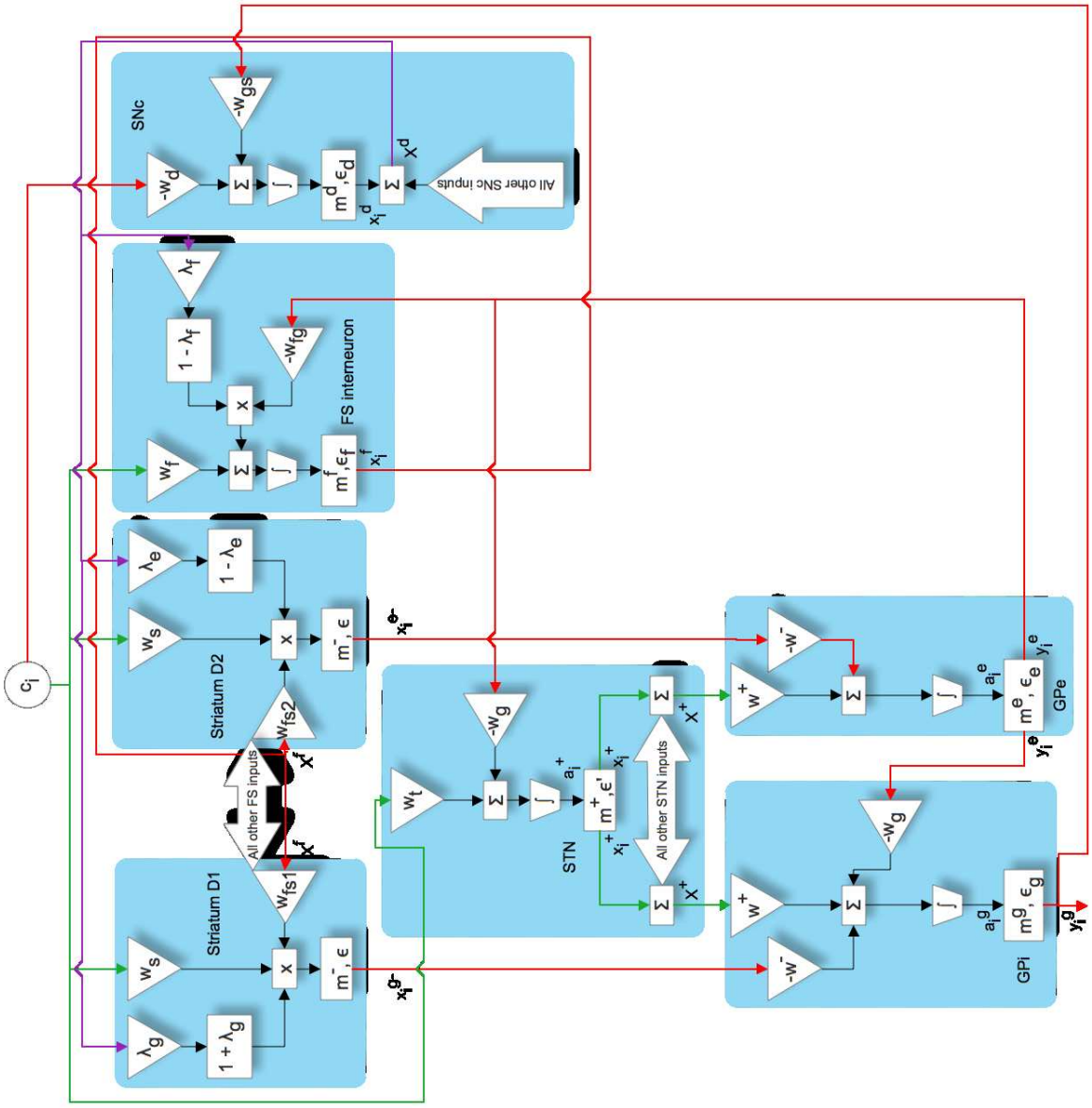


Figure 5.2: Schematic view of the newly modified model. For reference, the original model is illustrated in figure 2.16. It has the following significant structural changes: the addition of striatal fast spiking interneurons, which inhibit striatal MSNs and are themselves inhibited by GPe; an explicit source of dopamine, namely the SNc, which is tonically active but inhibited by both cortex and GPi. Green indicated excitatory projections; red inhibitory; and purple dopaminergic (modulatory).

$$x_i^e = m^- \left[w_s(1 - \lambda_e X^d)c_i - w_{f2} X^f - \epsilon \right] H_i^{\uparrow}(\lambda_e)$$

where w_{f1} and w_{f2} are the FSI-to-MSN inhibitory weights for the selection and control pathways respectively.

Fast spiking interneurons As FSIs are extensively interconnected via gap junctions, their input into each MSN is summed over all channels, i.e.

$$X^f = \sum_i x_i^f$$

where

$$x_i^f = m^f a_i^f H_i^{\uparrow}(a_i^f)$$

and

$$a_i^f = w_f c_i - w_{fg}(1 - \lambda_f X^d) y_i^e - \epsilon_f$$

The formula for a_i^f shows that FS interneurons are excited by a channel-specific cortical input, weighted by w_f, and an inhibitory pallidostratial input originating from GPe (y_i^e). The efficacy of this inhibitory influence is itself mediated via D2 receptors: as these are increasingly occupied, the effect of GPe input is diminished. ϵ_f and m^f are the activation threshold and gradient respectively.

Substantia nigra pars compacta (dopamine levels) In the original model, dopamine efficacy was reflected by the parameters λ_g and λ_e for selection pathway D1 receptors, and control pathway D2 receptors respectively. In this modified version, λ_g and λ_e become weights for an explicit dopamine nucleus activation level. Like the FSIs and STN (see next section), activation is considered on a channel-wise basis (a_i^d), though the output is summed over all channels:

$$X^d = \sum_i x_i^d \tag{5.1}$$

where

$$x_i^d = m^d a_i^d H_i^{\uparrow}(a_i^d)$$

and

$$a_i^d = \epsilon_d - w_d c_i - w_{gs} y_i^g \tag{5.2}$$

It has been shown that for those midbrain cells that target the striatum (as opposed to cortex), their afferent cortical projections target the inhibitory interneurons, rather than the dopaminergic neurons themselves [15, 529]. Therefore the net effect is one of a cortical suppression of subcortical dopamine release. In the modified model, this connection $w_d < 0$. Also, in order to maintain a stable tonic level of dopamine under quiescence, a negative activation threshold ϵ_d will be used. SNc activation is additionally regulated by a projection from GPi (y_i^g), weighted by w_{gs}.

Subthalamic nucleus The original STN implementation had a channel-wise activation of

$$\tilde{a}_i^+ = w_t c_i - w_g y_i^e + \epsilon'$$

which is summed over all channels to provide the output

$$X^+ = m^+ \sum_{i=1}^{n} (w_t c_i - w_g y_i^e) H_i^{\uparrow}(\tilde{a}_i^+)$$

Where w_t is the corticosubthalamic weight, and w_g the pallidosubthalamic weight, feeding in GPe inhibition. The STN remains the same for the modified version, except that the explicit activation threshold ϵ' is no longer necessary, as tonic activation is provided by the addition of background cortical activity.

External globus pallidus Originally, GPe activity was described as

$$y_i^e = m^e \tilde{a}_i^e H(\tilde{a}_i^e)$$

where

$$\tilde{a}_i^e = w^- \left\{ \delta X^+ - x_i^{e-} H_i^{\uparrow}(\lambda_e) \right\} + \epsilon_e$$

w^- is the striatopallidal weight, and δ is derived from $w^+ = \delta w^-$, where w^+ is the subthalamopallidal weight, and x_i^e is the striatal control pathway activation described earlier. The GPe remains fundamentally unchanged. However, as x_i^e now includes both FSI input and dynamic dopamine, its activation can be expressed as

$$\tilde{a}_i^e = w^- \left\{ \delta X^+ - \left[w_s (1 - \lambda_e X^d) c_i - w_{f2} X^f - \epsilon \right] \right\}$$

Note that the explicit activation threshold ϵ_e is no longer needed, as the new addition of background cortical activity maintains tonic pallidal activity via STN.

Internal globus pallidus Originally, GPi activation was described using

$$y_i^g = m^g \tilde{a}_i^g H(\tilde{a}_i^g)$$

where

$$\tilde{a}_i^g = w^- \left\{ \delta X^+ - x_i^{g-} H_i^{\uparrow}(-\lambda_g) \right\} - w_e y_i^e$$

w_e is the weight for GPe's inhibitory input into GPi. As for GPe, the changes to striatal activity means that this changes to

$$\tilde{a}_i^g = w^- \left\{ \delta X^+ - \left[w_s(1 + \lambda_g X^d) c_i - w_{f1} X^f - \epsilon \right] \right\} - w_e y_i^e \tag{5.3}$$

5.3.2 Model analysis

Before a simulation of the model could be implemented, it was first necessary to determine the constraints underlying stable operation of its three key new components: an explicit dopaminergic midbrain module (SNc), FSI-mediated inhibition of striatal MSNs, and a pallidostriatal pathway connecting GPe to the FSIs.

As this model operates, we are interested in two key states: when all channels are "quiescent", that is, at a background level of salience c_{ton}; and when two channels i and j are in direct Competition, with $c_i > c_j$. The interaction that this has with Selectivity must also be established.

5.3.2.1 Nucleus activities during quiescence

Striatal MSNs and FSIs As in the original model, the relatively high activation threshold ϵ keeps striatal units silent until significant activity was present. The new introduction of explicit background cortical activity, c_{ton}, therefore requires that $c_{ton} < \epsilon$ and $c_{ton} < \epsilon_f$ for striatal output to remain zero under quiescence.

External globus pallidus This nucleus is the regulatory core of the model, with the modified version expanding its sphere of influence beyond STN and GPi, to include striatal FS interneurons. During quiescence, it is important that the GPe (and GPi) remain tonically active, in order for this system to be considered a disinhibitory selection mechanism. Following the convention of Gurney et 2001 [28], each GPe channel falls into one of three classes:

$$\begin{aligned} i \in \xi_0 \quad &: \quad y_i^e = w^- \delta X^+ \\ i \in \xi \quad &: \quad y_i^e = w^- \left\{ \delta X^+ - \left[w_s(1 - \lambda_e X^d) c_i - w_{f2} X^f - \epsilon \right] \right\} \\ i \in \xi^* \quad &: \quad y_i^e = 0 \end{aligned} \tag{5.4}$$

where ξ_0 is the set of channels for which $c_i = c_{ton}$, ξ are the channels for which $c_i > c_{ton}$, but $y_i^e > 0$, and ξ^* is the set of channels for which $c_i > c_{ton}$ and GPe has been entirely suppressed ($y_i^e = 0$). During the quiescent state ξ_0, we can say that $y_i^e > 0$ so long as $X^+ > 0$.

Subthalamic nucleus The activation of a single STN channel is described by

$$\tilde{a}_i^+ = w_t c_i - w_g y_i^e + \epsilon'$$

Under quiescent conditions, $c_i = c_{ton}, \forall i$ and $i \in \xi_0, \forall i$. Tonic STN firing X_{ton}^+ is therefore defined as:

$$\begin{aligned} X_{ton}^+ &= n(w_t c_{ton} - w_g w^- \delta X_{ton}^+) \\ X_{ton}^+ &= \frac{n w_t c_{ton}}{1 + n w_g w^- \delta} \end{aligned} \tag{5.5}$$

Therefore STN (and consequently GPe) will always produce a positive tonic output so long as $c_{ton} > 0$.

Internal globus pallidus An individual GPi channel's activation is defined as

$$a_i^g = w^- \delta X^+ - w_e y_i^e$$

which when we enter X_{ton}^+ 5.5 and ξ_0 for y_i^e 5.4 becomes

$$a_i^g = \frac{n w^- \delta w_t c_{ton}(1 - w_e)}{1 + n w_g w^- \delta}$$

Therefore, for $a_i^g > 0$ when $c_i = c_{ton} \forall i$, $w_e < 1$. This was also true of the original implementation.

Substantia nigra pars compacta As dopamine has such a strong modulatory role of the BG model, it is first necessary to establish the conditions for useful SNc activity with respect to background cortical activity. The activity of a single SNc channel is defined as

$$x_i^d = \epsilon_d - w_d c_i - w_{gs} y_i^g$$

Under quiescent conditions, where $c_i = c_{ton}, \forall i$ and $i \in \xi_0, \forall i$, tonic X_{ton}^d is defined as

$$\begin{aligned} X^d_{ton} &= n\left(\epsilon_d - w_d c_{ton} - w_{gs}\left[\frac{nw^-\delta w_t c_{ton}(1-w_e)}{1+nw_g w^-\delta}\right]\right) \\ &= \frac{n\epsilon_d + n\epsilon_d n w_g w^-\delta - n w_d c_{ton} - n w_d c_{ton} n w_g w^-\delta - w_{gs} n w^-\delta w_t c_{ton}(1-w_e)}{1+nw_g w^-\delta} \\ &= \frac{n\left(nw_g w^-\delta\left[\epsilon_d - w_d c_{ton}\right] + c_{ton}\left[w_{gs} w^-\delta w_t\left(w_e - 1\right) - w_d\right] + \epsilon_d\right)}{1+nw_g w^-\delta} \end{aligned}$$

If $n \gg 1$, then $n^2 \gg n$, therefore X^d_{ton} can be approximated as

$$X^d_{ton} \approx n^2 w_g w^-\delta\left(\epsilon_d - w_d c_{ton}\right)$$

Therefore, for $X^d_{ton} > 0$,

$$\epsilon_d > w_d c_{ton} \tag{5.6}$$

Neuroanatomical evidence suggests that w_d is sparse [526]. Therefore this is justified in adopting a small weight. This would ensure that this constraint is satisfied over a broad range of conditions. Importantly, this formula reveals that as background cortical activity decreases, tonic dopamine levels will increase.

5.3.2.2 Activity during Competition

Competition is inversely related to the magnitude of $\Delta c = c_i - c_j$.

Internal globus pallidus First, the model was checked to ensure it still performed order-preserving mapping, that is, that if $c_i > c_j$ and $\Delta_{ji} y^g = y^g_j - y^g_i$, then $\Delta_{ji} y^g > 0$. This key function must remain if the BG are to be considered an action selection mechanism. Working with the GPi activation equation 5.3:

$$\begin{aligned} \Delta_{ji} a^g &= w^-\left[\delta X^+ - \left(w_s(1+\lambda_g X^d)c_i - w_{f1} X^f - \epsilon\right)\right] - w_e\left[w^-(\delta X^+ - \left(w_s(1-\lambda_e X^d)c_i - w_{f2} X^f - \epsilon\right)\right] \\ &\quad - w^-\left[\delta X^+ - \left(w_s(1+\lambda_g X^d)c_j - w_{f1} X^f - \epsilon\right)\right] + w_e\left[w^-(\delta X^+ - \left(w_s(1-\lambda_e X^d)c_j - w_{f2} X^f - \epsilon\right)\right] \\ &= w_s \Delta c\left(w^-(1+\lambda_g X^d) - w_e(1-\lambda_g X^d)\right) \end{aligned} \tag{5.7}$$

As in the original implementation, the model will remain an disinhibitory action selector so long as $w^- > w_e$.

Substantia nigra pars compacta (dopamine levels)

A key change from the original model implementation is that dopamine levels are now partly dependent on the demands being placed upon the system. This must be characterised in order to understand how it will affect the responses of downstream nuclei to Competition.

In order to determine how dopamine will fluctuate when the system is no longer quiescent, it is necessary to determine the change in the summed output of the SNc with respect to the change in salience, $d\Delta X^d/d\Delta c$, where $\Delta X^d = X^d_{c_i>0} - X^d_{c_j>0}$, $c_i > c_j$, and where ΔX^d is defined by equations 5.1 and 5.2.

$$\begin{aligned}
\Delta X^d &= (\epsilon_d - w_d c_i - w_{gs} y_i^g) - \left(\epsilon_d - w_d c_j - w_{gs} y_j^g\right) \\
&= \left(\epsilon_d - w_d c_i - w_{gs}\left[w^-\left(\delta X^+ - \left(w_s(1+\lambda_g X^d)c_i - w_{f1}X^f - \epsilon\right)\right) - w_e y_i^e\right]\right) \\
&\quad - \left(\epsilon_d - w_d c_j - w_{gs}\left[w^-\left(\delta X^+ - \left(w_s(1+\lambda_g X^d)c_j - w_{f1}X^f - \epsilon\right)\right) - w_e y_j^e\right]\right) \\
&= w_{gs}w^- w_s(1+\lambda_g X^d)c_i - w_d c_i + w_{gs}w_e y_i^e + w_d c_j - w_{gs}w^- w_s(1+\lambda_g X^d)c_j - w_{gs}w_e y_j^e \\
&= w_{gs}w^- w_s \lambda_g X^d \Delta c + \Delta c\left(w_{gs}w^- w_s - w_d\right) + w_{gs}w_e\left(y_i^e - y_j^e\right)
\end{aligned}$$

Determining Competition in GPe:

$$\begin{aligned}
y_i^e - y_j^e &= w^-\left[\delta X^+ - \left(w_s(1-\lambda_e X^d)c_i - w_{f2}X^f - \epsilon\right)\right] - w^-\left[\delta X^+ - \left(w_s(1-\lambda_e X^d)c_j - w_{f2}X^f - \epsilon\right)\right] \\
&= w^-\Delta c\left(w_s\lambda_e X^d - w_s\right)
\end{aligned}$$

Therefore replacing this in ΔX^d:

$$\begin{aligned}
\Delta X^d &= w_{gs}w^- w_s\lambda_g X^d(c_i - c_j) + (c_i - c_j)\left(w_{gs}w^- w_s - w_d\right) + w_{gs}w_e w^-(c_i - c_j)\left(w_s\lambda_e X^d - w_s\right) \\
\Delta X^d &= \frac{\Delta c\left(w_{gs}w^- w_s\left[1 - w_e\right] - w_d\right)}{1 - w_{gs}w^- w_s\Delta c\left(\lambda_g - w_e\lambda_e\right)} \qquad (5.8)
\end{aligned}$$

If we differentiate this with respect to Δc:

$$\begin{aligned}
\frac{d\Delta X^d}{d\Delta c} &= \frac{\left[\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)\left(w_{gs}w^{-}w_s\left[1-w_e\right]-w_d\right)\right]}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&\quad + \frac{\left[w_{gs}w^{-}w_s\left(\lambda_g-w_e\lambda_e\right)\Delta c\left(w_{gs}w^{-}w_s\left[1-w_e\right]-w_d\right)\right]}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&= \frac{\left[\left(1-w_{gs}w^{-}w_s\Delta c\lambda_g+w_{gs}w^{-}w_s\Delta cw_e\lambda_e\right)\left(w_{gs}w^{-}w_s-w_{gs}w^{-}w_sw_e-w_d\right)\right]}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&\quad + \frac{\left[\left(w_{gs}w^{-}w_s\lambda_g\Delta c-w_{gs}w^{-}w_sw_e\lambda_e\Delta c\right)\left(w_{gs}w^{-}w_s-w_{gs}w^{-}w_sw_e-w_d\right)\right]}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&= \frac{w_{gs}w^{-}w_s\left[1-w_e+\Delta c\left(w_{gs}w^{-}w_s\left[\lambda_g-\lambda_g\right]\right)\right]-w_d}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&= \frac{w_{gs}w^{-}w_s-w_{gs}w^{-}w_sw_e+w_{gs}w^{-}w_s\Delta cw_{gs}w^{-}w_s\lambda_g-w_{gs}w^{-}w_s\Delta cw_{gs}w^{-}w_s\lambda_g-w_d}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2} \\
&= \frac{w_{gs}w^{-}w_s\left(1+w_{gs}\Delta cw^{-}w_s\lambda_g\right)}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2}-\frac{w_{gs}w^{-}w_s\left(w_e+w_{gs}\Delta cw^{-}w_s\lambda_g\right)+w_d}{\left(1-w_{gs}w^{-}w_s\Delta c\left(\lambda_g-w_e\lambda_e\right)\right)^2}
\end{aligned}$$

Therefore dopamine levels will rise with increasing Competition i.e. as $c_i - c_j \to 0$, $d\Delta X^d/d\Delta c < 0$:

$$\begin{aligned}
w_{gs}w^{-}w_s\left(w_e+w_{gs}\Delta cw^{-}w_s\lambda_g\right)+w_d &> w_{gs}w^{-}w_s\left(1+w_{gs}\Delta cw^{-}w_s\lambda_g\right) \\
w_{gs}w^{-}w_sw_e+w_d &> w_{gs}w^{-}w_s \\
w_e+w_d &> 1
\end{aligned}$$

I.e., dopamine levels will increase with respect to Competition so long as the weights of the GPe-GPi and cortex-SNc connections sum to at least 1. The greater the excess, the broader the dynamic range of dopamine's response.

Striatum

FS feedforward inhibition The first study suggested that total striatal activation increased as $c_i - c_j \to 0$. One means by which this could occur is if the feedforward inhibition MSNs were subjected to was alleviated as Competition became more intense. Therefore the first step was to determine under what conditions X^f would decrease as $c_i - c_j \to 0$. As all quiescent channels will produce no striatal activation, the effects of changes in $c_i - c_j = \Delta c$ can be established by examining $\Delta X^f = x_i^f - x_j^f$.

$$\begin{aligned}
\Delta X^f &= w_fc_i - w_{fg}(1-\lambda_fX^d)y_i^e - \epsilon_f - \left(w_fc_j - w_{fg}(1-\lambda_fX^d)y_j^e - \epsilon_f\right) \\
&= w_f\Delta c - w_{fg}(1-\lambda_fX^d)(y_i^e - y_j^e) \qquad (5.9)
\end{aligned}$$

The difference between the two GPe channels $\Delta y^e = y_i^e - y_j^e$ is:

$$\begin{aligned}\Delta y^e &= w^- \left[\delta X^+ - \left(w_s(1-\lambda_e X^d)c_i - w_{f2}X^f - \epsilon\right)\right] - w^- \left[\delta X^+ - \left(w_s(1-\lambda_e X^d)c_j - w_{f2}X^f - \epsilon\right)\right] \\ &= w^- w_s \Delta c \left(\lambda_e X^d - 1\right) \end{aligned} \tag{5.10}$$

The change in X^d with respect to Competition is:

$$\begin{aligned}\Delta X^d &= \epsilon_d - w_d c_i - w_{gs} y_i^g - \left(\epsilon_d - w_d c_j - w_{gs} y_j^g\right) \\ &= w_{gs}\Delta y^g - w_d \Delta c\end{aligned}$$

Note that $\Delta y^g = y_j^g - y_i^g$. From equation 5.7 we know:

$$\Delta y^g = w_s \Delta c \left(w^-(1+\lambda_g X^d) - w_e(1-\lambda_g X^d)\right)$$

Therefore ΔX^d is:

$$\begin{aligned}\Delta X^d &= w_{gs} w_s \Delta c \left(w^-(1+\lambda_g X^d) - w_e(1-\lambda_g X^d)\right) - w_d \Delta c \\ &= \frac{\Delta c \left(w_{gs} w_s \left[w^- - w_e\right] - w_d\right)}{1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]}\end{aligned} \tag{5.11}$$

When this is placed back in Δy^e in 5.10:

$$\begin{aligned}\Delta y^e &= w^- w_s \Delta c \left(\lambda_e X^d - 1\right) \\ &= \frac{w^- w_s \Delta c \Delta c \lambda_e \left(w_{gs} w_s \left[w^- - w_e\right] - w_d\right)}{1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]} - w^- w_s \Delta c \\ &= \frac{\left(w^{-2} w_s^2 w_{gs} \Delta c^2 \lambda_e - w^- w_s^2 w_{gs} w_e \Delta c^2 \lambda_e - w^- w_s w_d \Delta c^2 \lambda_e\right)}{1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]} - w^- w_s \Delta c \\ &= \frac{w^- w_s \Delta c^2 \left(w_s w_{gs} \left[\lambda_e \left(w^- - w_e\right) - \lambda_g \left(w^{-2} w_s \Delta c + w_e\right)\right] - w_d \lambda_e\right)}{1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]}\end{aligned} \tag{5.12}$$

Placing 5.11 and 5.12 back in 5.9:

$$
\begin{aligned}
\Delta X^f &= w_f\Delta c - w_{fg}(1-\lambda_f X^d)(y_i^e - y_j^e) \\
&= w_f\Delta c \\
&\quad -w_{fg}\left(1-\lambda_f\frac{\Delta c\left(w_{gs}w_s\left[w^- - w_e\right]-w_d\right)}{1-w_{gs}w_s\Delta c\left[w^-\lambda_g - w_e\lambda_g\right]}\right)... \\
&\quad ...\frac{w^-w_s\Delta c^2\left(w_sw_{gs}\left[\lambda_e\left(w^- - w_e\right)-\lambda_g\left(w^{-2}w_s\Delta c + w_e\right)\right]-w_d\lambda_e\right)}{1-w_{gs}w_s\Delta c\left[w^-\lambda_g - w_e\lambda_g\right]} \\
&= w_f\Delta c - w_{fg}... \\
&\quad ...\left(\frac{w^-w_s\Delta c^2\left(w_sw_{gs}\Delta c\left[\lambda_g+\lambda_f\right]\left(w_e-w^-\right)+w_d+1\right)\left(w_sw_{gs}\left[\lambda_e\left(w^- - w_e\right)-\lambda_g\left(w_sw^{-2}\Delta c + w_e\right.\right.\right.\right.}{w_{gs}w_s\Delta c\lambda_g\left(w_{gs}w_s\Delta c\lambda_g\left(-w^- - w_e\right)^2 - 2w^- + 2w_e\right)+1}\right. \\
&\quad \left.+\frac{w_{fg}w_d\lambda_e}{w_{gs}w_s\Delta c\lambda_g\left(w_{gs}w_s\Delta c\lambda_g\left(-w^- - w_e\right)^2 - 2w^- + 2w_e\right)+1}\right)
\end{aligned}
$$

If we eliminate the small terms w_{gs} and λ_x that have been squared or multiplied:

$$\Delta X^f \approx w_f\Delta c - w_{fg}\left(\frac{w^-w_s\Delta c^2\left(w_d+1\right)}{w_{gs}w_s\Delta c\lambda_g\left(-2w^- + 2w_e\right)+1}\right)$$

Differentiate with respect to Δc, and again removing squared small terms:

$$
\begin{aligned}
\frac{d\Delta X^f}{d\Delta c} &\approx w_f - w_{fg}w^-w_sw_d\Delta c - w_{fg}w^-w_s\Delta c \\
&\approx w_f - w^-w_sw_{fg}\Delta c\left[w_d+1\right]
\end{aligned}
$$

For $\Delta X^f \to 0$ as $\Delta c \to 0$, this needs to be positive. If w^- and w_s are 1:

$$w_f > w_{fg}\Delta c\left[w_d+1\right]$$

The $[w_d+1]$ term on the right-hand side means w_f will need to take a maximal value, and both $w_{fg} < 1$ and $w_d < 1$, otherwise the model will not behave appropriately at low Competition levels, i.e. when Δc is large.

Selection pathway MSNs The weights connecting the FS interneurons to MSNs must be carefully set so as not to overwhelm the MSNs altogether and stifle any activation. For a given change in a single input c_i, we would expect striatal activation x_i^g to positively increase.

$$x_i^g = w_s(1+\lambda_g X^d)c_i - w_{f1}X^f - \epsilon \tag{5.13}$$

With a single active channel, X^f can be summarised as:

$$X_1^f \sim w_f c_i - w_{fs}(1-\lambda_f X^d)y_i^e - \epsilon_f \tag{5.14}$$

GPe is:

$$y_i^e = w^- \left[\delta X^+ - \left(w_s(1-\lambda_e X^d)c_i - w_{f2}X^f - \epsilon\right)\right] \tag{5.15}$$

Where STN with a single active channel can be summarised as:

$$X_1^+ \sim w_t c_i - w_g y_i^e \tag{5.16}$$

Replacing 5.16 and 5.14 within 5.15 gives:

$$\begin{aligned} y_i^e &= w^- \left[\delta\left(w_t c_i - w_g y_i^e\right) - \left(w_s(1-\lambda_e X^d)c_i - w_{f2}\left(w_f c_i - w_{fs}(1-\lambda_f X^d)y_i^e - \epsilon_f\right) - \epsilon\right)\right] \\ &= w^- \left[y_i^e\left(w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right) + \delta w_t c_i - w_s c_i + \lambda_e w_s c_i X^d + w_{f2}w_f c_i - \epsilon_f + \epsilon\right] \\ &= \frac{w^-\left[\delta w_t c_i - w_s c_i + \lambda_e w_s c_i X^d + w_{f2}w_f c_i - \epsilon_f + \epsilon\right]}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]} \end{aligned} \tag{5.17}$$

Replacing 5.17 in 5.14:

$$X_1^f \sim w_f c_i - w_{fs}(1-\lambda_f X^d)\frac{w^-\left[\delta w_t c_i - w_s c_i + \lambda_e w_s c_i X^d + w_{f2}w_f c_i - \epsilon_f + \epsilon\right]}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]} - \epsilon_f \tag{5.18}$$

When 5.18 is reinserted into x_i^g from 5.13:

$$x_i^g = w_s(1+\lambda_g X^d)c_i - w_{f1}w_f c_i - w_{fs}(1-\lambda_f X^d)\frac{w^-\left[\delta w_t c_i - w_s c_i + \lambda_e w_s c_i X^d + w_{f2}w_f c_i - \epsilon_f + \epsilon\right]}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]} - \epsilon_f - \epsilon \tag{5.19}$$

If dopamine is held constant, and differentiate with respect to c_i:

$$\begin{aligned} \frac{dx_i^g}{dc_i} &= w_s(1+\lambda_g X^d) - w_{f1}w_f - \frac{w_{fs}(1-\lambda_f X^d)w^-\left[\delta w_t - w_s + \lambda_e w_s X^d + w_{f2}w_f\right]}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]} \\ &= \left[w_s(1+\lambda_g X^d) + \frac{w_{fs}w^-\left(w_s X^d\left[\lambda_e + \lambda_f\right] + \delta w_t + w_{f2}w_f\right)}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]}\right] \\ &\quad - \left[w_{f1}w_f + \frac{w_{fs}w^-\left(X^d\lambda_f\left[\delta w_t + \lambda_e w_s + w_{f2}w_f\right] + w_s\right)}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]}\right] \end{aligned}$$

This shows that the parameters contributing to a positive or negative response are quite evenly distributed, except for w_{f1}, which appears exclusively on the negative side. Assuming that the distributed factors partially cancel out, and that w_s and w_f are equal, then $dx_i^g/dc_i > 0$ so long as $w_{f1} < w_s$.

External globus pallidus The pallidal response to both Selectivity (positive) and Competition (negative) was sufficiently captured by the original model. Here we seek to maintain that account, and determine the parametric constraints permitting $\sum y^e \to 0$ as $\Delta c \to 0$. To assess the response of GPe to Competition, we evaluate $\Delta y^e = y_i^e - y_j^e$, which we know from 5.12 is:

$$\Delta y^e = \frac{w^- w_s \Delta c^2 \left(w_s w_{gs} \left[\lambda_e \left(w^- - w_e\right) - \lambda_g \left(w^{-2} w_s \Delta c + w_e\right)\right] - w_d \lambda_e\right)}{1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]}$$

If we differentiate this with respect to Δc, and assume that the small terms w_{gs} and λ_x are rendered insignificant when raised to a power greater than two:

$$\begin{aligned}
\frac{d\Delta y^e}{d\Delta c} &= \frac{\left[\left(1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]\right) w^- w_s \Delta c \left(w_s w_{gs} \left[\lambda_e \left(w^- - w_e\right) - \lambda_g \left(w^{-2} w_s \Delta c + w_e\right)\right] - w_d \lambda_e\right)\right]}{\left[1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]\right]^2} \\
&\quad - \frac{\left[-w_{gs} w_s \left[w^- \lambda_g - w_e \lambda_g\right] w^- w_s \Delta c^2 \left(w_s w_{gs} \left[\lambda_e \left(w^- - w_e\right) - \lambda_g \left(w^{-2} w_s \Delta c + w_e\right)\right] - w_d \lambda_e\right)\right]}{\left[1 - w_{gs} w_s \Delta c \left[w^- \lambda_g - w_e \lambda_g\right]\right]^2} \\
&\approx w_s^2 w_{gs} w^{-2} \Delta c \lambda_e - w_e w^- w_s \Delta c \lambda_e - w_s^3 w_{gs} w^{-3} \Delta c^2 \lambda_g - w_s^2 w_{gs} w_e w^- \Delta c \lambda_g - w_d w^- w_s \Delta c \lambda_e \\
&\approx w^- w_s \Delta c \left(\left[w_s w_{gs} w^- \lambda_e\right] - \left[w_s w_{gs} \lambda_g \left(w_s w^{-2} \Delta c + w_e\right) + \lambda_e \left(1 + w_e + w_d\right)\right]\right)
\end{aligned}$$

For $\sum y^e \to 0$ as $\Delta c \to 0$, $d\Delta y^e / d\Delta c > 0$. If w_s and w^- are 1, it should be the case that:

$$\left[w_{gs} \lambda_e\right] > \left[w_{gs} \lambda_g \left(\Delta c + w_e\right) + \lambda_e \left(1 + w_e + w_d\right)\right]$$

As it has previously been asserted that $w_e + w_d > 1$, and w_{gs} should be small, it would appear that keeping this relationship positive will be challenging. This will be explored further during the simulations.

5.3.2.3 Activity under Selectivity

Selectivity is manipulating by altering the connection between GPe and STN (w_g). In order to determine how the modified nuclei will vary with increasing Selectivity demands, their activation with respect to w_g needs to be determined.

Dopamine and Selectivity As has been noted during the assessment of tonic quiescent (equation 5.6) and competitive (equation 5.11) SNc activity, dopamine levels are not affected by w_g.

Pallidus and Selectivity A key characteristic of both the original model and fMRI data was that as Selectivity increased, i.e. as $w_g \to 0$, total pallidal activity also increased. It was therefore necessary to determine if the modified model was still capable of behaving accordingly. Considering only one single active channel, GPe is defined as:

$$y_i^e = w^- \left\{\delta X^+ - \left[w_s(1-\lambda_e X^d)c_i - w_{f2}X^f - \epsilon\right]\right\}$$

and similarly STN is defined as:

$$X^+ = w_t c_i - w_g y_i^e$$

FSI activity approximated as:

$$X^f \approx w_f c_i - w_{fg}(1-\lambda_f X^d)y_i^e - \epsilon_f$$

SNc approximated as:

$$X^d \approx \epsilon_d - w_d c_i - w_{gs} y_i^g$$

and GPi as:

$$y_i^g = w^- \left\{\delta X^+ - \left[w_s(1+\lambda_g X^d)c_i - w_{f1}X^f - \epsilon\right]\right\} - w_e y_i^e$$

Therefore X^d becomes:

$$\begin{aligned} X^d &\approx \epsilon_d - w_d c_i \\ &\quad -w_{gs}\left[w^-\left\{\delta\left[w_t c_i - w_g y_i^e\right] - \left[w_s(1+\lambda_g X^d)c_i - w_{f1}\left(w_f c_i - w_{fg}(1-\lambda_f X^d)y_i^e - \epsilon_f\right) - \epsilon\right]\right\} - w_e y_i^e\right] \\ &\approx \frac{c_i\left[w_{gs}w^-\left(w_s - \delta w_t - w_{f1}w_f\right) - w_d\right] + w_{gs}y_i^e\left[w^-\delta w_g + w^- w_{f1}w_{fg} + w_e\right]}{1 - w_{gs}w^-\left[w_s\lambda_g c_i + w_{f1}w_{fg}\lambda_f y_i^e\right]} \\ &\quad + \frac{\epsilon_d + w_{gs}w^- w_{f1}\epsilon_f - w_{gs}w^-\epsilon}{1 - w_{gs}w^-\left[w_s\lambda_g c_i + w_{f1}w_{fg}\lambda_f y_i^e\right]} \end{aligned}$$

and X^f becomes

$$\begin{aligned} X^f &\approx w_f c_i - \epsilon_f - w_{fg}y_i^e \\ &\quad + w_{fg}y_i^e\lambda_f\left[\frac{c_i\left[w_{gs}w^-\left(w_s - \delta w_t - w_{f1}w_f\right) - w_d\right] + w_{gs}y_i^e\left[w^-\delta w_g + w^- w_{f1}w_{fg} + w_e\right]}{1 - w_{gs}w^-\left[w_s\lambda_g c_i + w_{f1}w_{fg}\lambda_f y_i^e\right]}\right] \\ &\quad + w_{fg}y_i^e\lambda_f\left[\frac{\epsilon_d + w_{gs}w^- w_{f1}\epsilon_f - w_{gs}w^-\epsilon}{1 - w_{gs}w^-\left[w_s\lambda_g c_i + w_{f1}w_{fg}\lambda_f y_i^e\right]}\right] \end{aligned}$$

When X^f and X^d are inserted into y_i^e, only the denominator contains w_g terms. Differentiating this with respect to w_g then gives:

$$\frac{dy_i^e}{dw_g} = -w^- y_i^e \delta w_g^{-2}\left(1 - w_{gs}\left[w_{fg}\lambda_f y_i^e\left(2w_{f1} + w^- w_{f2}\right) + w^- w_s c_i\left(\lambda_g + \lambda_e\right)\right]\right)$$

For this to allow pallidal activity to increase with increasing Selectivity (i.e. decreasing w_g), this must remain negative, i.e.:

$$1 > w_{gs}\left[w_{fg}\lambda_f y_i^e\left(2w_{f1} + w^- w_{f2}\right) + w^- w_s c_i\left(\lambda_g + \lambda_e\right)\right] \tag{5.20}$$

Generally, this will be satisfied so long as $w_{gs} < 1$, and as w_{gs} is small for anatomical reasons, this shall be the case. However, as $dy_i^e/dw_g \sim w_g^2$, this may not hold true over all values of w_g - simulations will clarify this further.

Striatum and Selectivity One of the observations made from the data of the initial study with healthy controls was that Selectivity was having an effect on striatal activation. This was not something the original model could account for. In this modified implementation, the BG's key control node - GPe - is now able to modulate striatal FS interneuron firing. We now consider if and how the effects of Selectivity will be transmitted to striatum.

As determined earlier (equation 5.19), the formula for a selection pathway striatal MSN with one active channel is:

$$x_i^g = w_s(1+\lambda_g X^d)c_i - w_{f1}w_f c_i - w_{fs}(1-\lambda_f X^d)\frac{w^-\left[\delta w_t c_i - w_s c_i + \lambda_e w_s c_i X^d + w_{f2}w_f c_i - \epsilon_f + \epsilon\right]}{1 - w^-\left[w_{f2}w_{fs}\lambda_f X^d - \delta w_g - w_{f2}w_{fs}\right]} - \epsilon_f - \epsilon$$

As X^d is not affected by w_g, these terms can be ignored as being constant. Differentiating x_i^g with respect to w_g, and removing the terms w^-, w^t and w_s for clarity (as they both equate to 1):

$$\begin{aligned}
\frac{dx_i^g}{dw_g} &= \frac{-w_{fs}(1-\lambda_f)\left[\delta c_i - c_i + \lambda_e c_i + w_{f2}w_f c_i - \epsilon_f + \epsilon\right]}{\delta w_g^2\left(1-w^-\right)} \\
&= \frac{w_{fs}\left[w_f w_{f2} c_i\left(\lambda_f - 1\right) - \epsilon_f\left(\lambda_f - 1\right) + \epsilon\left(\lambda_f - 1\right) + c_i\left(1 + \lambda_f\lambda_e + \lambda_f\delta - \lambda_e - \lambda_f - \delta\right)\right]}{\delta w_g^2\left(1-w^-\right)} \\
&= \frac{w_{fs}\left(\left[\lambda_f\left(c_i\left[\delta + \lambda_e + w_{f2}w_f\right] + \epsilon\right) + c_i + \epsilon_f + \epsilon\right] - \left[\lambda_f\left(c_i + \epsilon_f\right) + c_i\left(\delta + \lambda_e + w_{f2}w_f\right)\right]\right)}{\delta w_g^2\left(1-w^-\right)}
\end{aligned}$$

Selectivity increases as the magnitude of w_g decreases. If modelled striatal activity is to increase with Selectivity, $dx_i^g/dw_g < 0$. The differential suggests that λ_f acts to diminish the positive aspects of the differential more than the negative, therefore the system seems theoretically capable of replicating the fMRI data in this regard. Also, the w_g^2 term in the denominator suggests a quadratic rather than a straightforward linear relationship, as for the pallidus.

5.4 Model simulations

To better assess the modified model's performance, a series of simulations were performed. The first steps were to establish if the model acted as a stable action selector between two competing channels, and whether it provided a better account for the fMRI data.

In a second series of simulations, the model was manipulated as a test bed for ideas concerning the pathophysiology of schizophrenia. The effects of two key hypothetical pathologies were examined: the level of background cortical activity, representing the general efficacy of glutamatergic transmission; and the efficacy of striatal D2 receptors. As well as considering how these factors altered nucleus activation (providing potential predictions for fMRI patient data), their effects on symptom-relevant behaviour were assessed using two measures: promiscuousness, that is, the ease with which two competing channels can be simultaneously selected; and distractibility - the ability of a selected channel to withstand a growing competitor.

5.4.1 Parameter initialisation

In all cases, the slope of the output relations m^-, m^+, m^g, m^e and m^d were set to 1. The sole exception was the gradient for FSIs m^f: as electrophysiological evidence suggests that these fire as much as four times faster for a given current injection [530], this was set to 3.

The firing thresholds were somewhat altered from the original model: the explicit inclusion of background cortical activity now acted to provide the tonic firing of STN, and hence GPe/GPi. The parameters ϵ', ϵ_e and ϵ_g were therefore set to zero. The striatal ϵ parameter remained unchanged at 0.2. Two new firing thresholds were set: for the striatal FS interneurons, $\epsilon_f = 0.4$, as electrophysiological evidence suggests these require a greater depolarisation than MSNs in order to generate action potentials [530]; and for the dopaminergic SNc units, $\epsilon_d = -0.1$, to provide tonic activation for a nucleus which is subject to multiple sources of inhibition in this model.

Most of the weights defined in the original implementation of this model remained unchanged. Here are listed initialisation values for the new weights, and any changes that were made to the original model parameter values:

- $w_e = w_d = -0.55$: During mathematical analysis, it was found that $|w_e| + |w_d| > 1$ for dopamine levels to increase with increasing Competition ($c_i - c_j \rightarrow 0$). Therefore the old GPe -> GPi value was set appropriately, and the new cortex -> SNc weight set to match. The new w_e value still conformed to the other constraints that $w^- > w_e$ for the model to remain an action selector, as well as $w_e < 1$ for GPi to be tonically active under quiescence. A w_d value of -0.55 is biologically justified, in that this connection is anatomically sparse but electrophysiologically potent [15, 531], but not to the same degree as, say, the cortex -> striatum connection of w_s, which is set to -1.

- $w_f = 1$ (cortex to striatal FSI): this was assumed to match w_s, the general corticostriatal pathway.
- $w_{f1} = -0.4$, $w_{f2} = -0.2$ (FSI to selection- and control-pathway MSNs respectively): during the analysis, it was determined that in order for $dx_i^g/dc_i > 0$, the magnitude of $w_{f1} < w_s$. w_{f2} was set to half the value of w_{f1} on the basis of evidence from tracing studies: FSIs have been found to preferentially target D1-type direct pathway MSNs within striatum, having a ratio of 53% vs 36% for functional synapses [519].
- $w_{fg} = -1$ (GPe to striatal FSIs): partly for simplicity, this was matched to w_g (the connection between GPe and STN), though representing this as such a significant connection is supported by tracer studies [521, 522].
- $w_{gs} = -0.1$ (GPi to dopaminergic SNc): this was set to match the magnitude of ϵ_d, so allow dopamine to fluctuate approximately between 0 and 1 in accordance with the conditions imposed by background salience and Competition.

5.4.2 Verifying the model's action selection function

Initially, a simple simulation was run that held all weights at default values, and varied both Competition and Selectivity, to determine if the model could still perform disinhibitory action selection.

5.4.2.1 Varying Competition

Background cortical activity $c_{ton} = 0.1$, and 2 channels (3 and 4, with 3 always having the higher salience) were manipulated over a range of Competition values 5.3. For the initial verification simulation, these Competitions were set up serially in time over 1000 time steps, each for 50 time steps, with a 50 step break in between.

Striatal activity The striatal selection and control pathways responded appropriately, and at least one channel was able to overcome FSI inhibition at all Competition levels (figure 5.4). As expected, the D1 selection pathway is more active than its D2 control pathway counterpart, due to the facilitative vs diminishing effects of dopamine.

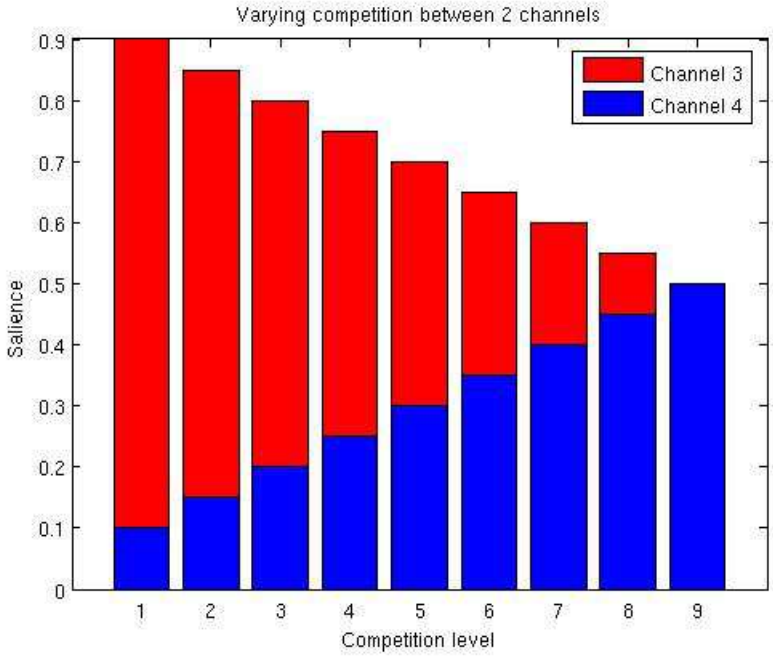


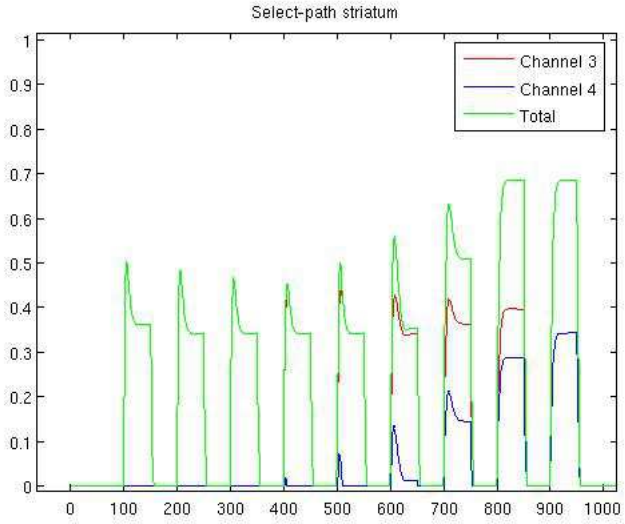
Figure 5.3: How Competition was manipulated over two channels. All Competitions summed to 1.

Internal globus pallidus To be considered an action selector, the channel with the highest striatal activation should force its respective GPi unit to the lowest value. Simulations suggest that this is indeed the case (figure 5.5). The channel with the highest salience was always selected (except during the 50:50 state when both were selected), with its competitor forced to a high value.

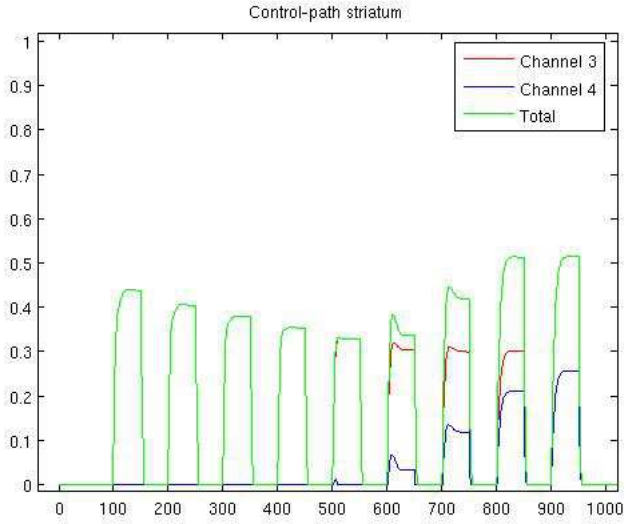
5.4.2.2 Varying Selectivity

One of the key predictions of the original model was that Selectivity and pallidal activation increased together - which was indeed born out by the fMRI data. To examine if this remains the case, a simulation was run where Competition was kept constant with $c_1 = 0.75$, and $c_2 = 0.25$, and w_g varied from -1.0 to -0.5.

An interesting quadratic relationship between these variables has emerged, as anticipated by the mathematical analysis (equation 5.20): pallidal activation does indeed increase with Selectivity so long as $-0.75 < w_g < -0.5$, though this is reversed if $-1.0 < w_g < -0.75$. In



(a)



(b)

Figure 5.4: Striatal responses to Competition for (a) the D1 receptor-dominated selection pathway and (b) the D2 receptor-dominated control pathway, for the modified model. The x axis displays time in model time steps, where each ”on” period corresponds to a Competition level in the order described in figure 5.3.

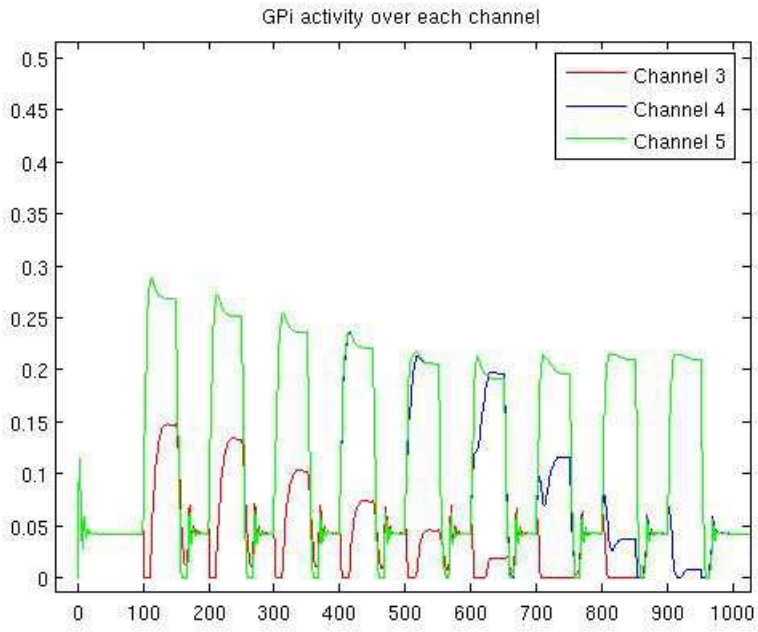


Figure 5.5: Responses of the internal globus pallidus to Competition for the modified model. In each case, the channel with the highest salience (channel 3) is forced to zero, whereas the losing competitor is forced to a high value. Only at the 50:50 Competition level are both channels selected. The x axis displays time in model time steps, where each ”on” period corresponds to a Competition level in the order described in figure 5.3.

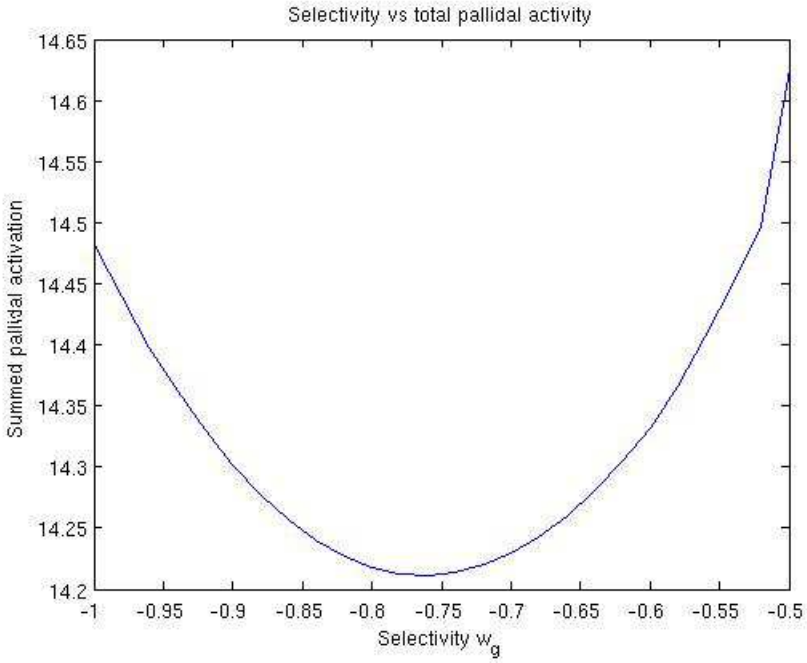


Figure 5.6: The effect of varying Selectivity on total pallidal activation for the modified model. Selectivity was varied between -1.0 and -0.5. Competition was held constant at 0.75:0.25.

order that the model matches the fMRI data, w_g will be kept within the range $-0.75 < w_g < -0.5$.

5.4.3 Comparing the model's responses to the fMRI findings

The next step was to determine if the model could better replicate the Competition x Selectivity interaction for both striatal and pallidal activation.

5.4.3.1 Pallidal activation

It has already been established that pallidal activity increases with Selectivity within a circumscribed range. However, its relationship with Competition was yet to be ascertained. With $w_g = -0.75$, Competition was varied as shown in figure 5.3. Pallidal activity as summed over all channels and both internal and external segments (figure 5.7). This showed that the modified model maintained its sufficient account for the fMRI data, in that pallidal activity decreased with increasing Competition.

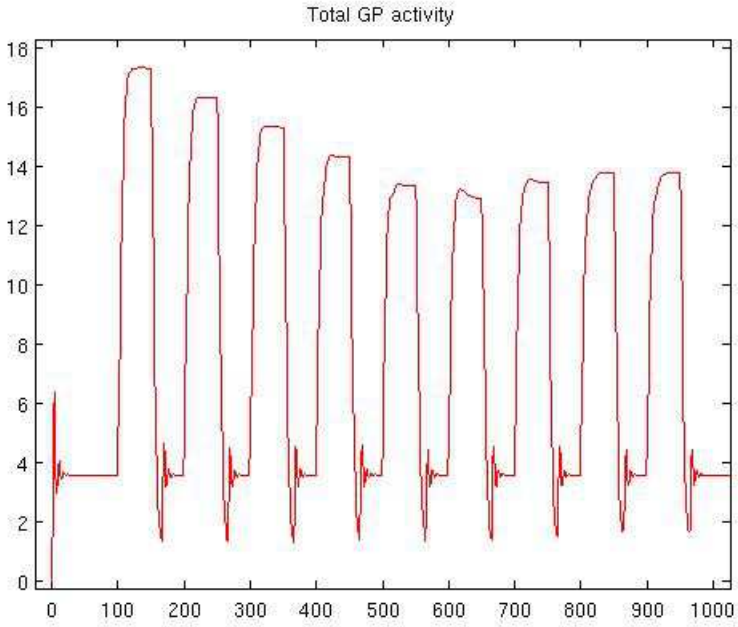


Figure 5.7: The effect of Competition on summed pallidal activation for the modified model. The x axis displays time in model time steps, where each ”on” period corresponds to a Competition level in the order described in 5.3. Pallidal activity decreases with Competition, as observed in the original fMRI data.

5.4.3.2 Striatal activation

The fMRI data suggested that striatal activation increased with both Selectivity and Competition. To assess the effects of Competition, Selectivity was held steady ($w_g = -0.75$), Competition was varied as described in figure 5.3, and striatal activity was summed over all channels for both the selection and control pathways (figure 5.8). It was found to increase with Competition, in accordance with the fMRI data.

To assess the effects of Selectivity, the same analysis as described in section 5.4.2.2 was performed, and activity summed over striatal MSN channels for both the selection and control pathways. The predicted quadratic relationship was found (figure 5.9), though the desired positive relationship was present when $-0.7 < w_g < -0.5$.

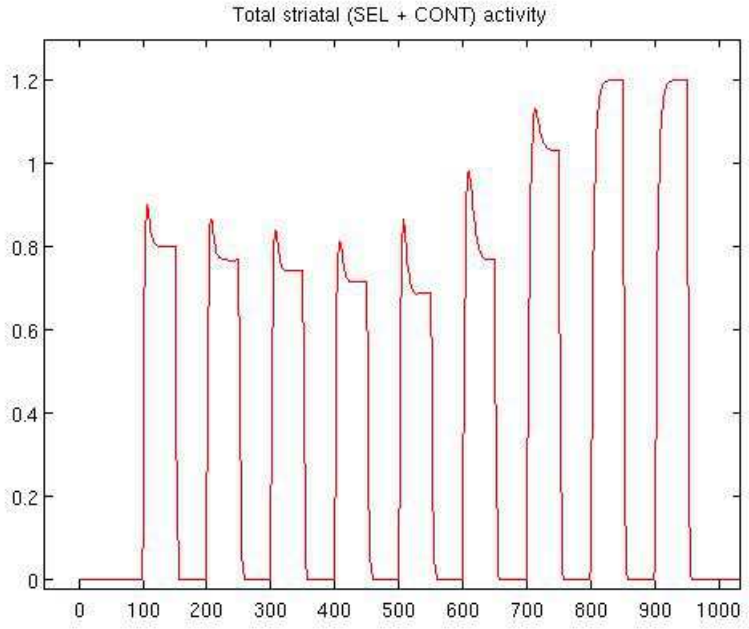
In principle, it appears that the modified model is able to provide a better account for the fMRI data extracted from the original study. A more rigorous assessment will be performed with the new data set acquired during the second study involving patients with schizophrenia.

5.4.4 Exploring model function

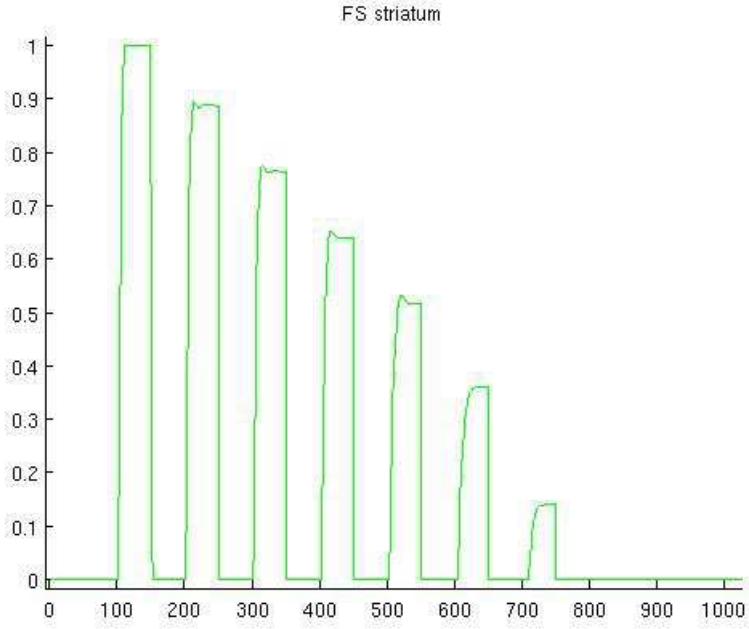
Given this modified, better-fitting model, it remains to examine why it behaves as it does, and what insights it may provide into the neurobiology of schizophrenia. The key additions to the model were:

- Tonic background cortical activity, to dynamically replace the previous biased activation threshold terms ϵ', ϵ_e and ϵ_g
- The explicit inclusion of a dopaminergic midbrain, regulated by both background cortical activity and internal globus pallidus
- A network of feedforward fast spiking inhibitory interneurons within striatum, regulated by a pallidostriatal projection originating from GPe

The initial model produced four hypotheses, some of which were supported by the data, whereas others were not. This modified version added new hypotheses to be tested during the second study: to differentiate them from the originals, these will be identified alphabetically.



(a)



(b)

Figure 5.8: The effect of Competition on (a) striatal MSN (selection and control pathways) and (b) FSI activity for the modified model, summed across all channels. The x axis displays time in model time steps, where each ”on” period corresponds to a Competition level in the order described in figure 5.3. Striatal MSN activity increases with Competition, in accordance with the original fMRI data. This is due to a drop off in inhibition imposed by FSIs.

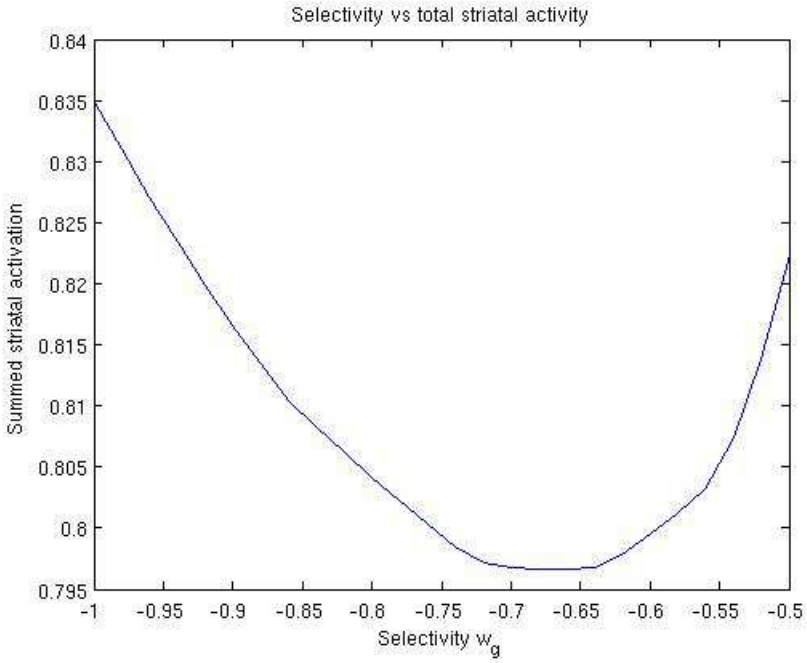


Figure 5.9: The effect of Selectivity on striatal activation, summed over all channels within the selection and control pathways. Again, the relationship is markedly quadratic, with activation and Selectivity having a positive relationship when $-0.7 < w_g < -0.5$.

5.4.4.1 Dopamine regulation

Dopamine has a profound affect on how this model performs. Activation along each SNc channel is defined by:

$$x_i^d = \epsilon_d - w_d c_i - w_{gs} y_i^g$$

Therefore it must be determined how dopamine release varies according to (a) background cortical activity c_{ton}., and (b) Competition between two competing channels, which will affect GPi activity, y_i^g. With w_g held constant at -0.75, both Competition and c_{ton} were varied, and peak SNc activity sampled, summed over all channels. This revealed that both c_{ton} and Competition affect dopamine release, though there is no interaction between the two (figure 5.10). As tonic salience decreased, and Competition became narrower, SNc activity increased.

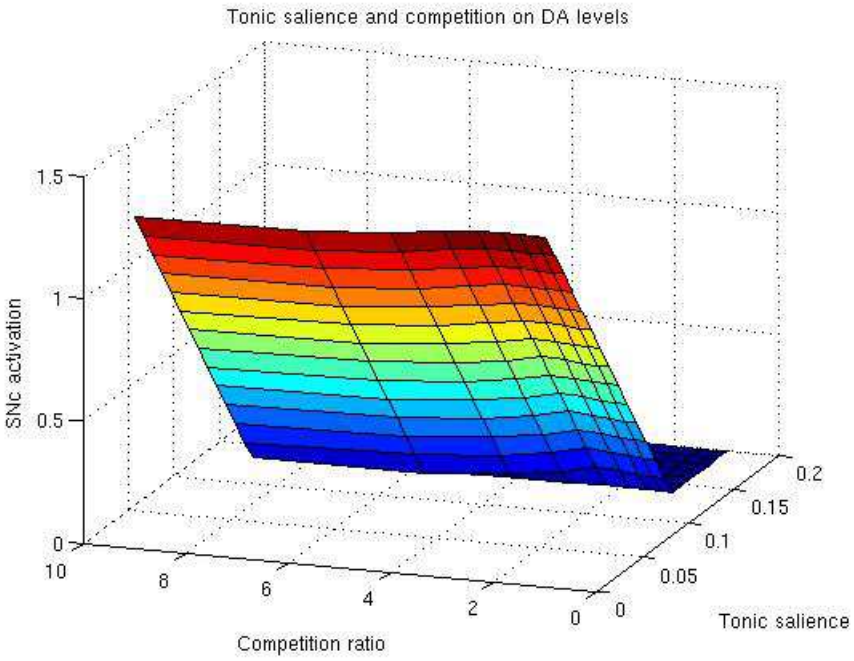


Figure 5.10: The effects of tonic background cortical activity and Competition on SNc activity, and hence dopamine levels. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. SNc activation is summed over all channels.

5.4.4.2 Dopamine action

Having established that both lowering tonic salience and increasing Competition raised dopamine levels, it was necessary to determine how this would affect selection performance. Two measures were used to investigate this:

- Point of promiscuity: The model was exposed to a series of finely-graded Competitions, where $c_1 + c_2 = 1.0$. The Competition for which dual selection was able to occur, that is, when the model disinhibited both channels c_1 and c_2 within the GPi, was noted as the point of promiscuity.
- Distractibility: Here, c_1 was held constant at a relatively weak value of 0.6, and c_2 gradually increased from zero until it was able to achieve dual selection, i.e. distracted the selection mechanism from maintaining the stronger signal in isolation.

Dopamine levels were modulated by varying tonic salience c_{ton}. As c_{ton} decreased, it became increasingly easy for two channels to be selected at once. Similarly, less salience was required to override an established channel (figure 5.11). Once tonic salience $c_{ton} \geq \epsilon_d$, dopamine levels were considerably reduced, the system became very strict, and no promiscuous selection

occurred. This produced a new hypothesis:

Hypothesis A *Both decreased cortical efficiency and increased Competition present situations where selection is made more difficult. They also both serve to disinhibit subcortical dopaminergic regions, with the resultant elevated dopamine acting to produce a more flexible, promiscuous selection mechanism*

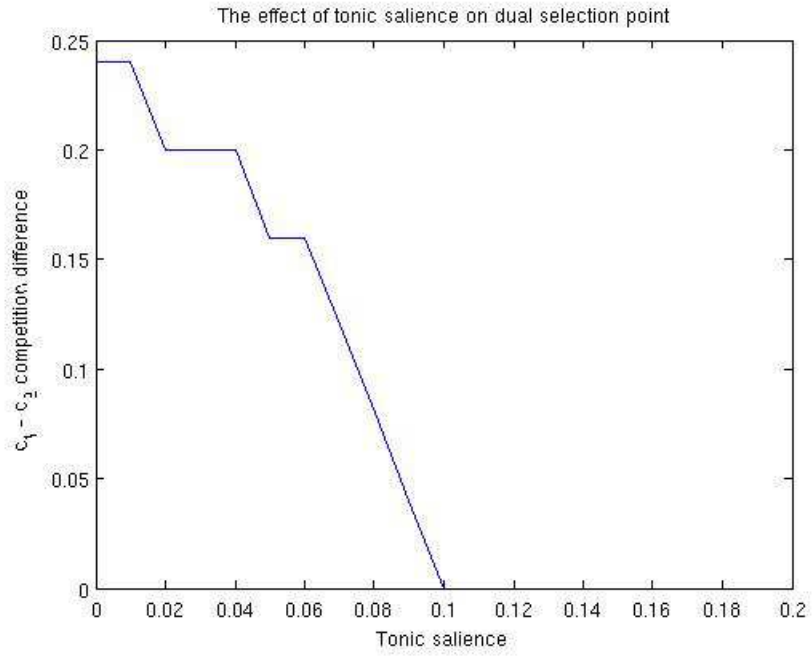
In order to allow for the model's behaviour to be more fully explored - in particular its dopaminergic interactions - c_{ton} was set to 0.07 for all subsequent simulations.

5.4.4.3 Fast spiking inhibitory interneurons

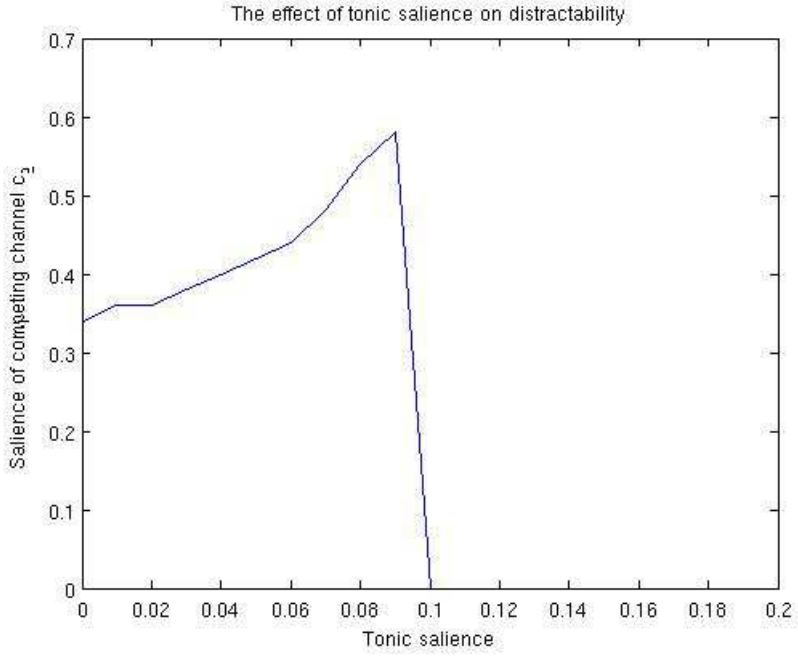
The addition of fast spiking inhibitory interneurons into the striatal system introduced several new parameters. The effects of each were addressed in turn, specifically (a) how they influenced the effects of Selectivity and Competition and (b) how this altered both summed striatal and pallidal activity, and the action selection performance measures. For all simulations, parameter in question was varied incrementally between a range of values, and the all non-manipulated parameters were held at their default values.

The cortico - FS interneuron projection, w_f

Striatal activation The degree of excitation the FS interneuron received was expected to have a significant impact on subsequent action selection function and BG activation. First, its impact on striatal activation was examined, and the interaction this had with Competition (figure 5.12). In general, striatal activity increased as w_f decreased, as expected. w_f was crucial in maintaining the positive relationship between Competition and striatal activation: if it dipped below 0.5, the relationship was reversed. w_f also had a similar profound effect on how Selectivity affected striatal activation: below 0.5, the effects of Selectivity were completely removed. In other words, once the influence of the new FS inhibitory interneurons was diminished below a certain point, the model's behaviour reverted to that of its original incarnation, and its ability to fit the fMRI data was lost. The FS interneurons would therefore seem to be a useful elaboration of BG architecture.

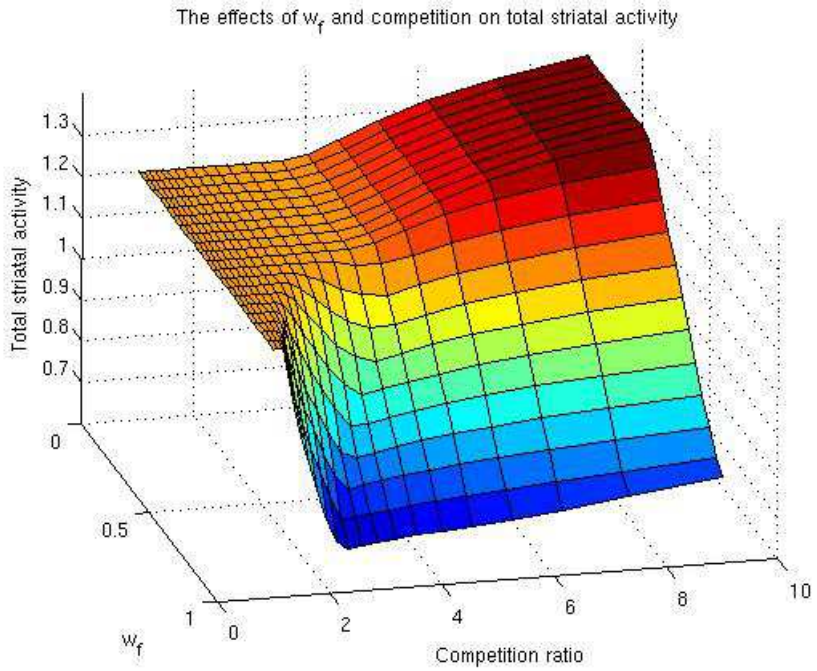


(a)

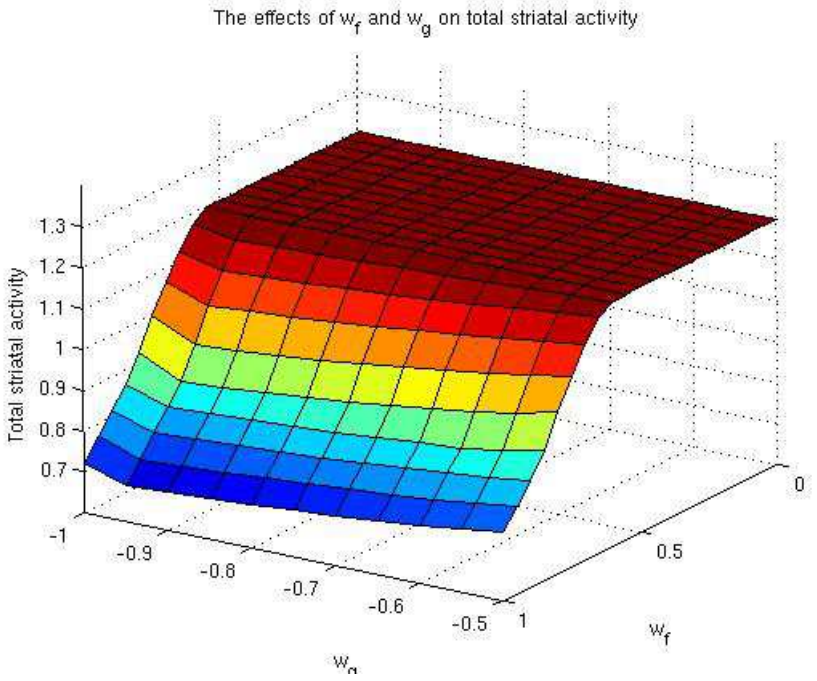


(b)

Figure 5.11: The effect of background cortical activity on selection performance. (a) Point of promiscuity: here the y axis states the difference between two competing channels for the point where both are simultaneously selected. As tonic salience decreases, the system becomes more promiscuous, and dual selection is made easier. (b) distractibility: Here $c_1 = 0.6$, and c_2 is gradually increased until dual selection occurs. The y axis denotes c_2 for the point when this happens. As tonic salience decreases, an increasingly weaker competitor is able to override the established c_1 channel.



(a)



(b)

Figure 5.12: The effects of w_f on total striatal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$. Therefore smaller values describe higher Competition. As noted earlier, striatal activation increases with increasing Competition (decreasing Competition ratio). However, as cortical excitation of the FS interneurons decreases, this no longer remains the case: if $w_f < 0.5$, the relationship between Competition and striatal activation is reversed. (b) The interaction with Selectivity. Selectivity increases as $w_g \rightarrow 0$. Again there is a pronounced shift as $w_f < 0.5$: below this, Selectivity has no effect on striatal activation.

Pallidal activation Increasing w_f acted to decrease total pallidal activation (figure 5.13). It also served to modulate the relationship this had with Competition, though in a less critical way than for striatum. Pallidal activation always decreased with Competition across the w_f range, though if $w_f > 0.8$, the intermediate Competitions resulted in pronounced pallidal suppression. With respect to Selectivity, w_f did not qualitatively alter its relationship with pallidal activation, but did serve to enhance it.

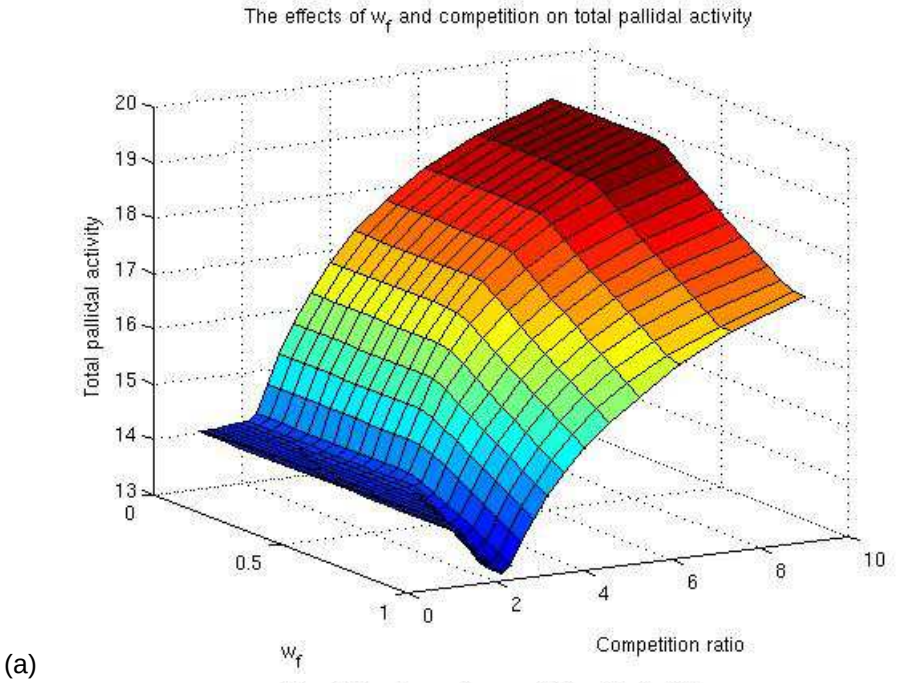
Action selection performance w_f was found to exert a fairly bimodal influence over selection (figure 5.14): when $w_f \leq 0.8$, the system had a degree of softness to it, with dual selection able to occur at a Competition ratio of 0.6:0.4; when $w_f > 0.8$, this decreased markedly.

The FS interneuron - MSN projections w_{f1} and w_{f2}

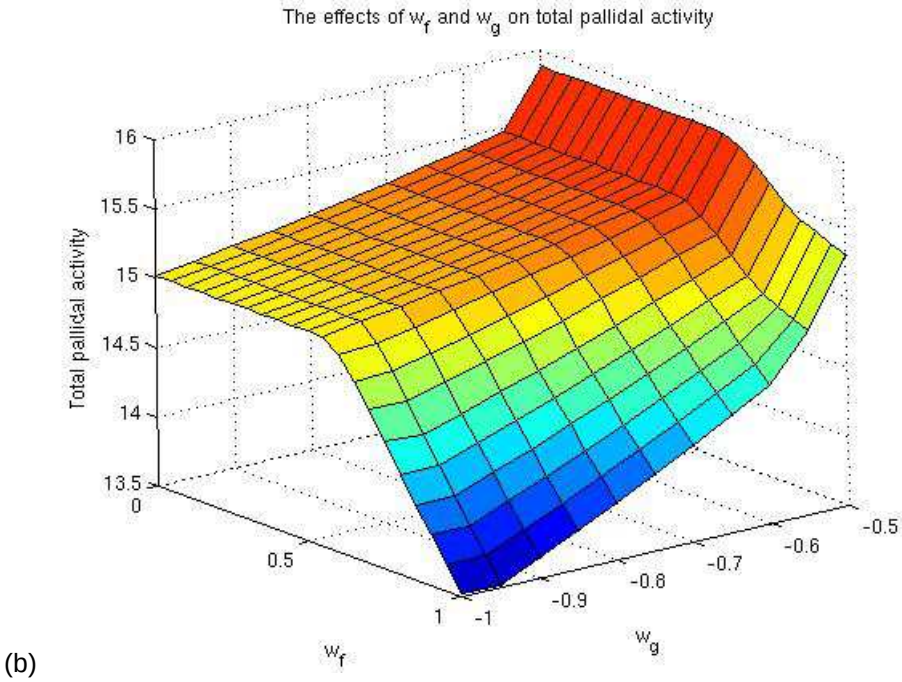
Striatal activation The projections from FSIs to the D1 receptor-dominated selection pathway (w_{f1}, figure 5.15) and D2 receptor-dominated control pathway (w_{f2}, figure 5.16) MSNs were found to decrease striatal activation as a whole. They enhanced, but didn't qualitatively alter the effect of Competition and Selectivity.

Pallidal activation w_{f1} raised pallidal activity as a whole, but did not affect the interaction with Competition or Selectivity (figure 5.17). w_{f2} on the other hand did strongly modulate the interaction between pallidal activation and Competition (figure 5.18): the desired relationship was lost if $w_{f2} > 0.2$.

Action selection performance w_{f1} and w_{f2} both served to gradually modify the system's strictness, though in opposite directions : w_{f1} increased strictness (figure 5.19), whereas w_{f2} decreased it (figure 5.20).

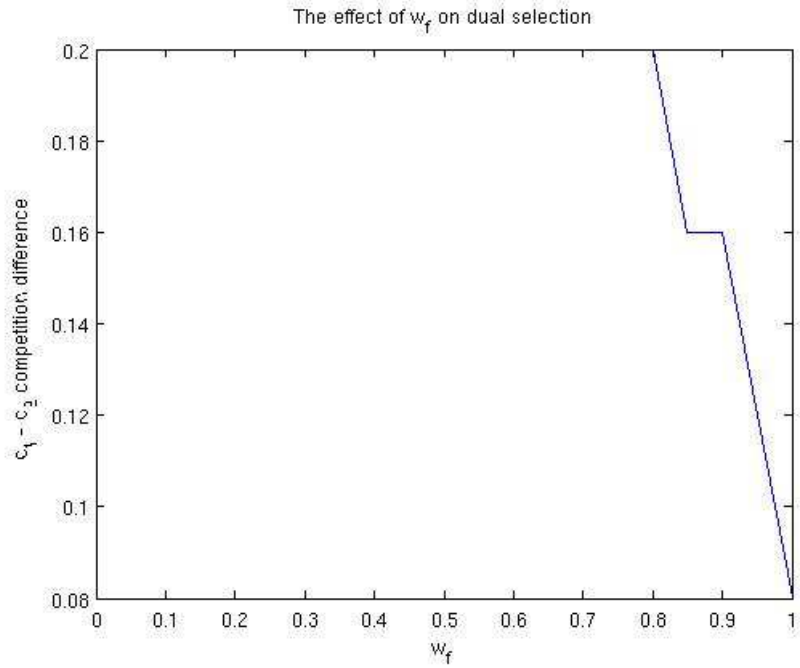


(a)

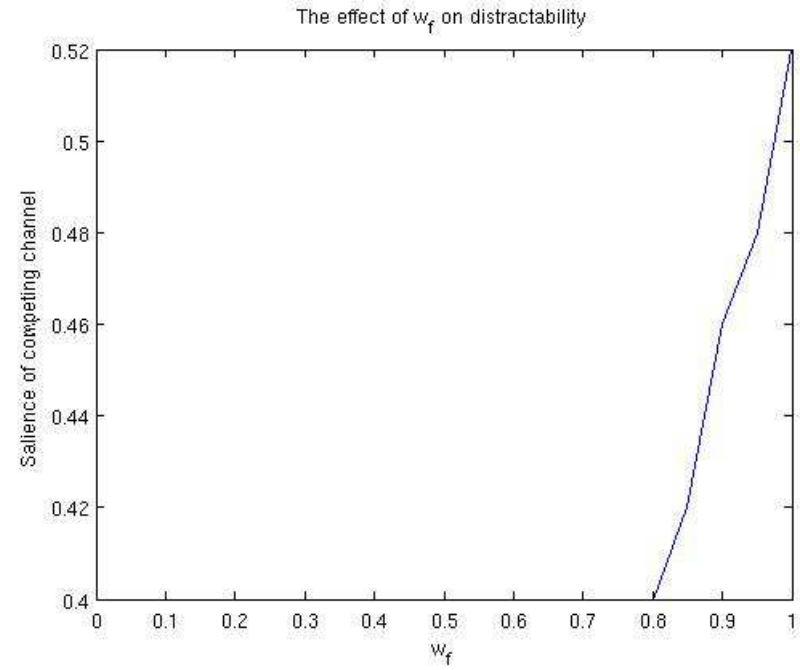


(b)

Figure 5.13: The effects of w_f on total pallidal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$. Therefore smaller values describe higher Competition. w_f acted to decrease pallidal activation generally - especially so at intermediate levels of Competition. However, the desired negative relationship between Competition and pallidal activation was maintained across the w_f range. (b) The interaction with Selectivity. Selectivity increases as $w_g \rightarrow 0$. Here, w_f appears to augment the positive relationship between Selectivity and pallidal activation.

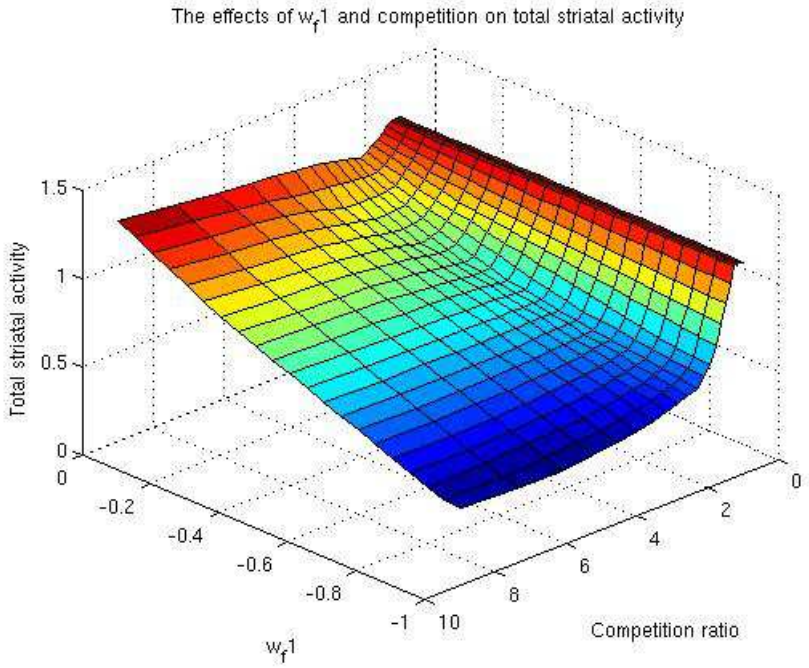


(a)

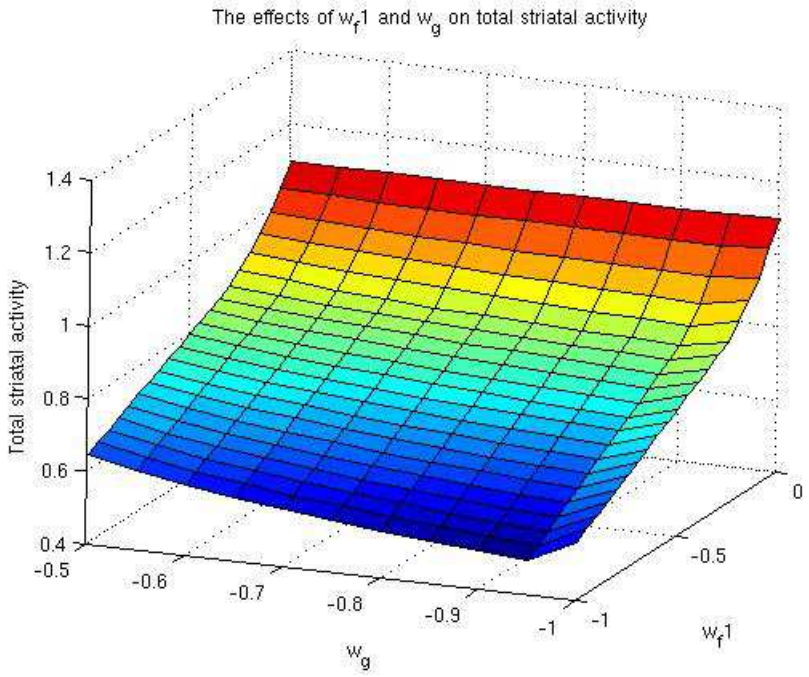


(b)

Figure 5.14: The effect of w_f on action selection performance. (a) The point of promiscuity: there is a small degree of permitted promiscuity for most of the range of w_f value, though the system becomes much stricter as $w_f > 0.8$. (b) distractibility: as for promiscuity, a competing channel with $c_2 = 0.4$ can override an existing channel with $c_1 = 0.6$ for most w_f values, until it exceeds 0.8.

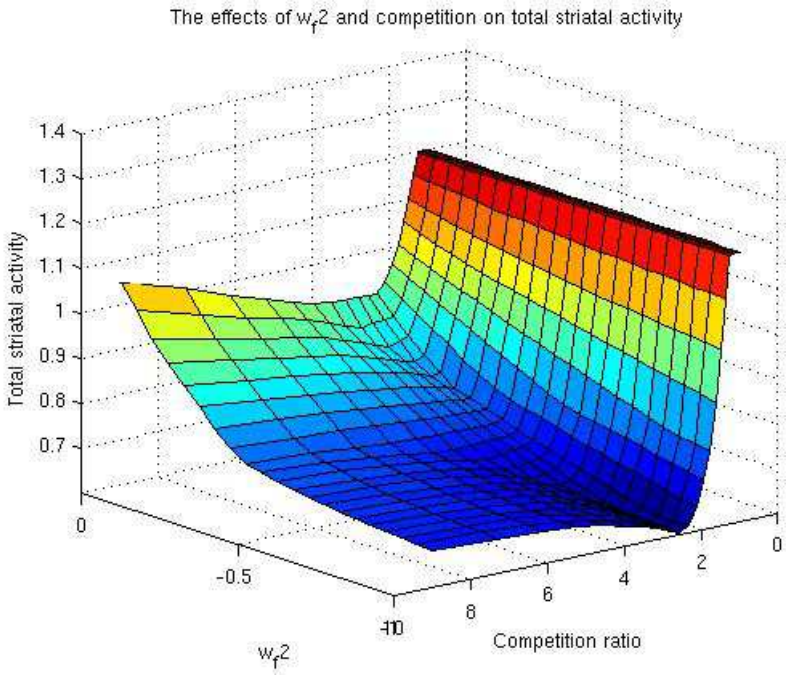


(a)

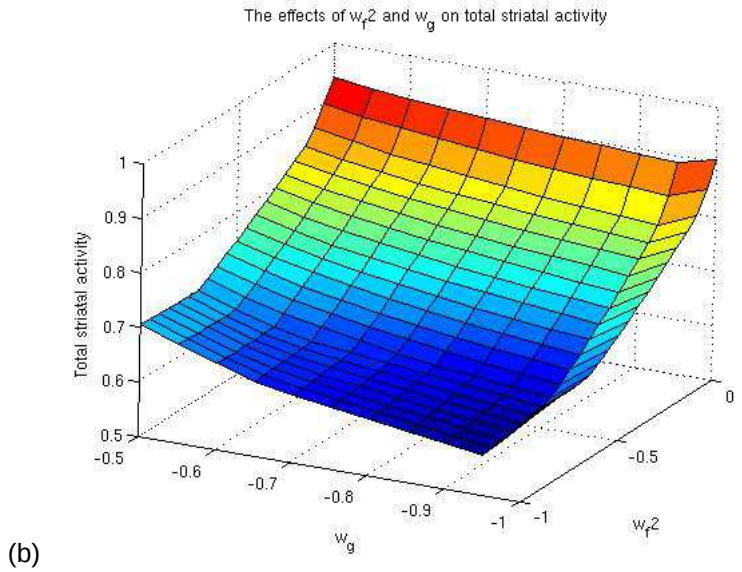


(b)

Figure 5.15: The effect of w_{f1} on striatal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. w_{f1} enhanced the positive relationship between Competition and striatal activation, but decreased striatal activation overall. (b) The interaction with Selectivity. Selectivity increases as $w_g \to 0$. There is no notable interaction with w_{f1}.

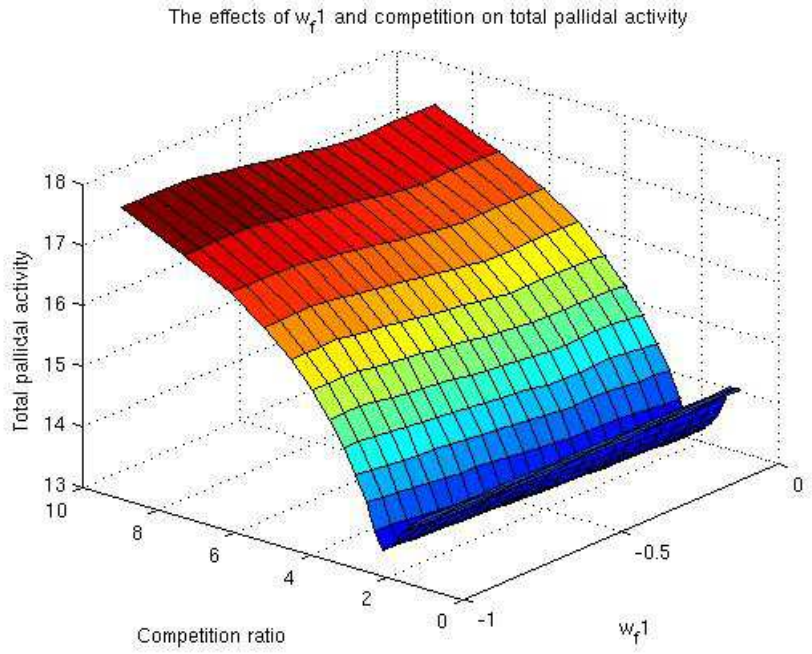


(a)

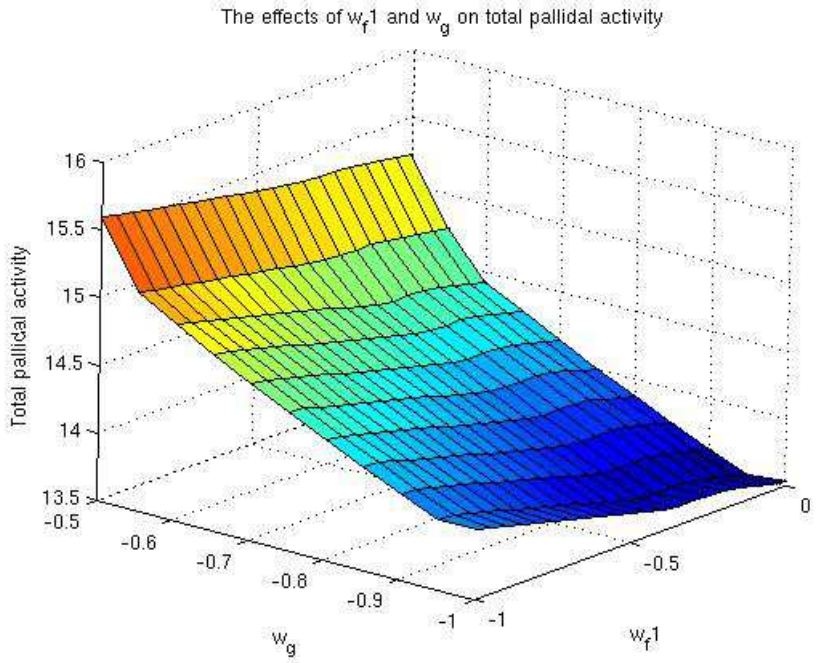


(b)

Figure 5.16: The effect of w_{f2} on striatal activation. Again, w_{f2} serves to decrease striatal activity, whilst enhancing the interactions with both (a) Competition and (b) Selectivity. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. Selectivity increases as $w_g \rightarrow 0$.

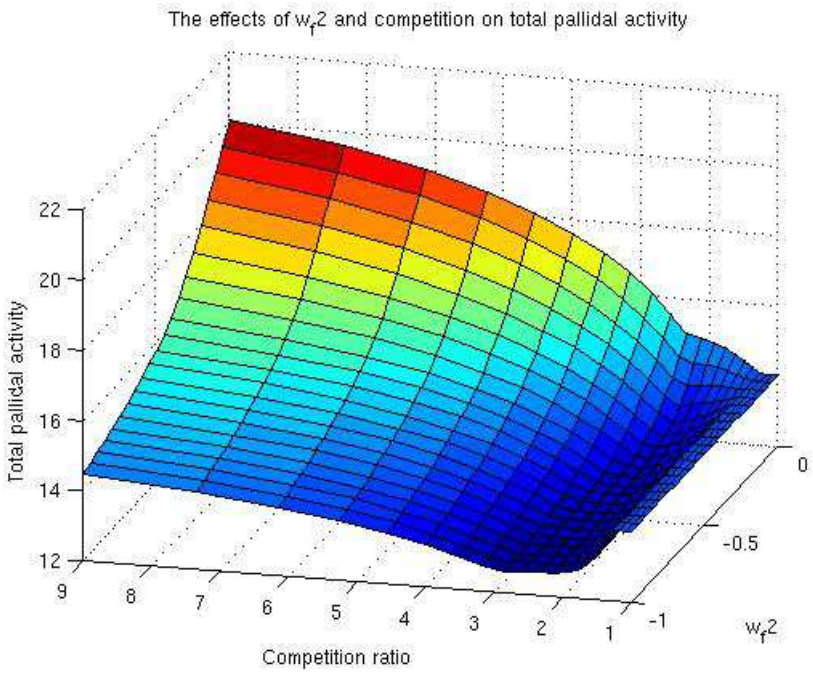


(a)

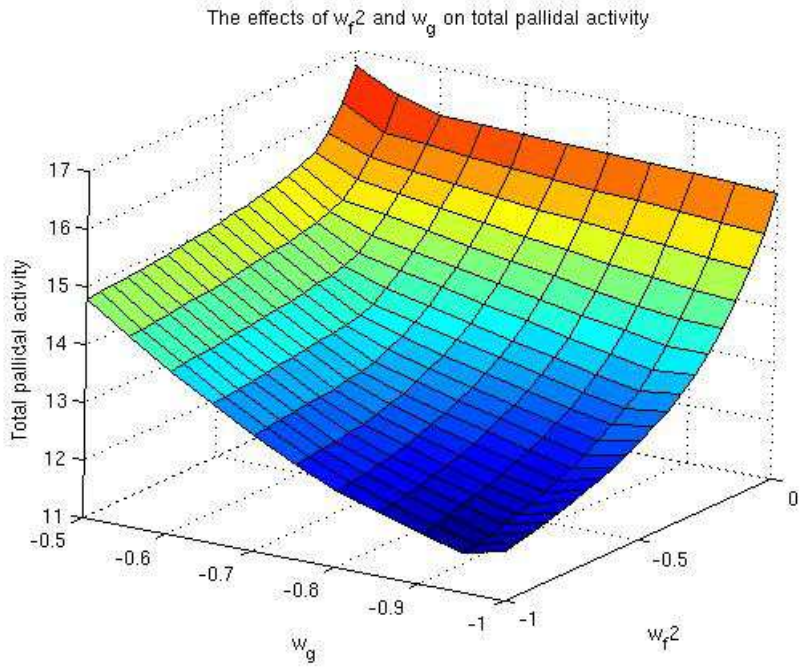


(b)

Figure 5.17: The effect of w_{f1} on pallidal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. w_{f1} is not affecting the interaction. (b) The interaction with Selectivity. Selectivity increases as $w_g \to 0$. There is no prominent interaction

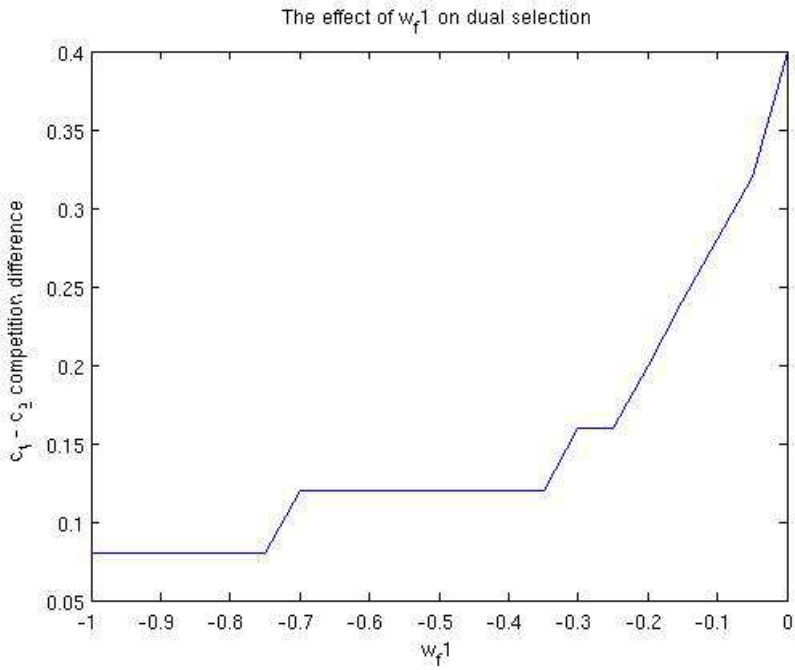


(a)

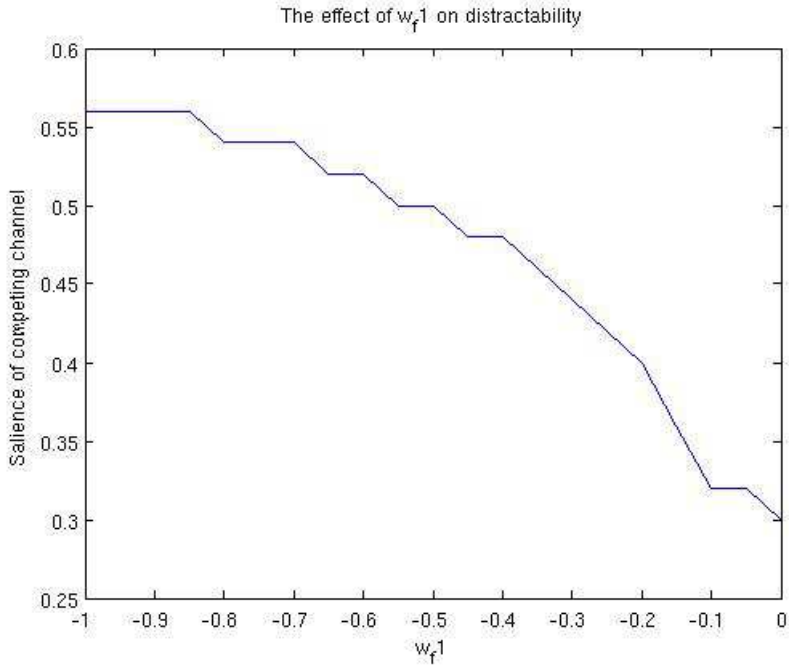


(b)

Figure 5.18: The effect of w_{f2} on pallidal activation: in general this exerts an inhibitory influence. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. The FSI - control pathway MSN projection markedly modulates this interaction, with the desired negative Competition-activation relationship only remaining at low values of w_{f2}. (b) The interaction with Selectivity. Selectivity increases as $w_g \rightarrow 0$. The interaction is enhanced by increasing w_{f2}.



(a)



(b)

Figure 5.19: The effect of the FSI - D1 receptor-dominated selection MSN (w_{f1}) projection on action selection performance. (a) The point of promiscuity and (b) distractibility. w_{f1} has a more gradual modulatory role, with decreasing values producing a much softer system.

5.4.4.4 The pallidostriatal projection

External globus pallidus' inhibition of striatal FSIs, w_{fg} The results of manipulating w_{fg} and its dopaminergic negative modulator λ_f were found to be identical: only the effects of w_{fg} are reported here.

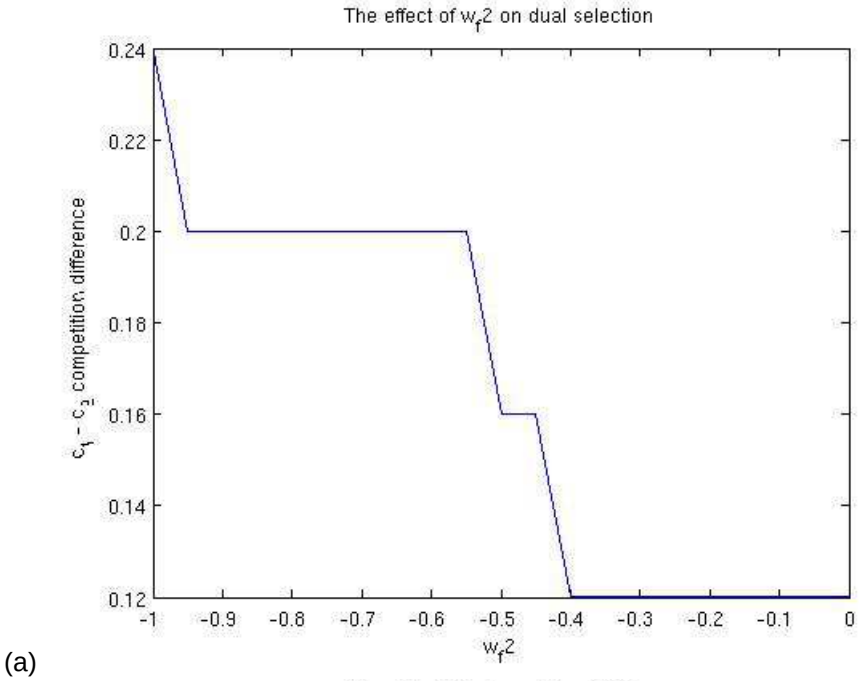
Striatal activation The desired positive relationship between Competition and striatal activation is maintained across a range of w_{fg} values, though once $-0.5 < w_{fg} < 0$, this response is increasingly flattened. A similar situation arises for the interaction with Selectivity, though there is no distinct cutoff (figure 5.21).

Pallidal activation w_{fg} was not found to significantly alter the relationship between pallidal activation and either Competition or Selectivity - indeed it barely had any effect at all (figure 5.22).

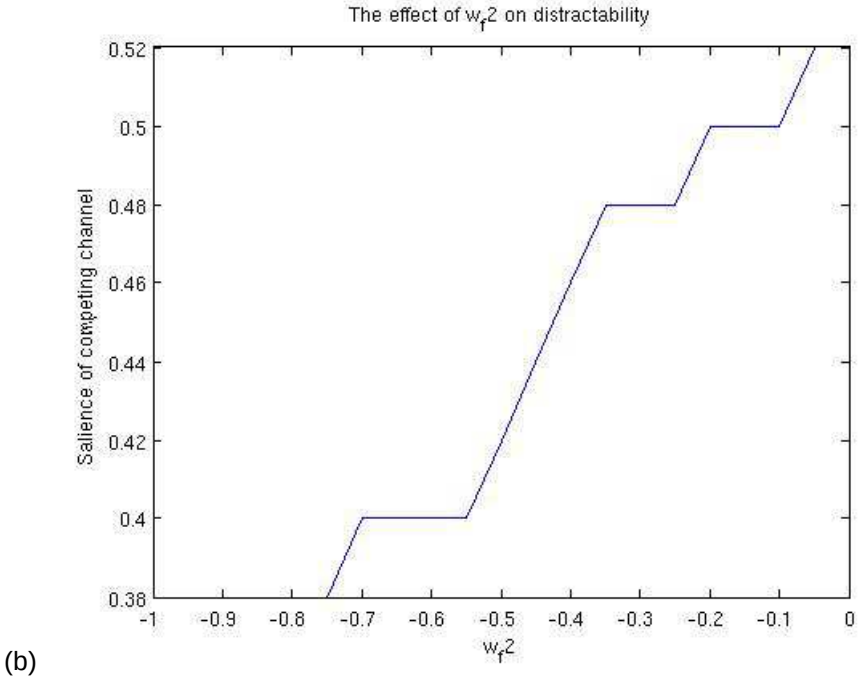
Action selection performance Like the FSI-MSN projections w_{f1} and w_{f2}, the pallidostriatal projection also seemed to be an effective modulator of action selection performance. As $w_{fg} \rightarrow -1$, the system became less strict, with less intense Competitions resulting in dual selection, and lower saliences able to perturb established channels (figure 5.23).

5.4.4.5 Summary

Increasing dopamine levels served to make the system less strict, and selection easier. Dopamine was increased by both cortical inefficiency (the presumed means by which tonic cortical activity is diminished), and when Competition is intense. These are both situations where selection will be more difficult. The relation that these conditions have with rising dopamine levels seems to allow the system to adapt and so facilitate selection in situations where it would otherwise become too strict, possibly underlying negative symptoms.

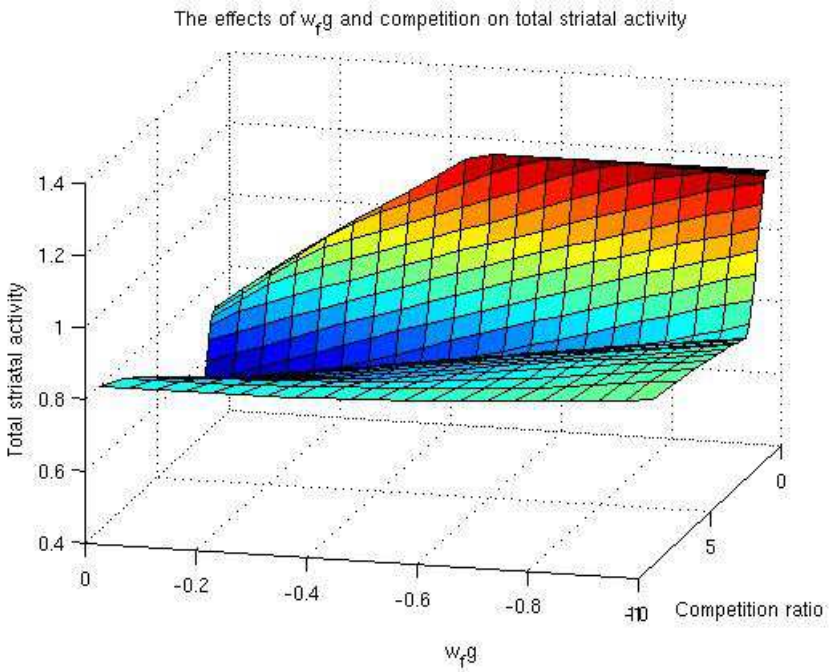


(a)

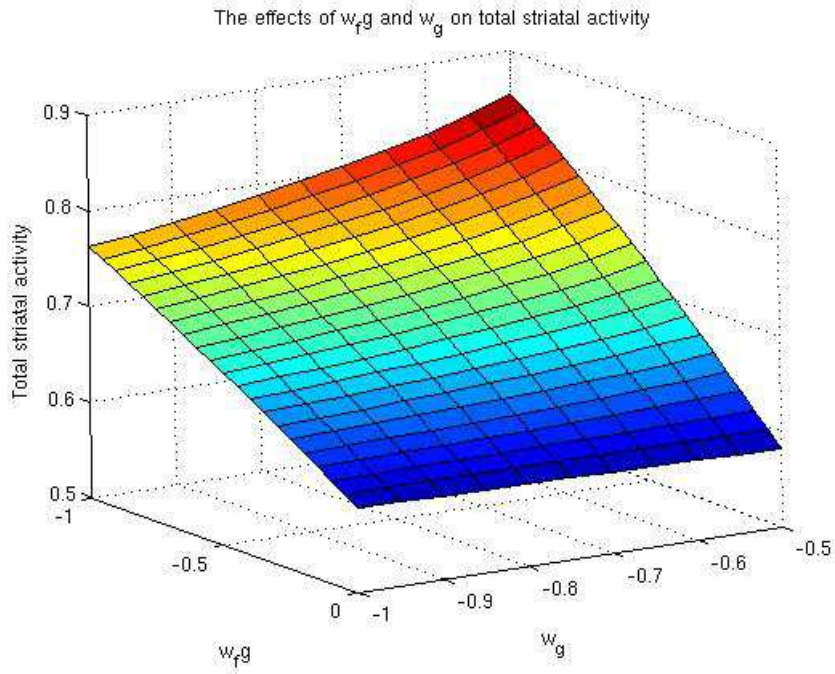


(b)

Figure 5.20: The effect of the FSI - D2 receptor-dominated control MSN projection (w_{f2}) on (a) the point of promiscuity and (b) distractibility. This too exerts a gradual influence, though in the opposite direction to w_{f1}: the system becomes less strict with increasing connectivity strength.

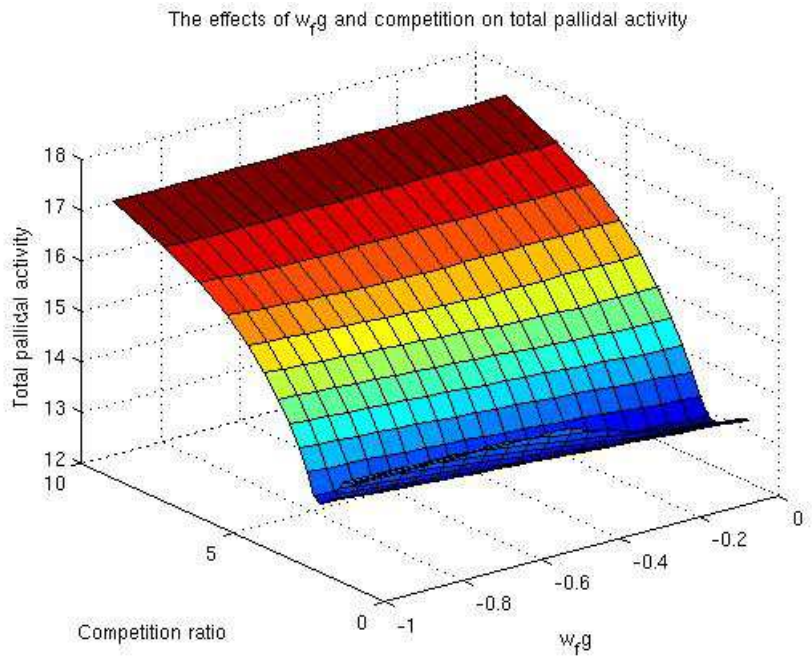


(a)

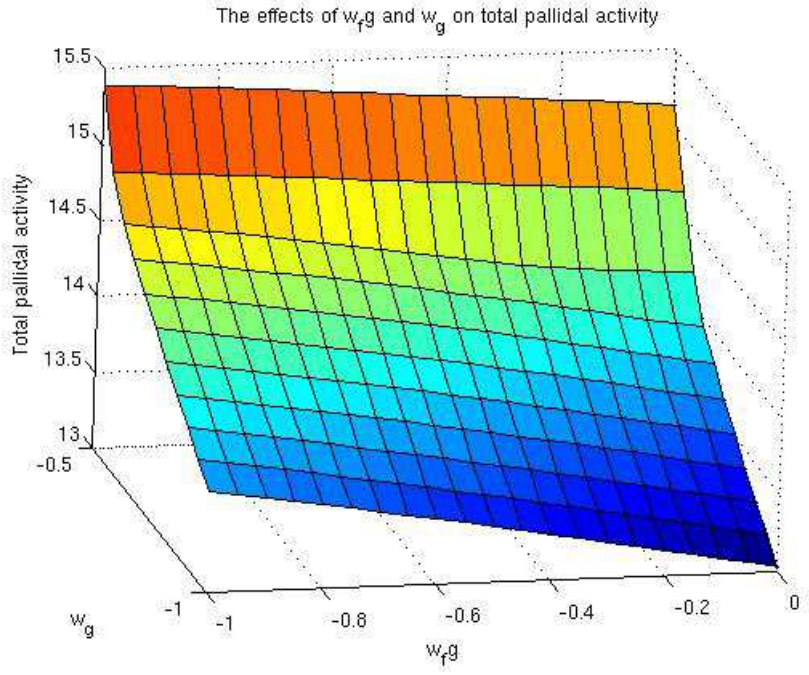


(b)

Figure 5.21: The effect of w_{fg} on striatal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. (b) The interaction with Selectivity.

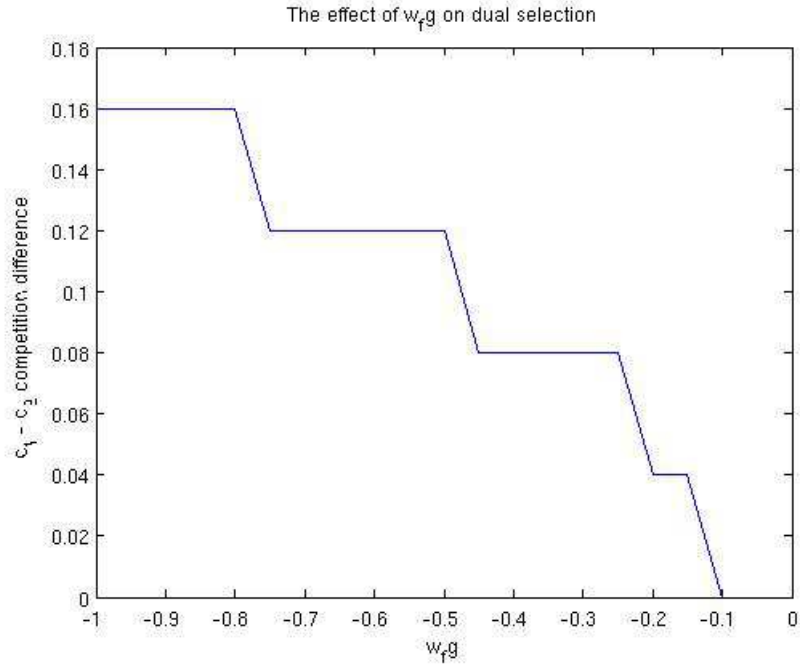


(a)

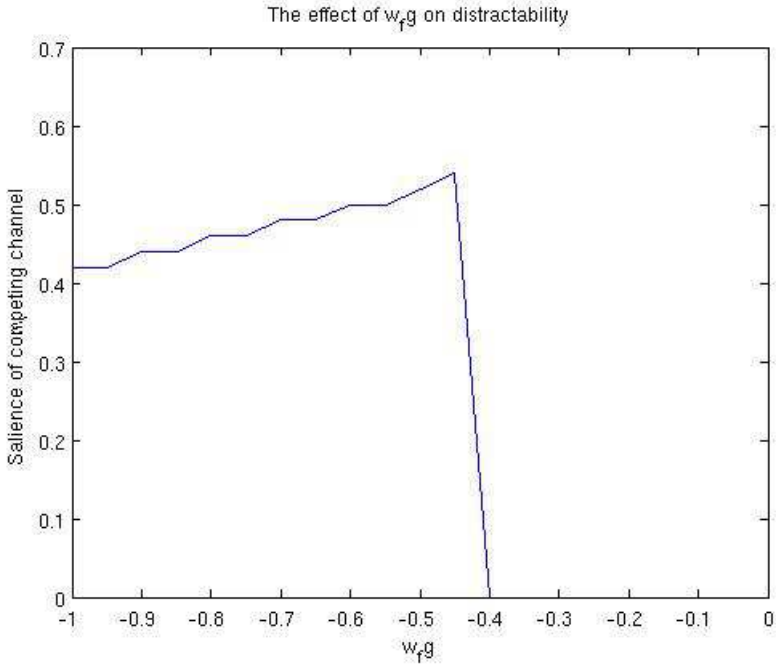


(b)

Figure 5.22: The effect of w_{fg} on pallidal activation. (a) The interaction with Competition. Competition ratio is defined as $c_1 : c_1$, therefore smaller values describe higher Competition. (b) The interaction with Selectivity.



(a)



(b)

Figure 5.23: The effect of the pallidostriatal projection (w_{fg}) on (a) the point of promiscuity and (b) distractibility. The system becomes less strict with increasing w_{fg}, with distraction becoming impossible when $-0.4 < w_{fg} < 0$.

Higher values of both the cortico-FSI (w_f) and pallido-FSI (w_{fg}) projection weights were crucial for improving the modified model's account of striatal activation with respect to both Selectivity and Competition. This supports the inclusion of FSI. As the desired patterns of activation were achievable over a significant range of parameter values, this suggests this architecture is a stable one.

Improving the model's fit for pallidus' relationship with Competition was dependent on one of the two FSI - MSN projections: that of the control pathway ($-0.2 < w_{f2} < 0$). This was a much more constrained parameter specification, but in line with neuroanatomical evidence that FSIs preferentially target the D1 receptor-dominated selection path [519], i.e. that $w_{f1} > w_{f2}$.

In terms of action selection performance, both the cortico-FSI projection w_f, and its subsequent suppression of D1 receptor-dominated selection pathway MSNs (w_{f1}) acted to increase strictness, whereas the pallido-FSI projection and D2 receptor-dominated control pathway suppression made the system more promiscuous. In the context of schizophrenia, if it is indeed the case that the primary pathology is one of glutamatergic inefficiency, leading to negative symptoms, then both disinhibition of subcortical dopamine levels and modulation of the control pathway could represent an attempted compensation. The question is, at what point does the compensation fail, or itself become part of the problem?

5.5 Considering the pathophysiology of schizophrenia

An aim of this project's modelling work was to assess how faults at different points of the network drive compensatory alterations in activity and behaviour over time. These changes are equated to specific symptomatic states, indicating excessively rigid (negativity, stereotypy and monoideation) or loose (hallucination, thought disorder and distractibility) selection mechanisms. By selectively "lesioning" this simplified representation of a recurrent, dynamical system, it is hoped that insight can be gained into some of the more paradoxical findings concerning schizophrenia (for example, hypo/hyperfrontality), and the confounding effects of medication.

The complex and multifactoral nature of schizophrenia means isolating the dominant factors is challenging. The plethora of neurophysiological and anatomical abnormalities that are associated with the disease could be conceptualised as falling into a relatively small number of camps. In this modelling work, the main underlying factors of schizophrenia were reduced to two areas: a generalised inefficiency of cortical glutamatergic function, resulting in both decreased cortical background activity and a reduced SNR; and an over-efficacy of striatal

dopamine transmission, especially related to D2 receptors.

5.5.1 Dopaminergic compensation for cortical inefficiency

That dopamine is dysregulated in schizophrenia is in little doubt. However, it remains controversial as to whether this is primary in itself, or a secondary consequence of upstream pathology and subsequent medication [532, 533]. Of particular relevance is the theoretical role of cortical glutamatergic hypofunction on midbrain dopamine activity [15]. Computational modelling provides us with the means to systematically investigate recurrently connected systems, in an attempt to address questions of this nature.

This model shows how increasing levels of Competition can lead to a disinhibition of dopaminergic regions, which acts to ”relax” the system and make selection easier. This would appear to be a facilitative function, allowing the system a degree of adaptivity in the face of varying demands. Similarly, this same principle applies to the potentially pathological situation of inefficient cortical glutamatergic transmission - this too is a situation where selection may be more difficult than usual, and can also induce an adaptive dopaminergic upregulation. In such a state, we would expect both tonic salience and the maximum salience any competing channel could take to be reduced - the overall effect would be one of reduced SNR. In figure 5.24 two situations were examined where peak salience operates at the maximal value of 1.0 (representing a ”healthy” state), and in a diminished state where it can only reach 0.6 (representing the patient state). If the default level of background tonic salience is taken to be 0.07, there is a marked increase in the system’s strictness for the diminished peak salience state, with a much smaller $c_1 - c_2$ interval being required for dual selection. However, this is corrected if tonic salience drops to 0.05, with the $c_1 - c_2$ interval being made equal to that of the ”healthy” model at $c_{ton} = 0.07$. So in the case of there being reduced glutamatergic efficiency, leading to reduced SNR, we see that without the dopaminergic compensation, selection would become much more difficult, perhaps underlying a predominantly negative symptom state in the affected patient. In this modified computational model, there is the innate capacity to compensate for these difficulties by disinhibiting dopamine release, thus normalising selection performance.

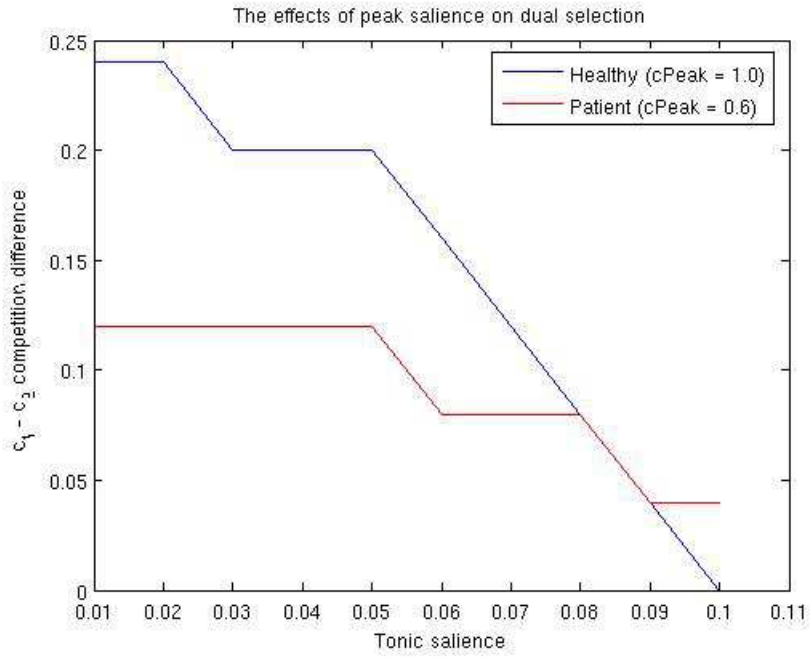


Figure 5.24: The effects of peak salience on the point of promiscuity, across the range of tonic salience values. Peak salience is the maximal activation a given cortical channel is able to achieve. The case where peak salience = 1.0 is denoted as the ”healthy” state, whereas the patient state is denoted with a peak salience of 0.6, reflecting reduced SNR. If a tonic salience value of 0.07 is taken as ”healthy”, then we can see that the patient system is more strict - a state that could be associated with negative symptoms and perseverance. However, if the patient system also experiences a reduction in tonic salience, say to 0.05, its performance is brought back into line with that of the healthy system, providing an effective compensation.

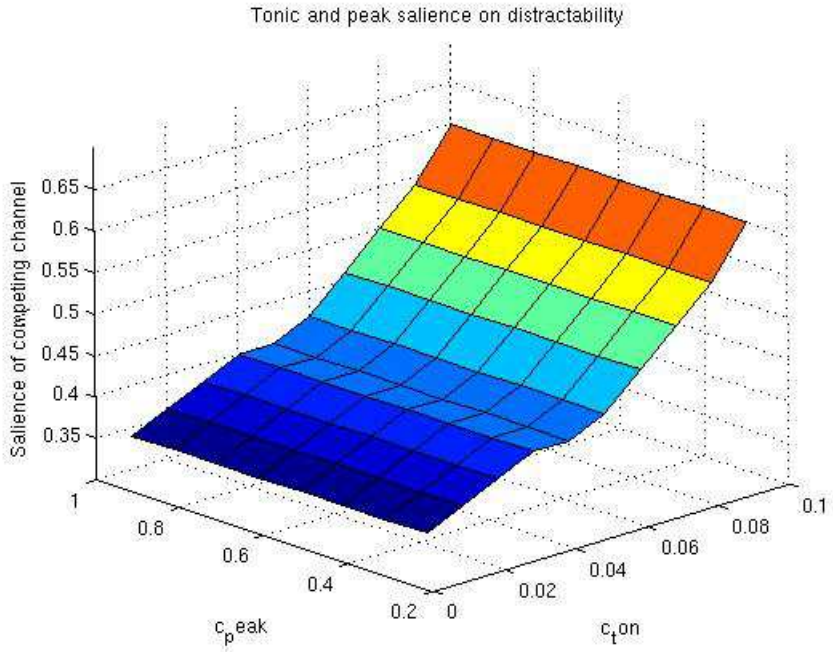


Figure 5.25: The effects of tonic and peak salience on distractibility. distractibility is measured as the minimum competing salience needed to override an established channel with salience 0.6. Peak salience has a negligible effect on distractibility, whereas tonic salience's effect is strong and positive - the higher dopamine that comes with lower background cortical activity leads to an overly-flexible and unstable selection mechanism.

5.5.2 Undesired consequences of the dopaminergic compensation

However, the effects of dopamine upregulation will not be isolated to altering the point of promiscuity. Figure 5.25 shows the effects of peak and tonic salience on distractibility. It can be seen than peak salience has very little effect here, whereas decreasing tonic salience acts to make the system much less stable. This could be interpreted as being analogous to more disorganised symptoms such as thought disorder, the experiencing of mental phenomena that would normally be suppressed, as well as additional negative symptoms such as loss of goal maintenance. This produces:

Hypothesis B *If taken to excess, the dopaminergic compensation induced by cortical inefficiency will lead to the manifestation of disorganised symptoms*

5.5.3 An alternative explanation: D2 receptor upregulation

Thanks to our understanding of the means by which antipsychotic medication works, increased D2 receptor efficacy also appears to drive the underlying pathophysiology of schizophrenia. Here we investigated what effects varying both the MSN- and FSI-associated D2 receptors have on action selection performance (figure 5.26). Increasing D2 efficacy on FSIs has very little effect on action performance, whereas increasing that of MSN neurons causes the system to become less strict. This is at odds with the view that the initial stages of schizophrenia are dominated by negative symptoms - the model suggests that D2 receptor upregulation is also part of a compensation mechanism to restore greater flexibility, rather than the primary insult itself.

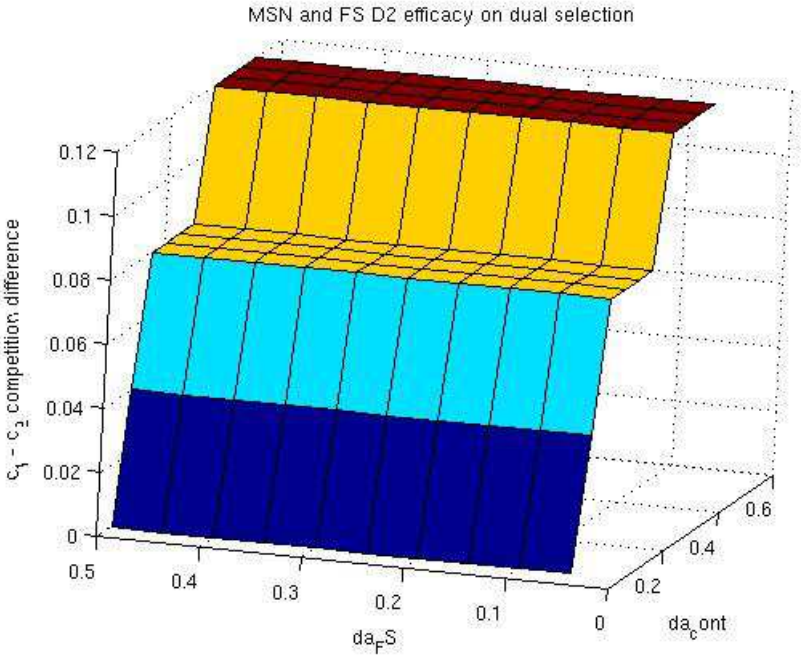
5.5.4 Modelled neuropathology and Competition/Selectivity

Although D2 receptor upregulation may be part of an attempted compensation rather than the original insult, it is still clearly part of the disease process. Further insight may be gained by establishing what roles cortical inefficiency and D2 upregulation have on Competition- and Selectivity-related activation. A series of simulations were run that examined the effects of reduced cortical efficacy and upregulated D2 efficacy in turn. In addition, the model's activation was considered under conditions of reduced D2 receptor efficacy, analogous to a state of overmedication.

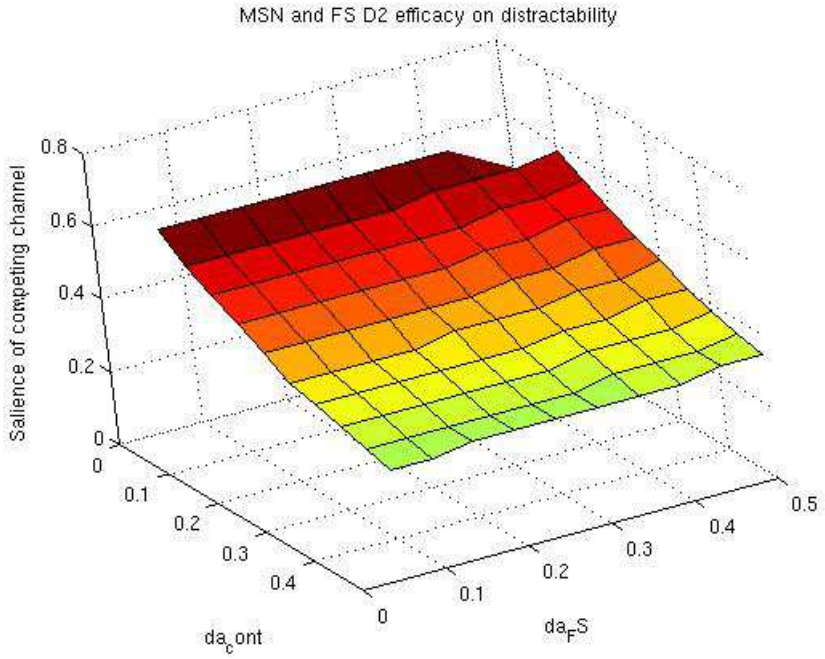
5.5.4.1 Striatum

In figure 5.27, it can be seen that the only scenario that maintains the "normal" relationship between total striatal activation and Competition involves altered D2 receptor efficacy. In situations involving reduced cortical efficiency, we see a marked alteration of this relationship, with striatal activation decreasing with increasing Competition. If the model proves able to provide a sufficient account of normal BG selection during action selection, then this finding presents an interesting opportunity, summarised as:

Hypothesis C *If patient striatal activation is seen to decrease with Competition, this suggests that reduced cortical efficiency is the more dominant factor in the illness' pathophysiology; if the*



(a)



(b)

Figure 5.26: The effects of D2 modulation of both MSNs (da_{cont}) and FSIs (da_{FS}) on (a) the point of promiscuity and (b) distractibility. FS interneuron D2 receptors play at best a marginal role in altering selection performance, whereas MSNs D2 receptors serve to make the system more loose/less strict, albeit subtly.

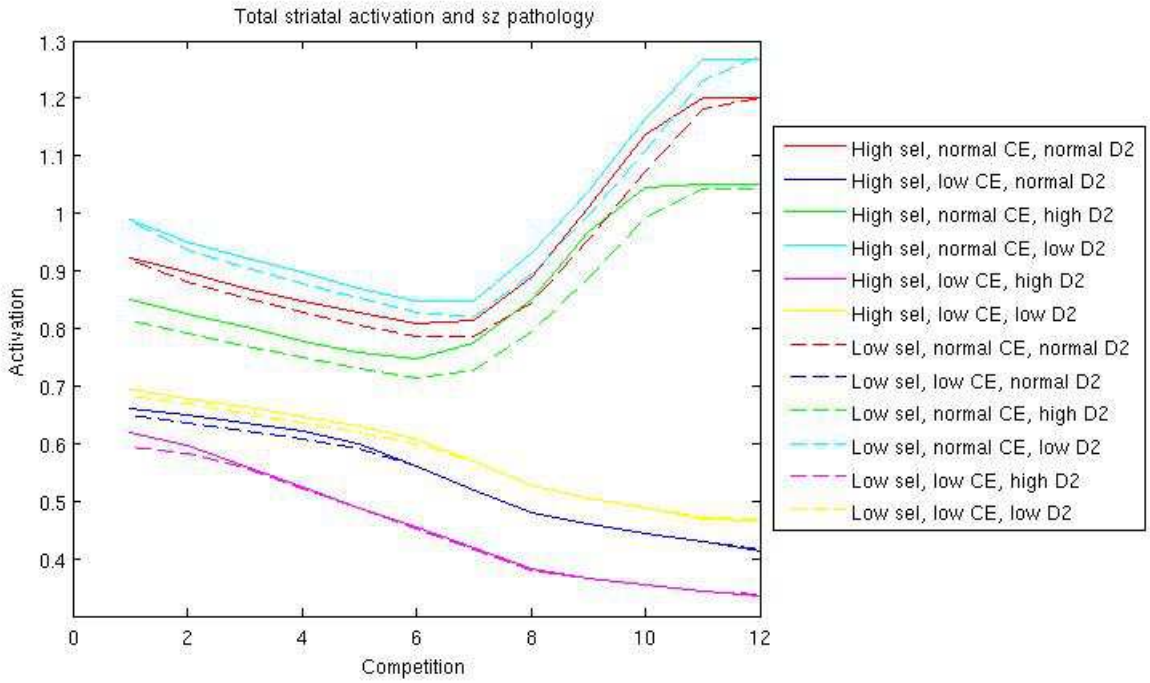


Figure 5.27: Exploring how varying parameters analogous to schizophrenia alter striatal responses to Competition and Selectivity. "Sel" means Selectivity; "CE" means cortical efficiency (reduced background activity and peak salience); "D2" means D2 receptor efficacy. The only situation that maintains the normal response to Competition is when cortical efficiency remains normal, and D2 receptors efficacy is modulated up or down (where down is the "overmedicated" state). All other scenarios involve reduced cortical efficiency, and demonstrate a qualitative shift in the response to Competition. The overmedicated state (low D2) paradoxically shows higher striatal activation regardless of cortical efficiency.

usual increasing relationship is observed, then cortical inefficiency is not playing the greater role.

This is particularly useful in that it is the patient's own relative responses that form the measure, with no reliance on group comparisons.

In all cases, increasing D2 efficacy - which is arguably an attempted compensation - acts to decrease striatal activation. Reducing D2 efficacy - which can be seen as analogous to overmedication - results in elevated striatal activation across the board. Again, pending model verification this could provide a useful means to fine-tune a patient's prescribed medication.

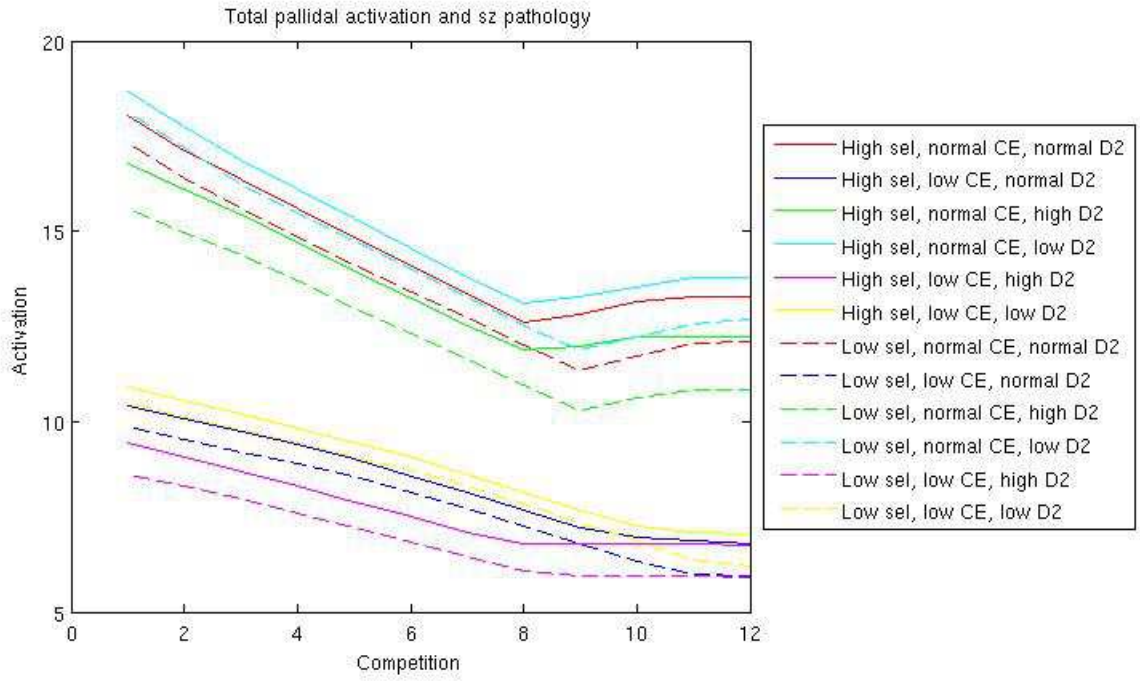


Figure 5.28: Exploring how varying parameters analogous to schizophrenia alter pallidal responses to Competition and Selectivity. ”Sel” means Selectivity; ”CE” means cortical efficiency (reduced background activity and peak salience); ”D2” means D2 receptor efficacy. For pallidus, no manipulation results in a qualitative shift in the response to Competition. The main effect is that cortical inefficiency results in reduced activation for all D2 manipulations.

5.5.4.2 Pallidus

Unlike striatum, pallidus did not display a clear bifurcation of Competition-related activation (figure 5.28). The only clear findings were that reduced cortical efficiency acted to reduced pallidal activation, and that the overmedicated pallidus was more active than its medication-free counterpart. As these effects would theoretically cancel each other out in a group of medicated patients, the model has less predictive value for pallidus.

5.5.4.3 Action selection performance

The next step was to establish what interaction cortical inefficiency and D2 dysregulation - and hence theoretical BG activation - would have with the point of promiscuity and distractibility. In these simulations, peak salience was set at the pathological value of 0.6. As Selectivity did not seem to have an interesting interplay with either cortical efficiency or D2 efficacy, this was

kept constant at the high level ($w_g = -0.75$).

Figure 5.29 examines the point of promiscuity. As noted earlier, both decreasing cortical efficiency and increasing D2 efficacy served to make the system less strict. There appeared to be an interaction: the effects of D2 became more pronounced as inefficiency increased (that is, tonic salience decreased). Indeed, when tonic salience was around "normal" levels ($c_{ton} \geq 0.07$), changes in D2 were only effective at very high levels.

The converse was also true when D2 dipped below the "normal" level of 0.2. If low D2 is seen as analogous to overmedication, the responses seen in figure 5.29 underlines the increased vulnerability patients with relatively low levels of schizophrenic neuropathology have to medication-induced Parkinsonism. Indeed the effects of overmedication are quite severe across the range, dramatically dragging the system down into a stiffer state.

The modulation of the point of promiscuity occurs over a relatively restricted range - this implies that the BG's ability to adapt to greater cortical inefficiency (via SNc disinhibition and D2 efficacy upregulation) is somewhat limited, and rendered ineffective if the neuropathology progresses beyond a certain point. It may be at these moments that a person becomes more floridly psychotic.

When considering distractibility 5.30, similar patterns are found. Increasing cortical inefficiency and D2 efficacy produce a looser system, though the changes become less effective at more pathological levels. Again, the pronounced effects of overmedication ($\lambda_e < 0.2$) suggest a slim margin for error when prescribing medication.

5.6 Fitting the model to the data

The majority of the model's parameters were set according to what was physiologically suggested, or in a manner that had consistency with previous parameter settings. However, the values of c_{ton} and λ_e, respectively representing cortical background activity and D2 receptor efficacy, were of especial interest given their presumed relationship with schizophrenic pathophysiology. Therefore it was decided that the optimal values for these parameters would be estimated by fitting the model to the fMRI data collected from healthy controls.

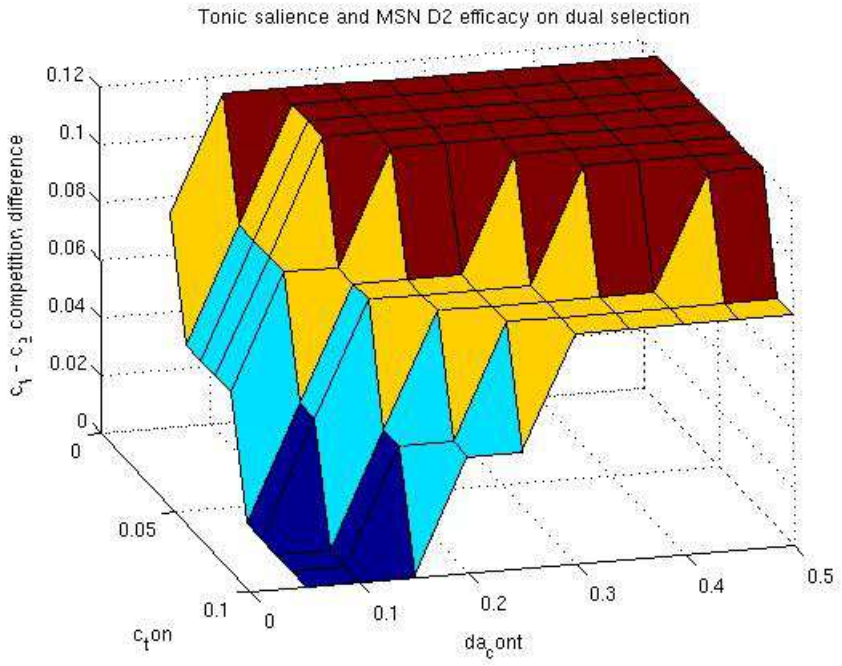


Figure 5.29: The effects of cortical efficiency and D2 dysregulation on the point of promiscuity. Both increasing D2 efficacy and decreasing cortical efficiency produces a less strict system. Interestingly, the effects of D2 modulation become more marked as cortex grows increasingly inefficient.

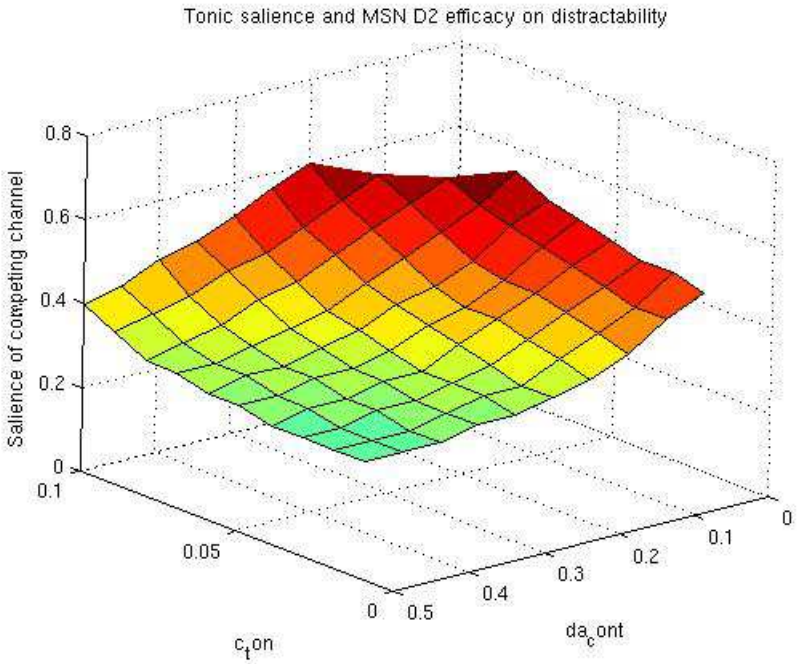


Figure 5.30: The effects of cortical efficiency and D2 dysregulation on distractibility. Both decreasing cortical efficiency and increasing D2 efficacy exert an equally distracting influence over the selection mechanism. The effects are more pronounced as $c_{ton} \rightarrow 0.1$, and $\lambda_e \rightarrow 0$, implying that as pathophysiology progresses, the effects of any attempted compensation by the BG are limited. As for the point of promiscuity, the increased gradient of the curve at low D2 levels implies that overmedication will exert a relatively sudden stiffening effect on system dynamics.

5.6.1 Methods

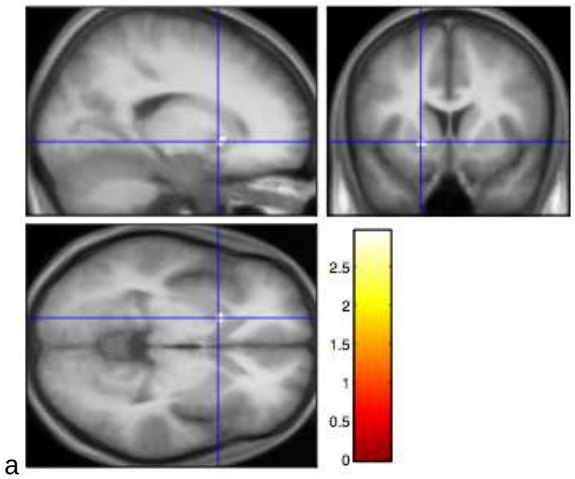
This was achieved by estimating multiple separate models for each participant, varying c_{ton} and λ_e.

- c_{ton} varied over 4 values: 0.025, 0.05, 0.075 or 0.1
- λ_e varied over 4 values: 0.1, 0.2, 0.3 or 0.4

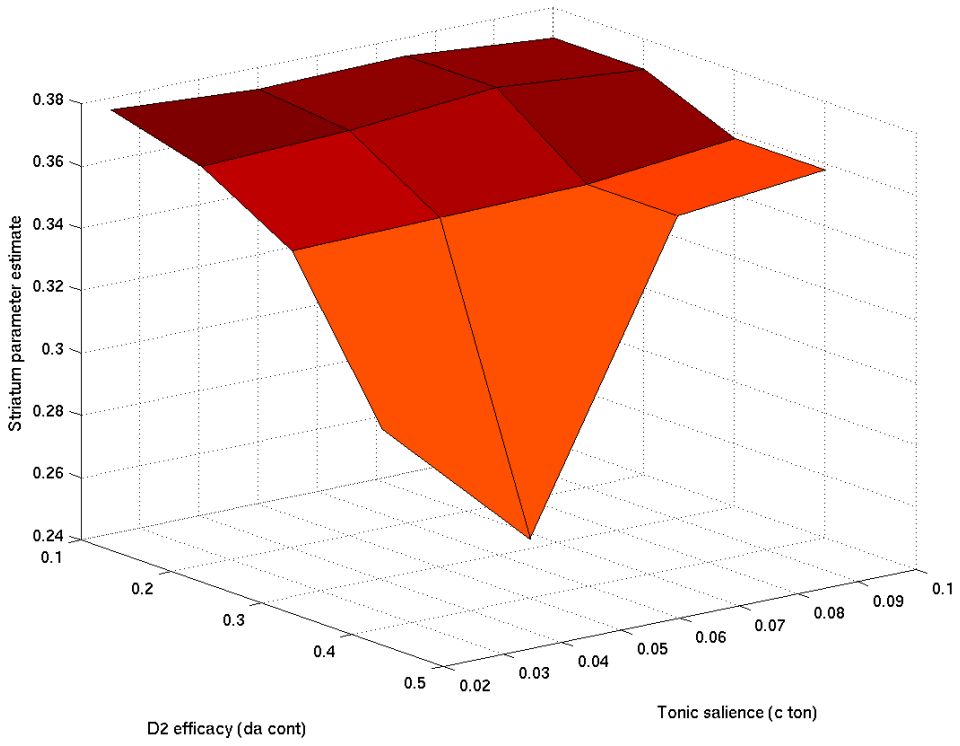
fMRI regressors for simulated striatal and pallidal activity were then constructed in the same way as described earlier (see section 4.1.4.5), based on each participant's trial order. These were both entered into single-subject design matrices, together with motion parameter estimates. Given the 4 x 4 parameter space, 16 different models were estimated for each participant. The t-contrasts for each of these points in parameter space were entered into a second-level t-test, and the voxel of peak activation located within either a striatal or pallidal ROI, appropriate to the regressor in question. Considering the striatum and pallidus separately, the contrast estimates of these peak voxel activations were then compared to determine which parameter set best represented the healthy control data set.

5.6.2 Results

For both the striatum and pallidus, the location of the voxels of peak activation barely varied across parameter space, ranging no more than 2 voxels from MNI -17 12 -2 for striatum, and MNI -18 -4 -4 for pallidus. This indicates a degree of stability in the model's representation of neural activation. The contrast estimates of these peaks were extracted separately for striatum (figure 5.31) and pallidus (figure 5.32). In this chapter, the neurobiology of schizophrenia was represented as a combination of cortical inefficiency (lowered tonic salience with concomitant reduced SNR), and elevated D2 receptor efficacy. This is partly born out by the results here, as the activation of healthy controls is better accounted for the further we move away from these theoretical pathological states.

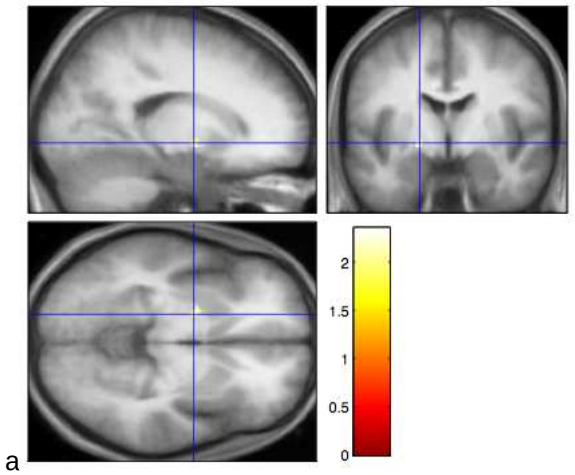


a

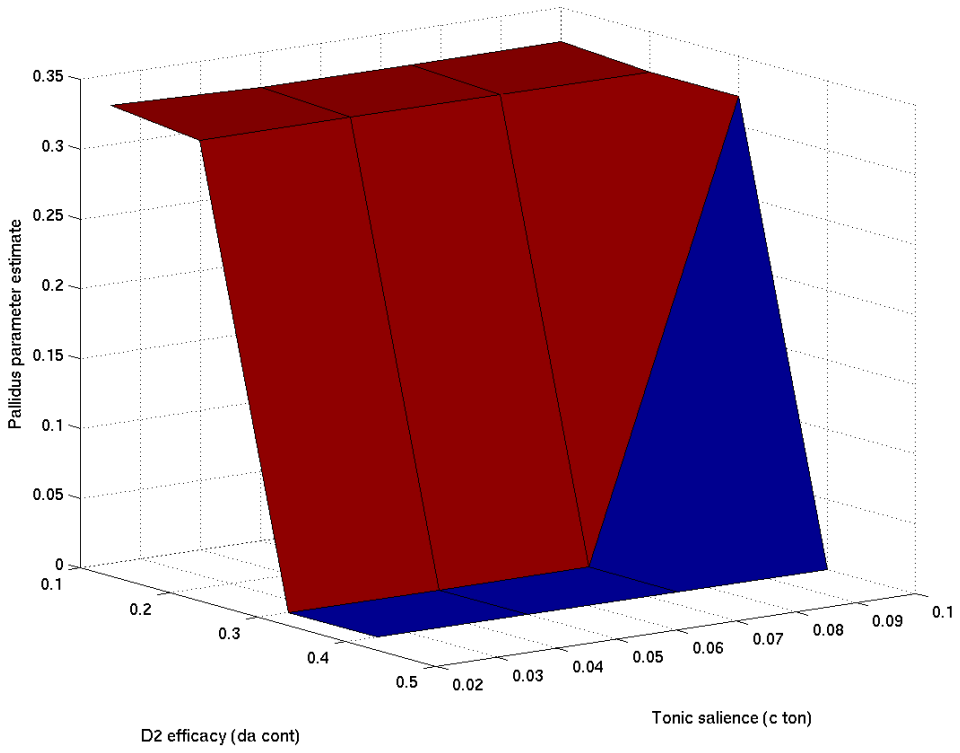


b

Figure 5.31: How varying the parameters c_{ton} (tonic salience) and λ_e (D2 efficacy) affected the new model's fit for the healthy controls' striatal activation. (a) Peak activation was observed in left anterior striatum, involving putamen and the cell bridges (MNI -17 12 -2). (b) In general, the model provides a better account for healthy activation as tonic salience increases, and D2 efficacy decreases.



a



b

Figure 5.32: How varying the parameters c_{ton} (tonic salience) and λ_e (D2 efficacy) affected the new model's fit for the healthy controls' pallidal activation. (a) Peak activation was observed in left ventrolateral pallidus (MNI -17 12 -2).(b) As was seen in striatum, the model's account improves as tonic salience increases, and D2 efficacy decreases.

5.7 Conclusions

In this section the original Gurney model has been modified in an attempt to provide a better account of the fMRI data obtained from healthy volunteers. While the modified model was able to retain the key characteristics of a disinhibitory action selector, both pallidus' response to Competition, and striatal responses overall, now bear a closer relation to the empirical evidence. Whether or not this modified model has been overfitted to the first data set remains to be tested: a second fMRI study involving both patients with schizophrenia and healthy controls should serve to elucidate this.

The modification comprised of a small number of components: the introduction of fast spiking inhibitory interneurons within striatum, whose firing was in part regulated by a GPe-based pallidostriatal projection; and the explicit implementation of the dopaminergic midbrain, regulated by both background cortical activation (a representation of cortical efficiency), and GPi. This introduced a number of new parameters, though only a few were found to exert a significant influence over the new model. Improving the fit for striatal data was found to be heavily dependent on the cortico-FSI (w_f) and pallido-FSI (w_{fg}) projections, remaining stable over a significant range of values. This indicated that the inclusion of FSIs was indeed a beneficial one. Improving the fit for pallidal data was found to be dependent on the control pathway's FSI-MSN projection (w_{f2}) being small, in line with the neurophysiological finding that FSIs preferentially target the D1-expressing MSNs of the selection pathway [519]. In terms of selection performance, it was found that increasing FSI activity (via w_f), and the effect it exerted on the selection pathway (w_{f1}), served to make the system more strict, whereas factors suppressing FSIs (w_{fg}), and mitigating its effect on the control pathway (w_{f2}) made it more flexible.

The inclusion of a dynamic dopamine system also imbued the system with greater adaptivity: dopamine levels rose in response to both high Competition, and decreased cortical efficiency. This acted to make selection easier, i.e. made the system more promiscuous, which would help to facilitate selection in these two difficult circumstances.

However, it was found that this modified model was vulnerable to the hypothetical features underlying the pathophysiology of schizophrenia. Both decreased cortical efficiency (in line with deficient glutamatergic transmission) and upregulated D2 receptor efficacy within striatum acted to make the system more flexible: the former by disinhibiting dopaminergic midbrain, and the later by undermining the control pathway. As cortical inefficiency was also associated with greater difficulty in selecting, these would seem to be a beneficial response at first. However, if raised too high, such a state could become associated with instability - and possibly symptoms

such as thought disorder, insertion, withdrawal, and a reduced ability to maintain a goal.

Further support was found for this representation of the neurobiology of schizophrenia by fitting the new model to the fMRI activation of healthy controls, optimising the parameters describing cortical efficiency (tonic salience, c_{ton}) and D2 efficacy (λ_e). The model's account for the data improved the further it moved away from the proposed schizophrenia-like states.

These changes were also associated with characteristic patterns of striatal activation - patterns that could potentially inform on both the primary insult in schizophrenia (glutamatergic vs dopaminergic), and whether or not a patient's current level of medication is appropriate.

Chapter 6

Study 2: Action selection in patients with schizophrenia

During the first study, the original Gurney model's ability to provide an account for BG activation was tested with a group of healthy volunteers. The varied results made it clear that some biologically-grounded modifications would need to take place These were subsequently implemented and used to interrogate the original fMRI data once more. However, the model's validity will remain highly questionable so long as it is only asked to provide an account for a single data set. Therefore, in order to further verify the new model, as well as test its specific predictions regarding schizophrenia, a second fMRI study was performed, during which matched groups of patients and healthy volunteers were scanned while performing the colour decision-making task. As this work proposes that schizophrenia results from the dysregulation of a central selection mechanism, a new PANSS-derived metric named "Instability" will be devised that will consider how each symptom could emerge. The appropriate PANSS item scores will then be used in the calculation of a rigid-unstable continuous metric.

6.1 Methods

6.1.1 Participants

Twelve patients with schizophrenia or schizoaffective disorder were recruited from the Royal Edinburgh Hospital, and twelve healthy volunteers from the staff, students and associates of

the University of Edinburgh. As the experimental task required a significant degree of sustained attention - especially during the training procedure - high-functioning, clinically-stable patients were targeted during recruitment, based on the assessments of an experienced consultant psychiatrist. Recruitment of healthy volunteers was guided by the need to match for age, gender and IQ. All participants were right-handed and native English speakers. Exclusion criteria included those normally associated with MRI; a history of head trauma or neurological illness; and abnormal colour perception (assessed using the Farnsworth D-15 test) or a history of eye disease. Additional exclusion criteria applying only to controls included a personal or familial history of mental illness; and the prescribed or recreational use of psychotropic drugs.

Patient recruitment Patients at the Royal Edinburgh Hospital (REH) were given an information sheet describing the study by their consulting psychiatrists, and asked if they would like to volunteer. It was made absolutely clear that they were under no obligation to do so, and may withdraw at any time. This information sheet outlined the purpose of the study, what it would entail, and described how they would be rewarded on the basis of their performance. If they showed an interest, the primary investigator met them in person to answer any questions and explain the risks associated with the magnetic resonance environment. Upon agreement, they were asked to sign a consent form, and an appointment for screening and behavioural assessment scheduled.

Control recruitment Controls were recruited from amongst the students and staff of both the REH and the University of Edinburgh via adverts placed on notice boards. Those expressing an interest were invited to the REH, to provide any further clarification, and to sign a consent form.

6.1.2 Behavioural procedure

6.1.2.1 Screening and clinical evaluation

Patients Given colour perception's crucial role in this experiment, the first stage was to assess their colour vision using an abridged version of the Farnsworth-Munsell 100-hue test [534, 411], with those showing no significant deficit continuing to the next stage. As the task required a degree of fine motor control, participants' motor skill were assessed using the standardised and objective nine-hole peg board test [535]. To establish if the participant was able to complete the experiment, they were introduced to a simplified version of the task, during

which the data glove and hand restraint were worn. If they demonstrated clear and correct responses to a small number of test trials, they went on to have their motor function assessed. This was necessary as both medication and schizophrenia itself can induce motor dysfunction, which needs to be quantified. This was done by Professor Cunningham-Owens using the Unified Parkinson's Disease Rating Scale (UPDRS) [536]. Their mental states were then assessed by Professor Johnstone during a structured interview according to the PANSS [537]. The clinical assessment stage took approximately one hour, and took place on the same day as scanning.

The computational model was modified to provide general predictions for the effects of altering (a) tonic dopamine levels and (b) D2 receptor efficacy - an analogue of medication level. In order to relate this to the patients' functional imaging data, a review of their medication was undertaken by Professor Johnstone, who summarised their antipsychotic histories over the past two years, based on their case note history. The summary included the mean dosage for the antipsychotic drug they had been taking over that period, or separate means if more than one drug was administered. These values were converted to chlorpromazine-equivalent doses, and summed for each patient.

Healthy participants Controls also completed a short simplified version of the task to ensure they were capable of performing it with ease. Controls' personal and familial psychiatric histories were determined by self-report. They also completed the colour vision and nine peg tests.

As the participants were deliberately selected to be high-functioning, it was deemed that the National Adult Reading Test (NART, [538]) would not provide a sufficient measure of intelligence, as the maximum score attainable in this test was 131. Instead the Wechsler Abbreviated Scale of Intelligence test (2-task version [539]) was used.

6.1.2.2 Stimulus set generation and pre-scan training

Participants completed the psychophysical calibration program to generate a customised series of colour cues to be used during the training and practice sessions. They were then accompanied to the Western General Hospital's MRI scanner. A radiographer confirmed that there was no reason why they should not be scanned. They then underwent training. This procedure was very similar to that used during the first study, in that participants wore the data

glove and hand restraint, and were allowed to learn the cue-response associations though repeated presentations of unambiguous colour stimuli. However, a key modification was made: the movements were analysed in greater detail, and the participant provided with more detailed feedback. For example, to try and extend a particular finger further, or to remind them of the correct cue-response associations if they made repeated mistakes. This procedure took no more than 15 minutes. After they reached an accuracy level of 90%, and reported feeling comfortable with the task, they were taken into the scanner.

6.1.3 Experimental design

The previous study significantly activated the pallidus both when considering the main effect of Selectivity, and when regressing the data by the computational model's trial-by-trial predictions. However, no main effect of Competition was found within the striatum. Potential problems in this area had been anticipated [407], which is why participants were informed that three of their trials would be selected at random and rewarded on a basis of £5 each. However, it seemed apparent that this did not provide a sufficient motivational aspect to the task. To improve this, it was decided that trial-by-trial feedback should be provided to the participant during the scanning session. It was hoped that this would help amplify the immediacy of the potentially positive/negative outcome, keep participants motivated and on-task, and so accentuate the BOLD response to Competition. This called for significant further development to allow both pre-scan glove calibration, and rapid trial-to-trial assessment of glove data within E-Prime (see section 3.4). The introduction of feedback was obviously a significant alteration to the original paradigm, and given the potential for confounding learning influences, was not a decision taken lightly. However, as participants would be well-trained beforehand, and efforts taken to ensure that feedback was as accurate as possible, it is believed that such confounds should be kept to a minimum.

6.1.4 Functional imaging

6.1.4.1 Data acquisition

The imaging protocol was essentially identical to that used during the first study. However, the original protocol called for 28 minutes of functional scanning, which would likely be an excessively long time for patients to endure. To establish whether this period could be shortened, the original control data was reanalysed using only the first three runs, to determine whether the

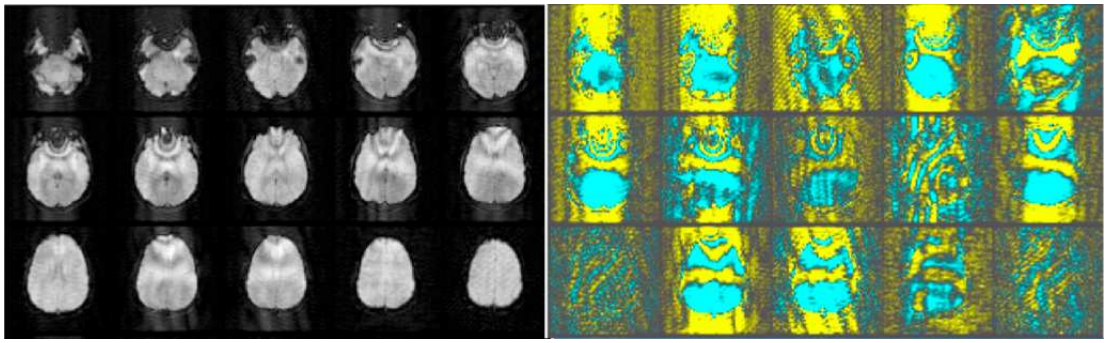


Figure 6.1: An example of a volume severely affected by interference. On the left is the raw data. On the right, the volume has been contrasted to the 5th volume in the time series, with the contrast enhanced using ArtRepair (http://spnl.stanford.edu/tools/ArtRepair/ArtRepair.htm)

previously-detected effects of interest remained significant. The effect of Selectivity remained significant within pallidus (MNI -18 -6 -6, p = 0.037 corrected within the pallidal ROI). Likewise, the effect of Competition within IFG also remained significant (MNI -34 8 0 and 34 28 0, p = 0.018 corrected for the whole-brain volume). It was therefore deemed that shortening the experimental acquisition from 28 minutes (comprising four runs) to 21 minutes (comprising three) was an acceptable trade-off between detection power and participants' comfort.

6.1.4.2 Data reconstruction and quality assurance

As the study progressed, data was reconstructed and checked for quality immediately after acquisition. The raw DICOM files were reconstructed into .img and .hdr format using SPM5's reconstruction functions. They then underwent the same quality assurance procedure as for the first study. Unfortunately, a severe artifact was present on one of the control's scans (figure 6.1).

Considerable QA work was undertaken by the staff at the Western General Hospital to attempt to isolate the cause. Repeated phantom scans both with and without the data glove and its associated hand restraint failed to elicit the artifact a second time. It was concluded that this was likely a one-off event, and scanning could proceed. This participant agreed to return for a repeat scan.

Unfortunately the artifact returned on a further two occasions, this time during patient scans. The experiment paused at this point, and the scanner was given a full service. Following artifact-free testing, scanning then recommenced, and thankfully the patients kindly returned

for full repeat scans. The full data set of twelve patients and twelve controls was successfully collected.

6.1.4.3 Preprocessing

SPM5 was used instead of SPM2 for the preprocessing and analysis of this data (`http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`). This was driven by the updated version's improved segmentation and warping techniques, which would offer more accurate normalisation. The neuropathological process of schizophrenia, and the medication taken chronically to treat it, are known modulators of cortical morphology [74]. Therefore comparisons between ill and healthy groups always entail extra nuisance variance at the stage of normalisation. As this study aimed to compare the functional activation between such groups, SPM5's improved normalisation techniques offered a useful means to reduce this source of error. The associated preprocessing and analysis scripts from the previous study were therefore rewritten for SPM5.

The initial steps were identical to those of the first study. Data were first slice-time corrected, and realigned to the mean. The participant's high-resolution T1 structural image was then coregistered to the mean EPI volume, and the parameters applied to the entire fMRI data set.

The key point of departure concerned normalisation. The T1 structural image was segmented using SPM5's improved technique, that incorporates spatial priors of tissue classification, nonlinear warping and bias correction into a more refined normalisation process. The 12-parameter affine transform was applied to the functional data to bring the participant's complete data set into MNI space, with the images written at a resolution of 3mm isotropic. The resliced images were finally smoothed at a resolution of 8mm FWHM, isotropically.

6.1.4.4 Analysis

Considering Selectivity and Competition These were the key factors of interest for evaluating the BG in the context of the modified computational model. The analysis proceeded in the same way as described for the first study.

The modulatory effects of medication The patients' mean medication dosage over the past two years was converted into chlorpromazine equivalents using accepted methods [540], and summed over different medications for each participant, if applicable. It was believed that taking a longer-term measure of medication would better capture its effects on synaptic and

neurochemical adaptation. This was was important for this study, as medication was not simply included as a confounding factor, but was the subject of specific hypotheses in itself.

Considering symptoms as an expression of selection dysfunction Finding *a priori*-predicted correlations between task-modulated brain activation and symptom ratings offers a powerful means to better understand the underlying neurobiology of schizophrenia. This is frequently done by considering, say, delusional symptoms and their relationship to association learning-related activation [541], or by considering a negative-to-positive continuum [542]. Here, as the emphasis is on a theoretical dysfunctional central selection mechanism, a more focused approach would be better suited. The PANSS items were considered in terms of the presumed inflexibility-versus-instability of the underlying selection mechanism. The following items were identified before scanning commenced.

Unstable Here five items were deemed to appropriately reflect an inherently unstable selection mechanism: one that either selected too freely, switched erratically, or failed to maintain an appropriate selection.

- P2: Conceptual disorganization. Described as a disorganized process of thinking characterized by disruption of goal-directed sequencing, e.g., circumstantially, tangentiality, loose associations non sequiturs, gross illogicality, or thought block. This could result from the central selection mechanism being unable to maintain an appropriate selection, with minor competing elements being allowed to overwhelm the supposedly dominant goal/action.
- P4: Excitement. Hyperactivity as reflected in accelerated motor behavior, heightened responsivity to stimuli hypervigilance, or excessive mood lability. These symptoms suggest a lowering of the threshold at which stimuli are attended to, or thoughts/actions executed.
- G9: Unusual thought content. Thinking characterized by strange, fantastic, or bizarre ideas, ranging from those which are remote or atypical to those which are distorted, illogical, and patently absurd. This again is indicative of lowered threshold for gating, with irrelevant ideas/thoughts being able to intrude into normal function.
- G11: Poor attention. Failure in focused alertness manifested by poor concentration, distractibility from internal and external stimuli, and difficulty in harnessing, sustaining, or shifting focus to new stimuli. distractibility was defined as an important metric of selection dysfunction within chapter 5, and is broadly analogous to the symptoms defined here.

- G14: Poor impulse control. Disordered regulation and control of action on inner urges resulting in sudden, unmodulated, arbitrary, misdirected discharge of tension and emotions without concern about consequences. This suggests a lowering of selection thresholds.

Rigid Seven items were chosen, intended to capture a mechanism that had difficulty in initiation selection, or remained perseverative once selection had occurred.

- N1: Blunted affect. Diminished emotional responsiveness as characterized by a reduction in facial expression, modulation of feelings, and communicative gestures. This indicates an excessively high selection threshold, that is preventing appropriate actions from occurring.
- N4: Passive/apathetic social withdrawal. Diminished interest and initiative in social interactions due to passivity, apathy, anergy, or avolition. This leads to reduced interpersonal involvement and neglect of activities of daily living. Again, this could be indicative of high thresholds preventing any selection.
- N6: Lack of spontaneity and flow of conversation. Reduction in the normal flow of communication associated with apathy, avolition, defensiveness, or cognitive deficit. This is manifested by diminished fluidity and productivity of the verbal-interactional process.
- N7: Stereotyped thinking. Decreased fluidity, spontaneity, and flexibility of thinking, as evidenced in rigid, repetitious, or barren thought content. A restricted repertoire of thoughts indicates perseveration, or a lack of flexibility within the selection mechanism.
- G7: Motor retardation. Reduction in motor activity as reflected in slowing or lessening of movements and speech, diminished responsiveness to stimuli, and reduced body tone. This suggests high selection threshold, and a difficulty in maintaining unimpeded gating once selection has occurred.
- G13: Disturbance of volition. Disturbance in the wilful initiation, sustenance, and control of one's thoughts, behavior, movements, and speech. A difficulty initiating appropriate thoughts/actions suggests excessively high selection thresholds.
- G15: Preoccupation. Absorption with internally generated thoughts and feelings and with autistic experiences to the detriment of reality orientation and adaptive behavior. Similar to N7: Stereotyped thinking: a perseverative thought process.

These items were condensed into one final metric of selection dysfunction: instability

$$instability = \frac{\sum(P2 + P4 + G9 + G11 + G14)}{5} - \frac{\sum(N1 + N4 + N6 + N7 + G7 + G13 + G15)}{7}$$

The instability measure was calculated for each patient, and entered into second-level regressions of the key contrasts of interest.

Computational modelling-directed analysis The first stage was to verify that the updated model provided a sufficient account for healthy control data. Based on the parameter fitting carried out on the data from the first study, it was decided that control's data was best represented with a tonic salience $c_{ton} = 0.1$, and D2 efficacy $\lambda_e = 0.2$. This also provided some room for manoeuvre in terms of simulating "overmedicated" (i.e. reduced D2 efficacy) and "unmedicated" (i.e. increased D2 efficacy) states. fMRI regressors were generated by modelling the BG performing each participant's experimental run, in the same way as before. These were then convolved with the HRF, entered into first-level design matrices, and estimated.

Patient data analysis The patient data was considered using a set of *a priori*-defined parameters representing various illness states:

- Healthy ($c_{ton} = 0.1$, $\lambda_e = 0.2$): normal cortical efficiency, normal D2 receptor efficacy
- Ill and unmedicated ($c_{ton} = 0.05$, $\lambda_e = 0.4$): cortical efficiency is reduced, and D2 efficacy is elevated
- Ill and appropriately medicated ($c_{ton} = 0.05$, $\lambda_e = 0.2$): the proposed primary disturbance of cortical efficiency remains unaltered, whereas D2 efficacy has been brought back to normal levels
- Ill and overmedicated ($c_{ton} = 0.05$, $\lambda_e = 0.1$): Again, the proposed primary pathology of diminished cortical efficiency remains present, though D2 efficacy has now been overcorrected to a level below normal.

Initially, the first-level t-tests of the "Healthy" model regressors were examined in a second-level t-test. In order to better represent the heterogeneity of the group, medication and the "Instability" symptom measure were included as covariates. This allowed for the basic fit the model provided for the patient data to be assessed, as well as exploring how that fit varied according to the inclusion of these key clinical parameters.

	Control mean (std)	Patient mean (std)	t	p (2-tailed)
WASI IQ	130 (6.7)	127 (7.5)	0.768	0.452
Age	32 (6.8)	34 (10.3)	0.677	0.505

Table 6.1: IQ and age for the patient and control groups.

In order to determine which of the four parameter sets provided an improved fit on the basic ”Healthy” implementation, first-level t-tests from all four models were entered into a single second-level ANOVA, again with medication and Instability included as covariates. T-test contrasts were performed for each each of the ill model variants were compared to Healthy (Ill and unmedicated > Healthy, Ill and medicated > Healthy, Ill and overmedicated > Healthy).

6.2 Results

6.2.1 Demographics and behaviour

Groups were matched for WASI IQ and age (table 6.1), and gender (controls 4F:8M, patients 5F:7M). To determine whether behaviour varied between groups, response time was assessed within a repeated measures ANOVA with Competition and Selectivity as within-subject factors, and diagnosis as a between group factor showed significant effects for Competition ($F(1.383) = 8.788$, $p = 0.003$) and Selectivity ($F(1) = 14.387$, $p = 0.001$), but no significance for group ($F(1) = 1.670$, $p = 0.210$) or its respective interactions with Competition ($F(1.383) = 0.626$, $p = 0.485$) or Selectivity ($F(1) = 0.172$, $p = 0.682$), indicating that groups were matched for performance. This removes one potential confounding factor from our interpretation of the data. There was no evidence of Parkinsonism in any of the patients, with ten patients having a UPDRS score of zero, and two having a score of one.

6.2.2 Selectivity and Competition

6.2.2.1 Basal ganglia responses

The models suggested that reducing cortical efficiency (hypothesised here to be the primary dysfunction) would result in a general reduction in both striatal and pallidal activation, whereas decreasing D2 efficacy (analogous to the effects of medication) would act to increase activity in both nuclei. Given the medicated status of the patient sample, this would suggest that

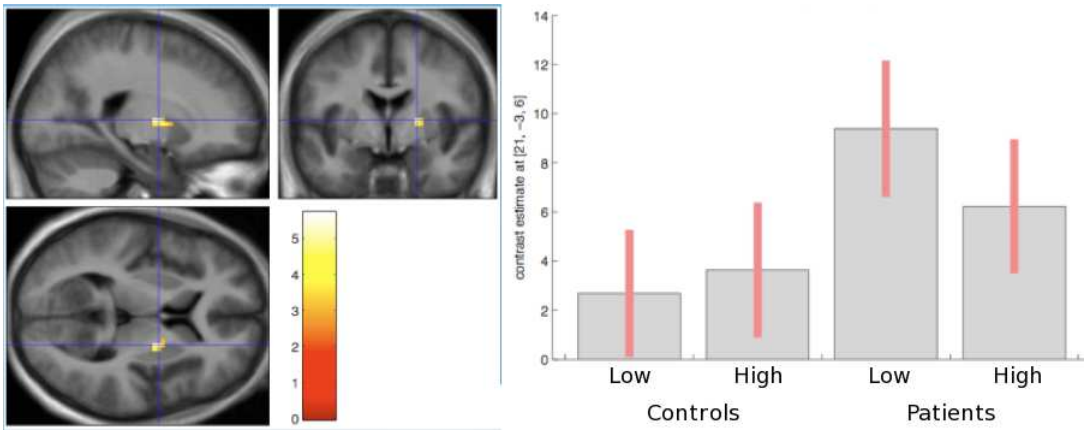


Figure 6.2: The effect of Selectivity within right pallidus. The controls show no significant effect, though the effect is in the expected direction of High > Low. Patients on the other hand show the opposite effect to what is normally expected of a disinhibitory selection mechanism. The peak voxel (MNI coordinates 21 -3 6) is for the contrast Low > High in patients, t = 4.44, p = 0.001 correcting within the pallidal ROI).

a general ”cancelling out” of these effects would be expected when comparing groups, and indeed there is no main effect of group in either nuclei, with or without the medication/Instability covariates ($p > 0.429$ corrected).

The modified computational model made quite specific predictions concerning Selectivity and Competition, and their effects on striatal/pallidal activation (see section 5.5.4).

Pallidal responses The strongest result from the previous study was that of Selectivity within pallidus. The contrast of High > Low Selectivity did not reveal anything significant in either patients or controls, perhaps due to diminished power (group size was reduced from 16 to 12, and scan time from 28 to 21 minutes). However, looking at the inverse contrast of Low > High Selectivity, a pallidal region was found to be significantly active in patients, at MNI coordinates 21 -3 6, t = 4.44, p = 0.001 (FWE corrected within bilateral pallidus, figure 6.2). This result is quite surprising, and must be further clarified by determining what effects medication has on this relationship. It is made more interesting by the observation that Low > High Selectivity in patients does not show any significant activation at the whole-brain level ($p < 0.005$ uncorrected), making it unclear what might be driving this difference.

The relationship between these key covariates and the patients' abnormally inverted Selectivity response was investigated further. A 3mm sphere was centred on this cluster, and the parameter estimates extracted for each cell of the 2 x 3 Selectivity x Competition factorial design.

These were entered into a repeated measures GLM within SPSS (www.ibm.com/software/uk/analytics/spss/), with medication and Instability then being entered as covariates, to assess the interactions between them. The analyses then performed were orthogonal to the original contrast that selected the region, therefore ”double-dipping” was not an issue. The Selectivity x Competition x Instability interaction was significant ($F(2, 18) = 7.116$, $p = 0.005$, whereas the respective interaction with medication was not ($F(2, 18) = 0.702$, $p = 0.509$)). Post hoc contrasts showed that this was being driven by the difference between High and Low Competition ($F(1,11) = 16.908$, $p = 0.003$). This interaction and its relationship with Instability is shown in figure 6.3. The model predicts that these values should be negative, as simulated pallidal activity decreased with increasing Competition, and an interaction with Selectivity was not expected. However, figure 6.3 shows that there is a marked dichotomy between Low and High Selectivity. For Low Selectivity, patients with low Instability show the ”correct” pattern of activation, whereas for High Selectivity, this is only true of those with high Instability. None of the parameter sets used during the modelling phase in chapter 5 produced an output that matched this pattern of activation - this is clearly an area where further work will need to take place to provide a better explanation. Modelled work suggested that increasing cortical inefficiency - associated with high instability - would act to diminish pallidal activation across the board, though no interaction with Selectivity was anticipated. This shifting dichotomy determined by the Instability measure is suggestive of an ”inverted-U” shaped response, with a more ideal activation being found at the centre of the distribution, i.e. when Instability = 0. This would be expected if Instability is indeed correlated with the degree of selection dysfunction within the BG.

Striatal responses To better understand what might be causing this Selectivity-determined interaction within pallidus in patients, the upstream influence of striatum was examined. The region of peak pallidal activation was identified using an effects of interest F test within the striatal ROI, and then extracting parameter estimates for each cell of the 2 x 3 factorial design from a 3mm sphere centred on these coordinates. A voxel centred at MNI -27 0 3 was found to be maximally activated by the experimental conditions (figure 6.4), and this also showed an unexpected interaction between Selectivity and Competition. During Low Selectivity, this striatal region appeared to increase with Competition as expected, although this was not the case when the demand for Selectivity was increased. At the point where the demands placed upon the system are maximal - i.e. when both Selectivity and Competition are at their highest,

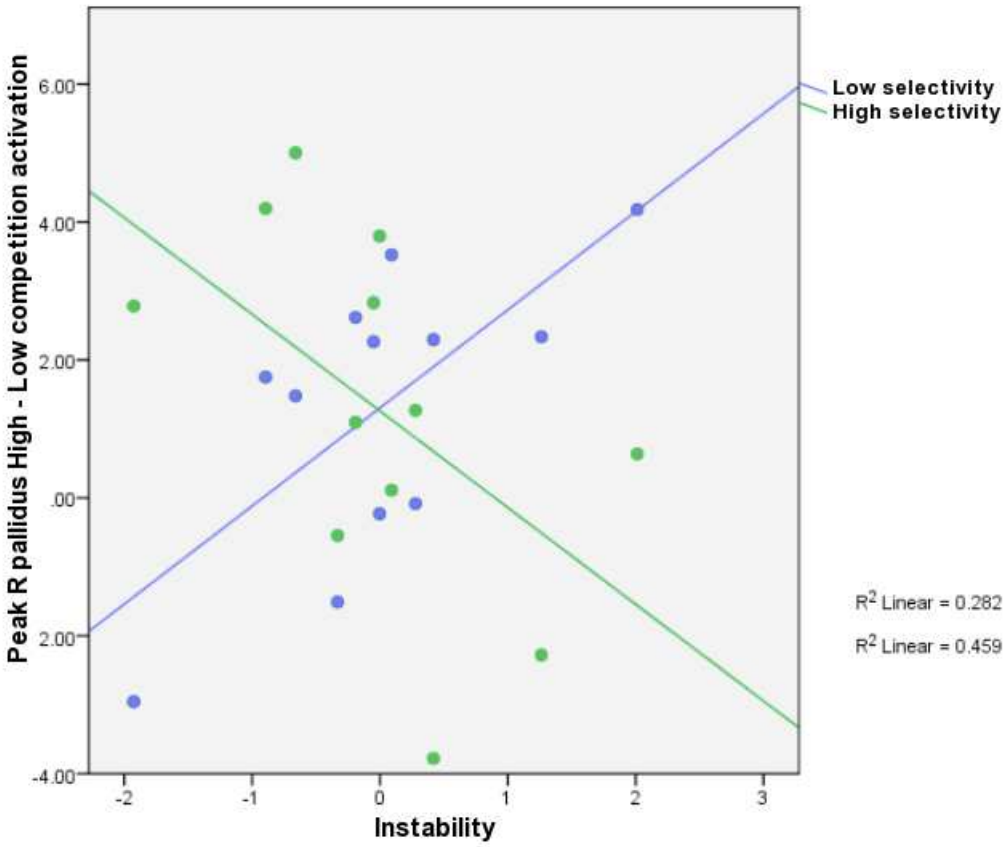


Figure 6.3: The Selectivity x Competition interaction within right pallidus in patients, and its relationship with Instability. The modified model predicts that the High - Low Competition pallidal response should always be negative (i.e., Low > High). However, the fMRI data suggests that for patients, this is only true in certain circumstances: those with low Instability demonstrate this when Selectivity demands are low, whereas high Instability patients only show this when Selectivity demands are increased.

there is a marked drop-off of striatal activation in patients.

A repeated measures GLM did not find significant effects of Selectivity, Competition or the interaction ($p > 0.148$), although the Selectivity x Competition x Instability interaction did show a strong trend ($F(2, 18) = 3.409$, $p = 0.056$): the degree to which a patient's striatal activation "failed" and entered an inverted state increased with Instability. The modelling work in chapter 5 (see section 5.5.4.1) predicted this inversion of striatum's response to Competition when cortical efficiency (reduced background activity and SNR) was pathologically reduced - above and beyond the effects of medication. It also suggested that Instability would increase with cortical inefficiency (due to the induced compensatory upregulation of dopamine). We therefore have a model-derived triangulation of three factors: an inverted striatal response to Competition and increased Instability are both predicted to indicate cortical inefficiency.

Instability seemed to provide a partial account for the inverted Selectivity response in general, with the Selectivity x Instability interaction being $F(1, 9) = 4.186$, $p = 0.071$. This finding was also true of the maximally task-activated region of right striatum (MNI 24 0 6), with the Selectivity x Competition x Instability interaction achieving a trend here also ($F(2, 18) = 2.948$, $p = 0.078$).

6.2.2.2 Cortical responses

Expanding beyond the BG, these analyses were conducted and corrected at the whole-brain level using cluster-level statistics, initially thresholded at $p < 0.005$ uncorrected unless otherwise stated.

Neither the controls nor patient groups showed any significant whole-brain effects of Selectivity.

The first study involving healthy controls highlighted the particular involvement of IFG in representing increasing Competition. The High > Low Competition response seen in controls is replicated in this new data set. In addition, the inclusion of trial-by-trial feedback has also elicited activation of anterior cingulate (table 6.2). The same contrast in patients shows no significant activation.

However, when we examine the contrasts estimates for the areas that are significant in controls, an interesting pattern is revealed. In accordance with theories that NMDA hypofunction will lead to elevated cortical activity overall (due to interneuron-mediated disinhibition [543]), and a general reduction in SNR, we find that activity is elevated and non-responsive to Competition (figure 6.5). In all cases, these contrast estimates were determined by comparing each

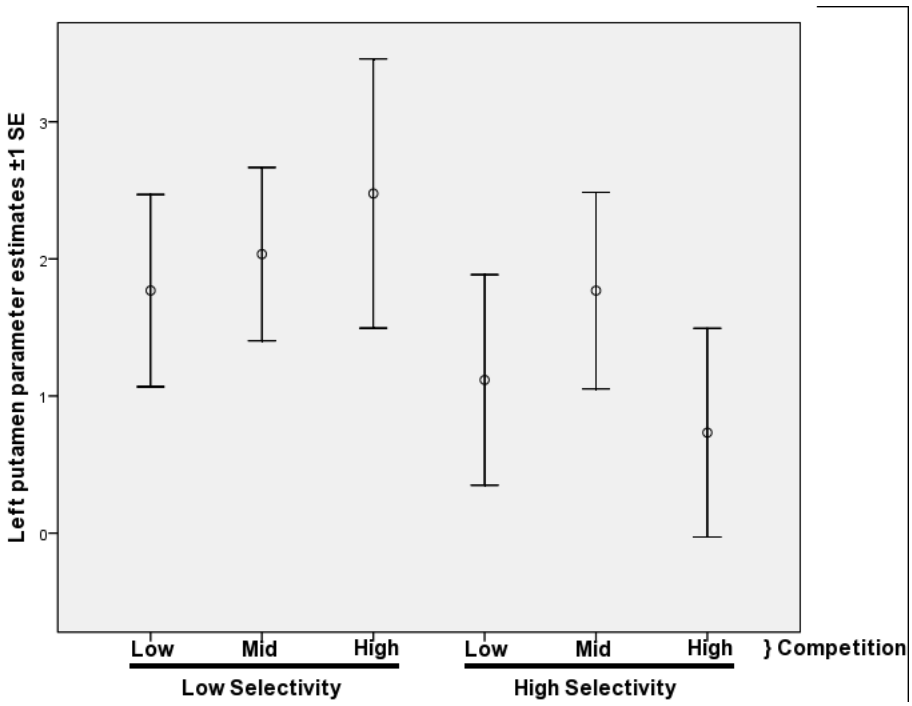


Figure 6.4: Parameter estimates extracted from peak-activated putamen of patients for each level of Selectivity and Competition. There were no significant main effects of Selectivity ($p = 0.148$), Competition ($p = 0.815$) or a Selectivity x Competition interaction ($p = 0.559$), although there was a strong trend for the Selectivity x Competition interaction being modulated by Instability ($p = 0.056$).

	MNI coords	Z	p	cluster size	cluster-corrected p
Left inferior frontal gyrus	-33 18 -3	4.79	0.000	1642	0.000
Left occipital cortex	-27 -87 15	4.59	0.000	357	0.001
Anterior cingulate cortex	-3 18 48	4.50	0.000	401	0.001
Right parieto-occipital cortex	36 -57 54	4.29	0.000	374	0.001
Right inferior frontal gyrus	45 24 15	4.05	0.000	386	0.001

Table 6.2: Within control subjects, areas that were significant for High > Low Competition (thresholded at 0.005 uncorrected).

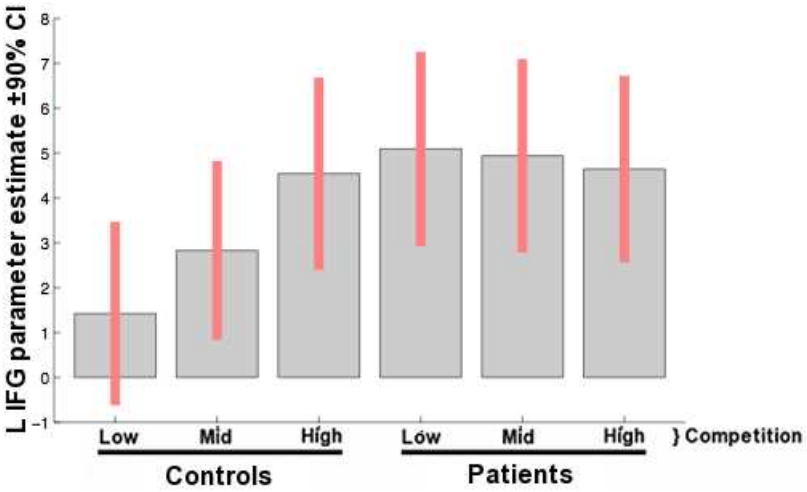


Figure 6.5: Contrast estimates extracted for Competition for controls and patients, from left inferior frontal gyrus (MNI -27 15 6).

experimental condition to the low-level visual control (checkerboard) trials.

To better understand what might be causing this deviation in the usual cortical encoding of Competition - and how this may in turn alter BG activation - the contrast of High - Low Competition was regressed by medication and Instability for the patient data in isolation: two results emerged.

First, the Competition response was found to be positively correlated with Instability in the right substantia nigra (figure 6.6, MNI 9 -15 -18, p = 0.025 cluster-corrected), one of the brain's key dopaminergic regions. Medication was not found to exert a notable effect here (Z = 1.69, P = 0.045 uncorrected). The model simulations suggested that both increased cortical inefficiency and increasing Competition lead to a disinhibition of dopamine release. Indeed, within controls there is a strong whole brain trend for a High > Low Competition response in left midbrain (MNI

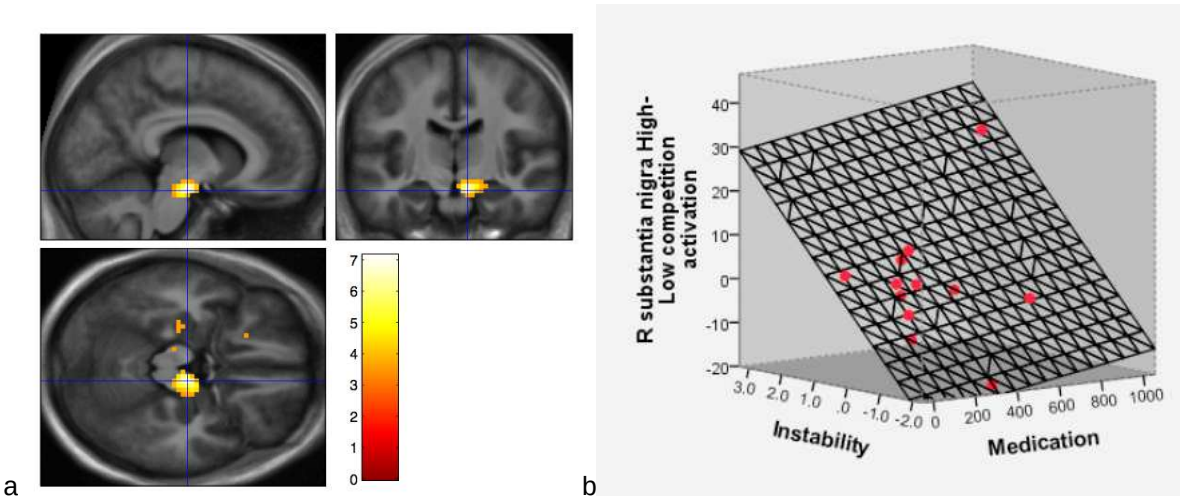


Figure 6.6: A modulation of the patients' midbrain response to Competition by Instability. (a) The Competition X Instability interaction, thresholded at p < 0.001 uncorrected. The right substantia nigra is significantly activated at the whole brain level (MNI 9 -15 -18, Z = 5.30, k_E= 79, voxel-level p = 0.003 corrected, cluster-level p = 0.025 corrected. (b) The relationships between medication, Instability and midbrain activation. There is a positive correlation between High-Low midbrain activation, and Instability. Medication has no notable influence (P = 0.045 uncorrected).

-12 -15 -9, voxel-level p = 0.054 corrected). Whether by tonic or demand-related disinhibition, this would act to facilitate looser, less strict selection, which would theoretically underpin the symptom set defining Instability. The data suggests that this Competition-dependent upregulation of dopaminergic activity is more pronounced for those patients with higher Instability - i.e., those with theoretically higher cortical inefficiency, inducing an exaggerated compensatory response.

Second, a negative correlation was found between Instability and Competition within right dorsolateral prefrontal cortex (figure 6.7, MNI 21 33 45, p = 0.013 cluster-corrected). Again, medication was not found to exert an influence here (P = 0.011 uncorrected). This is an especially notable region of prefrontal cortex within the perceptual decision-making literature: it has been found to encode the value of the selected decision across a variety of tasks and modalities [325, 330, 335]. In this context, it would therefore be expected to show a Low > High Competition response, as the correctly-selected stimulus in a Low Competition trial will by definition be dominated by one colour. Controls do show a Low > High Competition at these coordinates, albeit at a weak trend level of p = 0.104 corrected. The patients with low instability are therefore more concordant with the expected response. The model suggests that Low Instability

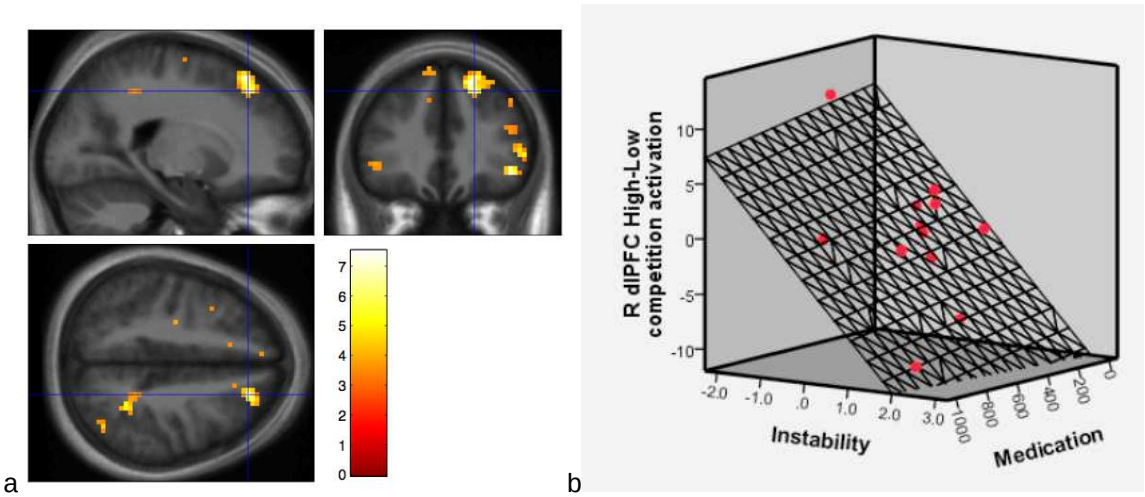


Figure 6.7: A modulation of the patients' prefrontal response to Competition by Instability. (a) The Competition X Instability interaction, thresholded at p < 0.005 uncorrected. Right dorsal prefrontal cortex is significantly activated at the whole brain level (MNI 21 33 45, Z = 4.13, k_E= 122, cluster-level p = 0.013 corrected. (b) The relationships between medication, Instability and right prefrontal activation. There is a negative correlation between High-Low midbrain activation, and Instability. Medication has no notable influence (P = 0.011 uncorrected).

patients will have relatively preserved cortical efficiency, and so remain able to maintain an acceptable cortical SNR. This tallies with the relationship seen between decreasing Instability and improved decision encoding within right dorsolateral prefrontal cortex.

Taking these two results together, we see that Instability - this proposed symptom-derived correlate of selection performance and a modelled indicator of cortical inefficiency - is associated with both an exaggerated dopaminergic compensation, and a diminished SNR in decision-encoding prefrontal cortex. This is above and beyond any corrective effects imposed by medication.

6.2.3 Computational model-directed analysis

During this phase of the analysis, the modified computational model was assessed to determine the goodness-of-fit it provided for the new set of healthy controls, as well as its ability to fit patient data based on a set of proposed pathological "states".

6.2.3.1 Healthy control verification

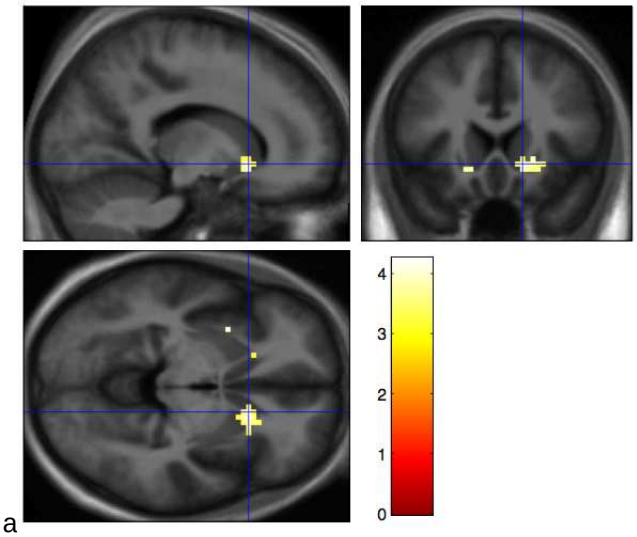
Using the ”healthy” parameter set of $c_{ton} = 0.1$, and $\lambda_e = 0.2$, separate regressors were generated for each participant, for both the striatum and pallidus. The account that this model was able to provide for striatal activation did achieve a higher level of significance than that of the old model for the original healthy control data: a region of ventral anterior striatum reached a ROI-corrected level of p = 0.126, MNI 15 18 -6 (figure 6.8). The original model was only able to reach a corrected level of P = 0.723 within left putamen. Whereas the old model's activation peak was in central putamen, the new model sees this shift to a more anterior-ventral aspect in both the original and new control data sets, indicating a degree of generalisability. However, with the original data set, the best account was seen in left striatum, whereas for this second verification data set the peak has moved onto the right side. The implications of this are unclear, as it is not clear what role laterality will be playing in these more anterior, associative areas.

Unfortunately for pallidus, a different picture emerged. The new model's pallidal regressors did not provide a sufficient account for activation in these areas within the new control data set - the uncorrected threshold had to be dropped to p < 0.21 before any activation appeared. Figure 6.9 displays the controls' parameter estimates for each experimental condition extracted from maximally-activated pallidus. This shows that the pallidal nuclei of this second set of controls are markedly different in their activation compared to both the original data, and model simulations.

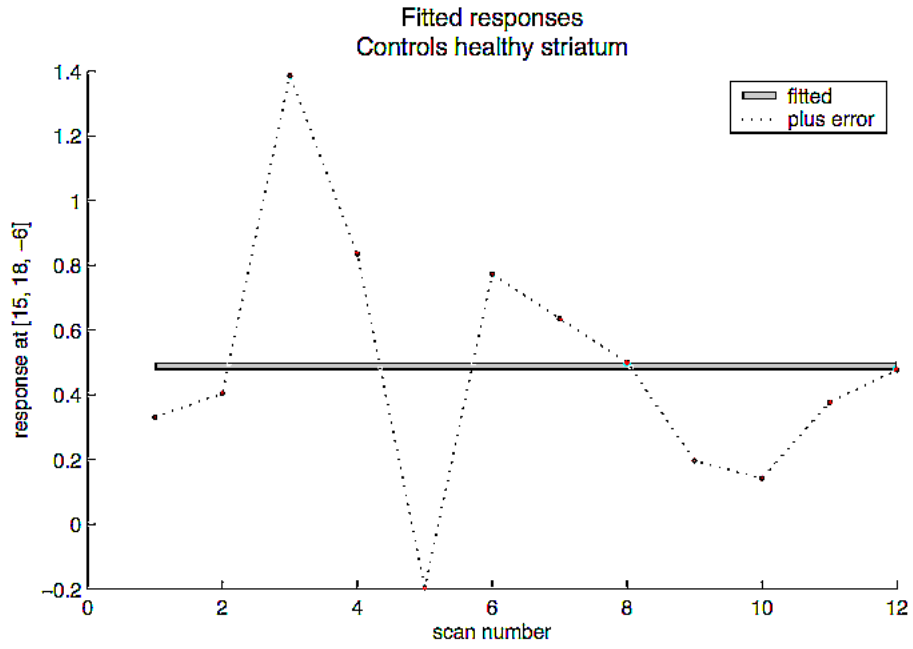
This is a step backwards from the original model - this may be due to the inclusion of the inhibitory pallidostriatal pathway, which in turn targets those striatal neurons that reengage with GPe. It may be that the balance of this recurrence needs to be adjusted to better represent the interplay between these two nuclei. An alternative explanation could be that the inclusion of trial-by-trial feedback has altered some of the fundamental processes underlying this task

6.2.3.2 Assessing the models' accounts for patient data

Striatum Within striatum, the suitability of the Healthy model was first assessed in a design matrix that included medication dosage and symptom-derived Instability as covariates. A simple t-test showed that the Healthy regressor showed a strong trend for providing an account



a



b:

Figure 6.8: Verifying the new model's capacity to account for striatal activation in a second set of healthy controls. Regressors were generated using the Healthy parameter set ($c_{ton} = 0.1$, $\lambda_e = 0.2$). (A) The new model provides some account for right anterior ventral striatum. Data is displayed at an uncorrected threshold of P < 0.005 within striatum. The crosshaired cluster is located at MNI 15 18 -6, p = 0.126 corrected. (b) Extracted parameter estimates for each participant. Only participant 5 shows a negative fit for the healthy striatal regressor, although this person is not an outlier on any of the behavioural measures.

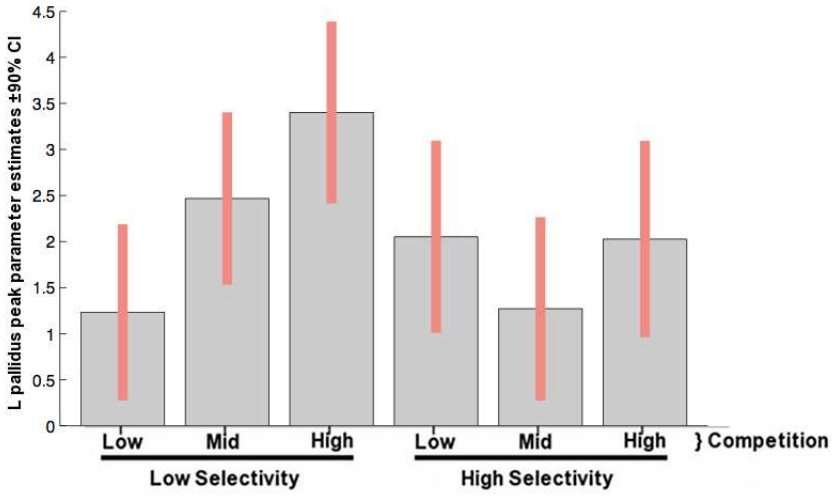


Figure 6.9: Parameter estimates extracted from maximally-activated pallidus in controls (MNI -15 6 -3). Low Selectivity is generally more active than High Selectivity, which contradicts both modelled predictions, and the results from the previous control data set. The relationship with Competition - which the previous study/modelling work suggested should be decreasing - is also discordant.

for right ventrolateral putamen activation (MNI 36 0 -3, p = 0.062 corrected within the striatal ROI, figure 6.10a).

To explore how this account varied under the influence of medication dosage and Instability, the data was estimated using the Healthy model regressors again, this time without the covariates, and parameter estimates extracted from a 3mm diameter sphere centred on the putamen activation peak. These were then correlated with the covariates in SPSS, and a 3D scatter plot produced to demonstrate the relationships (figure 6.10b). Both an increasing medication dosage ($F(1, 11) = 19.447$. $p = 0.002$) and an increasing Instability score ($F(1, 11) = 12.767$, $p = 0.006$) strongly improved the fit provided by the Healthy model for this striatal region, though medication and Instability were not themselves significantly correlated (Pearson correlation = 0.371, $p = 0.235$).

Simulations of the modified computational model suggested that increased cortical inefficiency (decreased tonic salience) led to an increasing instability, due to an upregulation of tonic dopamine levels. This compensatory mechanism would help to maintain useful selection behaviour in a system that would otherwise become increasingly rigid. This prediction is therefore supported by the observation that as this ”normalising” process becomes more evident (increasing instability), the Healthy model’s fit for striatal activation improves.

However, this work supposes that the compensation mechanism and its associated instability

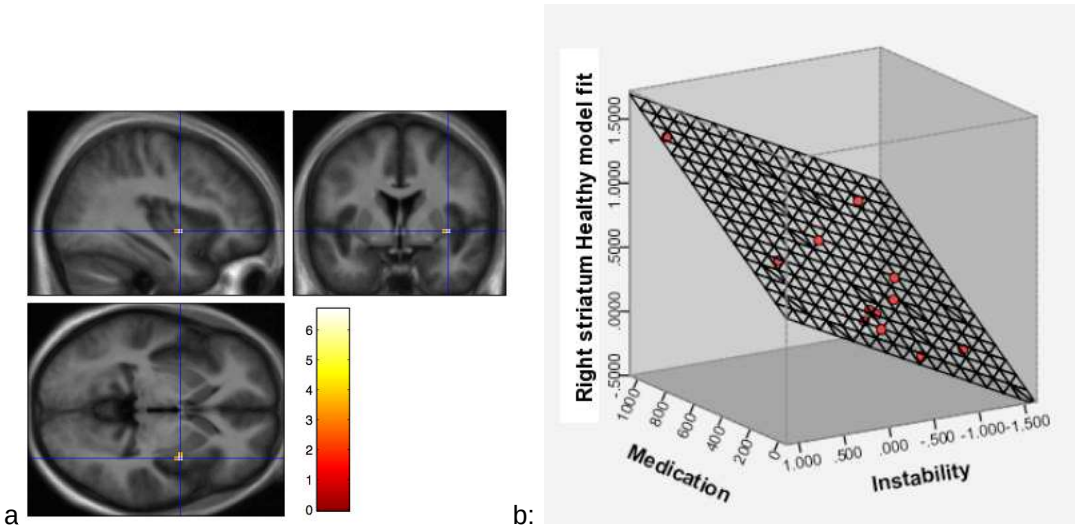


Figure 6.10: The account the ”Healthy” model provided for patients’ striatal activation, covarying for medication dosage and Instability. (a) The group t-test of the striatal Healthy regressor was thresholded at $p < 0.005$. A strong trend was found within right ventrolateral putamen (MNI 36 0 -3, $p = 0.062$ corrected within the striatal ROI). (b) A 3D scatter plot demonstrating the significant relationships between this striatal cluster’s Healthy model fit and both medication dosage ($F(1, 11) = 19.447$, $p = 0.002$) and Instability ($F(1, 11) = 12.767$, $p = 0.006$).

overwhelms the profile of the illness, which then must be kept in check with medication, which acts as a second ”normalising” influence. The scatter plot demonstrates that an increased medication dosage does indeed allow the Healthy model to provide a better fit for the data, supporting this hypothesis.

However, the *a priori* belief was that the Healthy model regressors should not provide an entirely sufficient account for BG activation within patients with schizophrenia, and that the pathological models would prove better. To evaluate this, parameter estimate maps for the different modelled pathological states - Ill-unmedicated, Ill-medicated, Ill-overmedicated - were entered into an ANOVA, together with the medication and Instability covariates. T contrasts of each of the pathological models versus Healthy were then examined:

- Ill-unmedicated > Healthy: this revealed no areas where its account surpassed that of the Healthy model. This is to be expected, as all but one of the patients was taking antipsychotic medication at the time of the scan.
- Ill-medicated > Healthy: an area of weak activation was noted in the right nucleus accumbens (MNI 3 9 -6, $p = 0.19$ corrected).

- Ill-overmedicated > Healthy: A similar region of right nucleus accumbens was activated, this time achieving a trend level of significance (MNI 6 18 -3, p = 0.099 corrected, figure 6.11). This suggests that the Ill-overmedicated model may provide a better account for caudate activation compared to the Healthy model, though the small sample size means this test is somewhat underpowered.

These results consist of a number of interesting findings: first, that the Healthy model can provide an account for the hand region of putamen in patients with schizophrenia. This implies that the neurobiology of schizophrenia does not involve an alteration in the more fundamental, motor-oriented functions of BG. Second: that the medicated versions of the Ill models are able to provide a better account for activation within more anterior/cognitively-associated striatum than the Healthy model. In fact, the more exaggerated the modelled effect of medication, the closer the resemblance.

Given that these models were differentiated by their degree of medication, it was expected that the improvement in fit over the Healthy model would be modulated by the participants' received dosage. This appeared to be the case, as the Ill-medicated > Healthy region showed a trend for a Model x medication interaction ($F(3, 9) = 2.614$, $p = 0.072$), whereas Instability was having no interacting effect ($F(3, 9) = 1.614$, $p = 0.209$). This suggested that the model was able to approximate the effects of diminished D2 receptor efficacy at a broad, nucleus-scale level.

Having found (a) an area of motor striatum that was sufficiently accounted for by the Healthy model, and (b) more anterior regions for which the pathological medicated and overmedicated models provided an improved account over Healthy in a medication-dependent manner, it now remained to see whether the fit provided by the models had an interaction with Instability, that would theoretically be indicative of the underlying disease process (greater cortical inefficiency leading to more pronounced upregulation of tonic dopamine). This was addressed using a second-level factorial design in SPM, with a single 4-level factor of Model, and the two covariates of medication and Instability, for which the interaction with Model was explicitly represented. This showed a trend for a Model x Instability interaction in the left nucleus accumbens (MNI -3 18 -6, $p = 0.077$ corrected within the striatal ROI, figure 6.12a). Using the same methods as before, parameter estimates for each of the models were then extracted from a 3mm sphere centred on these coordinates, with no covariates, and entered into a repeated measures GLM in SPSS to explore the relationships with medication and Instability. The Instability interaction was most pronounced for the contrast of Ill-medicated > Healthy, where $F(1, 11) = 30.427$, $p = 0.000$ (figure 6.12b), whereas no interactions with medication were found (p

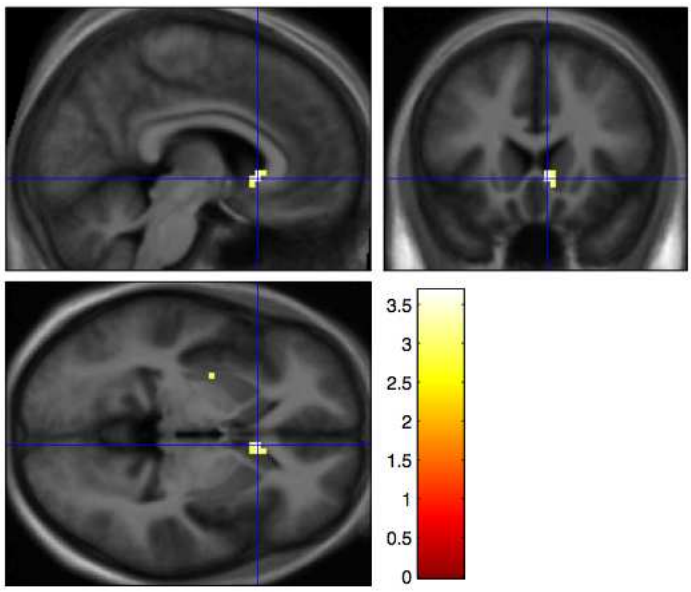


Figure 6.11: Ill-Overmedicated > Healthy model account contrast for patient data (covarying for medication and Instability), thresholded at $p < 0.005$. A region of right nucleus accumbens was found to be better approximated by the Ill-overmedicated modelled pathological state (MNI 6 18 -3, $p = 0.099$ corrected within the striatal ROI).

> 0.14).

Pallidus No region of pallidus was found to be accounted for by the Healthy model parameters estimates, covarying for medication and Instability. Similarly, when comparing the pathological versus Healthy models, no pallidal regions were accounted for. There is therefore no evidence that the new model can provide an informative account for the pallidal activation of patients with schizophrenia.

6.3 Conclusions

So far, the original Gurney BG model was used to derive specific predictions of activation for a perceptual decision-making task. These were tested during an fMRI study involving healthy volunteers, and subsequently used to drive changes to the computational model. Further predictions were then derived - this time concerning the hypothesised neuropathology of schizophrenia - and tested during this second fMRI study, involving both patients and controls. This iterative process has highlighted a number of interesting findings concerning cortical in-

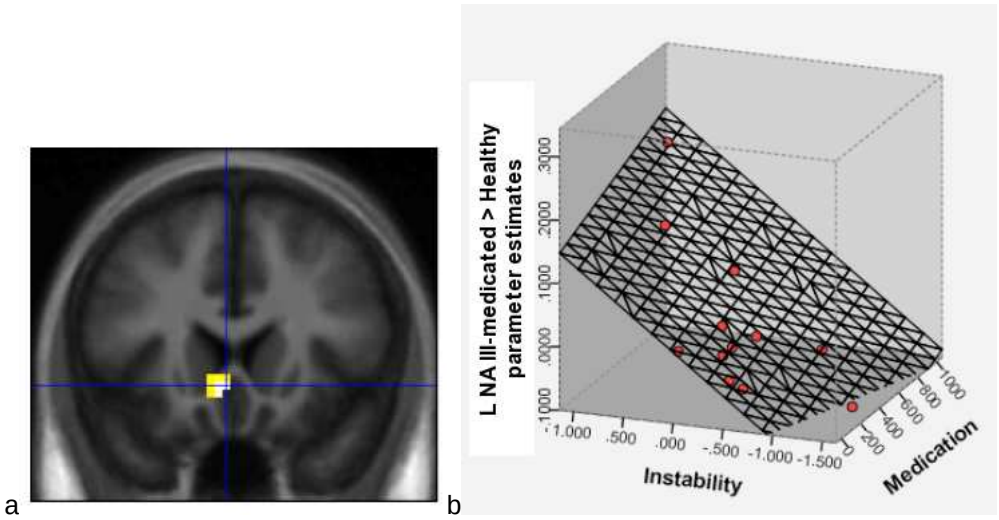


Figure 6.12: (a) Left nucleus accumbens (MNI -3 18 -6) showing a Model x Instability interaction (p = 0.077 corrected within the striatal ROI). (b) A scatter plot showing how medication and Instability affect the degree to which the Ill-medicated model is able to provide a better account for this region than the Healthy model. Instability exerted a profound influence over this fit ($F(1, 11) = 30.427$, $p = 0.000$), medication had no significant effect ($F(1, 11) = 2.683$, $p = 0.14$).

efficiency (a hypothesised consequence of primary NMDA receptor dysfunction within cortex, assessed using the symptom-derived correlate of Instability), its proposed consequence of compensatory dopamine disinhibition, and the corrective effects of antipsychotic medication.

6.3.1 There was no main effect of group within the basal ganglia

The updated model suggested that a drop in cortical efficiency would lead to a compensatory upregulation tonic dopaminergic activity, effectively acting to increase ”gain” within the striatum and so return cortico-BG function to normal. However, simulations suggested that this could lead to an instability in selection performance, with selection being initiated and switched with lower saliences, or less narrow Competitions. Here, simulated antipsychotic medication acted to push this overcompensation back the other way, again to normalise function. This combination of factors therefore made it unsurprising that no main effect of group was detected within the BG. A more subtle approach was required, involving the specific interactions between Selectivity, Competition and the key covariates of Instability and medication dosage.

6.3.2 The effects of Selectivity within pallidus were not replicated in controls

One of the key findings from the first study was that of pallidus' response to Selectivity. Unfortunately, this was not significantly replicated in the second group of controls, although extracted pallidal activity did suggest a High > Low Selectivity response. This could be due to several differences between this study and the last: the number of subjects is down from 16 to 12, and the scanning time itself was reduced from 28 to 21 minutes (to make the task more tolerable for patients). Both entail significant losses of power. Another key difference was the inclusion of trial-by-trial feedback, though it is unclear how this could interact with Selectivity to abolish the expected effect.

Given that the pallidus does not seem to be responding in the same way as before, it is no surprise that regressors extracted from the updated computational model are unable to provide a good account for its activation.

6.3.3 Patients' abnormal basal ganglia responses varied with symptoms suggestive of selection dysfunction

Patients showed a pronounced inversion of the Selectivity response within pallidus. By taking a broader view and incorporating the effects of Competition and Instability, a striking dichotomy was found: patients whose symptoms were suggestive of a somewhat rigid selection mechanism (i.e. with lower levels of dopamine disinhibition, or higher D2 blockade) showed "normal" pallidal responses only when Selectivity was Low, whereas patients with a more loose, unstable mechanism (i.e. with excessive dopamine disinhibition, or inadequate D2 blockade) were only "normal" when Selectivity was High. This shift in the point of selection breakdown was suggestive of an "inverted-U" response, and could be due to a complex interaction between the attempted corrections of both dopamine upregulation and antipsychotic medication. However, this is somewhat speculative. In order to gain a better understanding, upstream striatum was considered in the same way, and here the picture was a little clearer. Hypothesis C predicted that the hypothetical core pathology of schizophrenia - a reduction in cortical efficiency - would result in an inversion of striatum's response to Competition. This was indeed found to be the case when the selection mechanism was under increased pressure to suppress competing alternatives: moreover, this effect became more pronounced as a patient's Instability increased. This provided further support to the idea that cortical efficiency underwrote altered selection performance, as Instability was also predicted by hypothesis B to increase with cortical inef-

ficiency, due to a greater disinhibition of tonic dopamine release. How this in turn interacts with pallidus to produce the complex response observed there is unclear - especially given the recurrent relationship between these two nuclei. However, the fact that pallidal dysfunction is also dependent on Instability presents a degree of consistency.

6.3.4 Patients with theoretically-diminished cortical efficiency showed reduced SNR in decision-related cortex

Controls displayed a strong parametric response with respect to Competition in IFG. Although insignificant, patients' parameter estimates did show a complete flattening of this response, with its activation remaining tonically high. It has been proposed that this disinhibitory loss of SNR could be an expected consequence of NMDA receptor hypofunction, if acting via cortical interneurons [155]. In this work, such proposed pathophysiology was defined as "cortical inefficiency", which subsequent modelling simulations showed could lead to a less stable selection mechanism by virtue of disinhibited subcortical dopamine release. Therefore, this abnormal response to Competition had to be examined in the context of Instability. When the Competition x Instability interaction was assessed within patients, a strong effect was seen within right middle frontal gyrus - the very area that has been repeatedly implicated as crucial to decision-making in numerous perceptual paradigms [325, 330, 335]. Patients with high Instability - i.e. with hypothetically greater cortical inefficiency - showed a loss of the usual response to Competition in this area, which normally increased with the value of the selected response (i.e. decreased with Competition). This provided further support to hypothesis B.

6.3.5 Patients expected to have a greater disinhibition of dopamine activity showed an exaggerated midbrain response to Competition

Hypothesis A predicted that cortical efficiency led to selection Instability, via an overcompensatory disinhibition of tonic dopamine release. Simulations also suggested that dopamine levels would rise in response to increasing Competition - in both cases, the general action is to facilitate selection in arduous circumstances. It was therefore expected that unstable patients would show increased dopaminergic activity. *A priori*, it was not anticipated that these subtle tonic effects would be observed within midbrain, therefore no ROIs were applied. However, the Competition x Instability interaction was strong enough to be significantly detected at the whole-brain level, revealing that patients who hypothetically had the greatest cortical ineffi-

ciency (and so greater dopamine levels) as inferred from their symptoms, showed the greatest exaggeration of the expected dopamine response.

6.3.6 Instability appears to be a useful correlate of selection dysfunction

The three key findings in patients - that instability is related to BG selection failure, reduced cortical SNR and exaggerated dopaminergic midbrain responses - lend good support to this symptom-derived metric. Traditionally, symptoms were considered as ”positive” or ”negative”, with a third ”disorganised” factor emerging later, and now 5-factor models are frequently described. However, strong neurobiological correlates for these factors have not been forthcoming. More fruitful approaches have focused on the neurobiological mechanisms underlying a particular symptom, for example delusions, which is seen as resulting from a dysfunction in associative learning, isolating correlates in striatal and midbrain responses [544, 247, 545]. In this work, a selection of PANSS symptoms were combined in a way that attempted to directly assess the hypothetical consequences of a dysfunctional central selection mechanism. This yielded a trilogy of results that tallied well with what was expected based on computational simulations of the key selection mechanism of interest: the BG.

6.3.7 Direct regression by simulated nucleus activity yielded insights into medication

The computational model underwent considerable modification following the first study, including the explicit modelling of dopaminergic midbrain, a pallidostriatal projection, and a population of striatal FSIs. The benefit of these changes were partly verified in the second independent group of controls, who showed that the model’s account had improved for striatum, but diminished for pallidus. The reasons for pallidal deviance have been mentioned earlier, and as the model also failed to the pallidal activation seen in patients, it is concluded that this is an area for improvement in further modelling work.

However, more interesting results were found in model-regressed striatum. The healthy model was found to provide a partial account for the more simple, motor-oriented regions of putamen in patients, providing that the key covariates of medication dosage and Instability were included in the regression. Indeed, it was apparent that both these covariates acted to improve this account in the direction one would expect from their presumed ”normalising” influence, that is,

they were both positively correlated with Healthy model fit (though not with each other).

The pathological simulated models used to regress the data were differentiated by their level of D2 receptor blockade. This appeared to be a key parameter, as the Ill-unmedicated version was found to be specifically worse than Healthy (to be expected, given the positive medication histories of all the patients), whereas the Ill-medicated and Ill-overmedicated models provided an incrementally improved fit for striatal activation. Interestingly, this had shifted away from the motor region found earlier, and moved into the more cognitively- and motivation-relevant nucleus accumbens region. Further analysis showed that the degree of fit was modulated by the medication dosage of each patient. No significant effects of medication had been observed earlier. Therefore this presents a specific and theoretically-constrained means of assessing medication and its effects on BG function. The emphasis on motivational striatum is also of key relevance given the potentially detrimental effects medication can have on negative symptoms.

6.3.8 The Model x Instability interaction implies that model-fitting would be appropriate

Each of the pathological computational models did not vary the cortical inefficiency parameter, therefore it was not expected that a model x Instability interaction would be seen. However, one did emerge in left nucleus accumbens. This could suggest that the default parameter used to represent cortical inefficiency was too high, hence the improved fit in patients with higher Instability. In order to determine this, it would be helpful to fit the c_{ton} and λ_e parameters for each patient, just as was done earlier with the first set of healthy volunteer data.

Chapter 7

Discussion

This work has been an attempt to iteratively develop a system-level appreciation for the neurobiological dysfunction underlying schizophrenia, combining computational and experimental methods. At the outset, schizophrenia was proposed to result from two key alterations:

- A dysregulation of cortical activity, mediated primarily via glutamate transmission. The overall effects of this were proposed to be twofold: an reduction in synaptic transmission and subsequent firing; and a diminished SNR. This was termed *cortical inefficiency*.
- An upregulation of subcortical dopamine activity.

Indeed, the body of work concerning computational models of schizophrenia has coalesced around four themes: cortical stability, SNR, gating and motivational salience, all of which relate either to cortical efficiency, dopamine levels, or both.

Cortex and the dopaminergic midbrain interact recurrently - it is therefore unknown whether one of these alterations induces the other, or if they in fact represent independent pathological processes. Within this work, cortical inefficiency was proposed as the primary disturbance in schizophrenia, and a means by which this could lead to disinhibition of the midbrain was proposed, which was partly dependent on BG circuitry.

The striatum is a particularly important brain region when considering the neuropathology of schizophrenia, as highlighted by the affinity antipsychotic medication has for striatal D2 receptors. However, the striatum is an inextricable component of the broader BG circuitry - and this circuit seems optimised to perform action selection. If the term ”action” is generously defined to extend beyond behaviour, and additionally encompass phenomena such as plans, working-memory maintenance and internal dialogue, then dysfunctional action selection has a clear

relevance for schizophrenia. On the one hand, if selection performance developed a particular inertia and inflexibility, this could lead to symptoms traditionally classified as negative (or deficit): poverty of action, speech and thought; blunted affect; and avolition. Conversely, a selection mechanism that had become unstable might lead to more disorganised symptoms, such as thought disorder, distractibility and hypervigilance. This work therefore sought to relate altered BG responses to *Instability*: a symptom score-derived metric designed to capture hypothetical selection dysfunction.

The BG have a number of properties that make them suitable for the approach taken here:

- They are massively innervated by cortex and thalamus, and in turn exert a significant recurrent influence over those inputs. They could therefore serve as a crude ”bellwether” for the brain as a whole.
- Its larger nuclei are spatially resolvable using fMRI. The caudate, putamen and pallidus are all anatomically well-defined and delineated. This is in contrast to the majority of cortex, and other subcortical structures such as thalamic nuclei.
- The implications of intrinsic BG circuitry are beginning to be understood in broad terms. This has reached the point where meaningful computational models can be developed to relate nuclear-level activity to behaviour, most notably for action selection. Alternative models concerning BG function - for example the mathematical models of temporal difference learning - provide a strong account for function, but lack clarity as to how this directly relates, or could be implemented by, the underlying neuroanatomy.

These characteristics are succinctly captured by the model of BG developed by Gurney and colleagues. This was therefore used as a simplified but principled source from which to derive predictions of activation while performing tasks of varying selection demand.

7.1 The Competition-Selectivity decision-making task

To this end, a new perceptual decision-making task was developed that manipulated both the Competition between two possible actions, and the overall demand to select one and not the other (Selectivity). This task fulfilled a number of important criteria:

- It enabled both these factors to be manipulated parametrically in a 2 x 3 factorial design. For Selectivity, this was in contrast to previous tasks of response inhibition, which

compare the execution of an action to its successful withholding. Such comparisons are innately poorly controlled, having imbalanced motor components. One feature believed to be unique to this task is the inclusion of finger extension as a motor response - this allowed for a broader repertoire of comparable responses within the task, taking advantage of specific mechanical interdependencies to manipulate Selectivity. Correctly measuring the participants' responses required the use and programming of an unorthodox piece of hardware: the fibre-optic data glove. In this task, not only were all decisions associated with a motor response, but these responses were fully counterbalanced across the factors of Competition and Selectivity, removing any specificity from the motor elements underlying each condition. It is possible that the suppression demanded by movement incompatibility could be resolved intracortically. However, the task conceptualized these movements as the products of different action plans, called into explicit Competition. There was no cortical response to Selectivity, even following small volume correction with liberal thresholds. Although one cannot infer an absence of effect, the results argue against a direct role for cortex in mediating Selectivity. Previous investigations of response conflict/inhibition have not consistently shown pallidal activation ([471], though see also [32, 472]). This could be due to the choice of task, as Go/No-Go, Flanker and Stroop paradigms either suppress single motor plans, or select one action in the presence of distractors. These indirect competitions occur over a mixture of perceptual, planning and motor execution levels. By varying the degree of mechanical incompatibility and thus conflict between discrete movements, this task was able to focus on the basal ganglias' central role in action selection, offering an improvement on studies contrasting imbalanced motor responses, e.g. single versus multiple simultaneous finger movements [472].

- The parametric design also provided more variance for the computational model simulations - if the manipulation was too crude and coarse-grained, the model would be less able to provide an adequate account for the subtleties of the activation, including non-linearities and Competition x Selectivity interactions. As the stimuli were drawn from a continuous colour spectrum at fine intervals, the factor of Competition was sampled at a high rate. However, it should be noted that for the analyses that didn't involve directly regressing the data with simulation-derived time courses, Competition was considered at three levels in order to boost the power of difference detection, rather than optimising the efficiency of detecting the precise response to Competition across its full range.
- It enabled for a clear disambiguation of the Competition between actions from the overall

intolerance for dual (i.e. simultaneous) action selection. Indeed, Selectivity is of particular relevance for schizophrenia, as this is not a disorder best characterised as being dominated by ”impulsivity” (which would be more analogous to dysfunctional response inhibition), but more one of inappropriate associations, disturbed goal-maintenance and mismatches between what is predicted and experienced perceptually (for example, one’s inner voice). It could be argued that all of these are better described as resulting from a poor resistance to interfering ”actions”, i.e. inappropriate simultaneous selections. This could allow associations to be formed between erroneously concurrently-gated phenomena; currently-maintained plans to be derailed by phenomena of minimal salience; and (somewhat more speculatively) a failure to suppress mechanisms associated with external speech appreciation during internal speech generation. Therefore Selectivity is providing a new and more direct means of assessing a potential core feature of schizophrenia.

- The task had minimal working memory demands, given both the training participants received beforehand, and the ever-present onscreen cues, which were spatially-compatible with the associated motor responses. Working memory deficits are well-documented in patients with schizophrenia, and any task that called upon this faculty to be intact would innately be confounded in any group comparisons. Theoretically, these working memory deficits could be partly due to inadequate maintenance via BG selection mechanisms (see section 2.2.2), and so it was especially important that this confounding factor was eliminated when trying to assess the effects of Competition and Selectivity in isolation.
- The nature of the stimuli meant they could be finely calibrated on a participant-by-participant basis, enabling a more precise and fair manipulation of Competition. One issue frequently encountered in functional imaging studies comparing patients with schizophrenia and healthy controls is that of relative difficulty. This has resulted in some studies finding ”increased” activation in patients versus controls, and others finding the inverse. This is assumed to be dependent on variations in precisely where a particular task has sampled each groups’ ”inverted-U” response curve. In this task, the issue of difficulty has been partly addressed by individually determining each participant’s colour perception parameters, and then producing tailored stimuli sets that uniformly sample Competition-space. We can therefore be more confident that Group x Competition interactions are being assessed on a like-for-like basis. The stimuli also took advantage of the observation that colour perception does not appear to be intrinsically compromised in patients with schizophrenia. An alternative approach could have been to use stimuli comprised

of blended images of faces and houses [325]. This has the advantage of encoding the two dimensions of the Competition within different cortical regions, allowing for a better disambiguation of the underlying decision-making steps. However, facial processing is known to be especially affected in patients with schizophrenia, be it by basic perception [342], or through altered appreciation of emotions [546]. It therefore seems unsafe to rely on competitions that contain such an intrinsic intergroup bias.

This task - designed to directly interrogate Competition and Selectivity - was then used to validate the original Gurney model.

7.2 Developing the original model

The original model was found to provide a good account for pallidal activation, especially its response to Selectivity. This was an important step in the validation process, as contrasting Competitions involving finger flexions and extensions was novel, and its successful implementation crucial to the whole project.

However, the model was not readily able to provide a good account for the striatum's responses to either Competition or Selectivity. If this approach was to have any strength in producing meaningful comments on the neurobiology of schizophrenia, it would certainly need to be able to provide an adequate account for normal striatal activity. A number of changes were therefore made to the Gurney model, which were firmly grounded in physiological evidence, and intended to directly address the model's functional shortcomings.

7.2.1 The pallido-FSI projection

The first major discrepancy was the presence of a Selectivity effect on striatal activation: this original model predicted no such thing, and described no mechanism by which it could occur. It was therefore decided that a pallidostriatal pathway should be included, enabling the regulatory GPe to exert its adaptive influence over striatum. This pathway is usually omitted from conceptualisations of the BG. However it is substantial [522], conveys feedback from the STN [547], and is known to be of clinical relevance: dopamine depletion induces a loss of coordination between GPe and striatum, with striatum lagging GPe [548].

As neuroanatomical evidence suggested that the pallidostriatal pathway synapsed on striatal FSIs rather than MSNs, FSIs were also included. Given their place within feedforward inhibitory

striatal circuits, they provided an ideal means to implement pseudo winner-take-all functions within the striatum, offering a means to account for the Competition response. Indeed, mathematical analysis of the modified model found that the cortico-FSI and GPe-FSI projections were instrumental in improving the model's fit for striatal data. These modifications were also quite robust, as the model's ability to qualitatively fit the data remained stable over a broad range of parameters, with the fit reducing gradually as the new parameters were decreased. They were also both found to have a significant bearing on the system's performance, with increased FSI activity leading to a stricter, more negative mechanism. This remained compatible with and further augmented the indirect pathway's regulatory and disinhibitory role in competitive action selection, particularly its ability to adapt to prevailing conditions. Recent physiological work in rats supports the model findings reported here: FSI activity was shown to sharply increase at the moment of initiation. Acting in concert with neighbouring MSNs, with encoded the chosen action, whereas others nearby encoded the opposite action - supporting the idea that channels subserve actions. The moment of initiation also coincided with pallidus inactivation, which acted to globally disinhibit FSIs. The degree to which this occurred was important in determining how firmly competing actions were suppressed [549].

The Gurney group's independently-developed FSI implementation focuses on the architecture of MSNs and FSIs, and how this affects the extent of their dendritic fields, and the proximity of inhibitory influences. They describe how MSNs are most strongly inhibited by FSIs on the edges of their dendritic fields, and that these FSIs are greatly interconnected via gap junctions. This allows FSIs associated with other channels to instantiate a winner-take-all type function, which is compatible with modified model presented here, although one difference is that channels "self-suppress", as the factor of proximity has not been instantiated. The interconnectedness that results from the FSI gap junction network is also compatible with the summed and diffused nature of FSI output described in chapter 5.

7.2.2 Dynamic dopaminergic midbrain

For schizophrenia, the GPe's special relationship with D2 receptor action offers further insights into how medication exerts its corrective influence. Dopamine's role was elaborated on in a final modification that was key to exploring the potential interactions between diffuse cortical dysfunction and subcortical dopamine levels: the instantiation of dopaminergic midbrain. The original Gurney model implemented dopamine as a fixed parameter that could be raised or lowered to explore Parkinsonism vs psychosis. However it is clear that tonic dopamine levels

are not static, and do vary meaningfully on a trial-by-trial basis, for example to represent the uncertainty surrounding a given decision [550]. As for the proposed compensation involved in schizophrenia, this too represents a dopaminergic ”upregulation in the face of adversity”, enabling selection in arduous situations. In the modified model, the fine-tuning of this response was also partly determined by an indirect GPe influence, again allowing it to takes its integrated account of the system as a whole, and use this dynamic assessment to ultimately disinhibit the midbrain. Although this was not directly assessed during model simulations (and therefore represents a potential avenue of future work), the D2-antagonism of medication would act to increase control-pathway striatal MSNs activity, which in turn would inhibit GPe, reducing the disinhibition it is able to exert over midbrain, serving to lower dopamine levels overall: this represents a possible mechanism by which an overly disinhibited system can be reigned in, but this would not be without significant - possibly deleterious - effects on performance.

Having modified the model, it would have been interesting to more fully explore the roles of D1 and D2 receptors, as these were not assessed in isolation. The D1/D2 dichotomy is of great potential therapeutic interest [551, 552], and gaining a better understanding of how each could theoretically affect function and behaviour could lead to the development of useful clinical tools.

7.2.3 The modified model integrates with the function of the original

The inclusion of these additional anatomical features was motivated by a desire to provide a better account for activation observed in healthy controls. However, if such additions proved counterproductive in terms of the model's ability to perform action selection, then they would clearly lack face validity. None of the three major additions hindered absolute selection over a broad range of parameters, suggesting that all were compatible with the BG's fundamental role. However, they all served to modify action selection performance: the cortical input into dopaminergic midbrain, cortico-FSI (w_f) and FSI-direct pathway MSN (w_{f1}) projections served to increase the strictness of selection, whereas elevated dopamine release, the GPe-FSI feedback (w_{fg}) and FSI-indirect pathway MSN (w_{f2}) projections engendered a more flexible system (figure 7.1). These additions have served to augment to original architecture - disinhibition of the indirect pathway serves as the primary means of regulating how the system responds to varying demands. Modulation of the direct pathway serves to focus selection, preventing the selection of multiple channels. In the context of schizophrenia, this model highlights the particular importance of the D2-mediated indirect pathway in allowing for a looser, more flexible system. In the event that any compensatory mechanism operating along these lines enters a

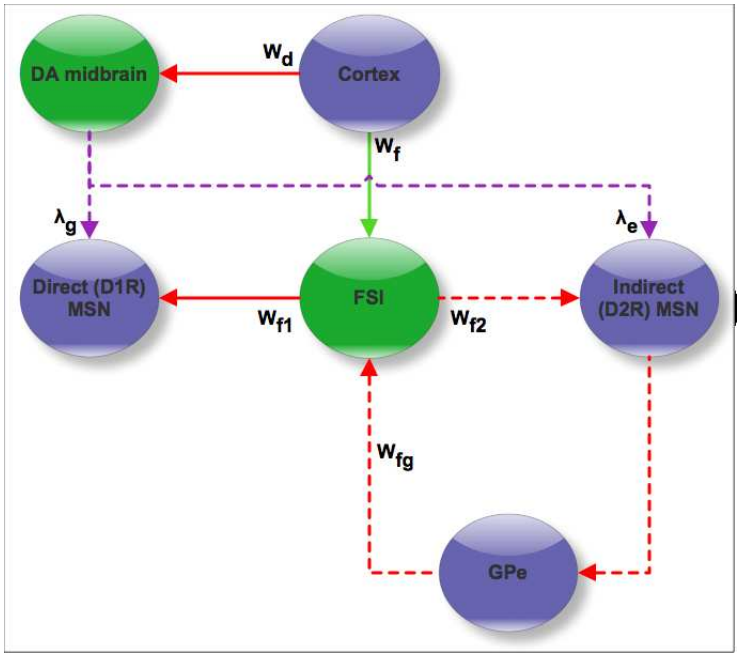


Figure 7.1: The influence of the new model nodes and parameters on selection performance. New nodes are shown in green, original nodes in blue. Excitatory connections are green, inhibitory connections red, and modulatory connections purple. Connections which were found to increase the strictness of selection are represented with solid lines, whereas dashed connections are those that produced a more flexible, promiscuous system.

runaway state, the specific role of D2 antagonists becomes clear.

7.2.3.1 Assessing performance

Action selection performance was assessed using two related measures, designed to capture different aspects:

- The point of promiscuity: the precise salience levels for which dual selection was permitted. Given two competing channels with saliences c_1 and c_2, a more flexible system would allow for a larger $c_1 - c_2$ interval, given that $c_1 + c_2 = 1$, i.e. less competitive situations could result in dual selection. A smaller $c_1 - c_2$ interval would reflect a more rigid system - this measure was intended to be more analogous to negative symptoms.
- distractibility: the minimum salience value that was able to override a weak but established channel with $c = 0.6$. Again, a more flexible system would permit smaller salience values to destabilise the established channel. This measure was intended to focus more

on selection stability, having a stronger bearing on symptoms coming under the umbrella of disorganisation.

Simulations found that these measures were closely correlated: systems that were highly distractable/unstable were also prone to dual selection at lower Competition levels. Therefore each of these measures made a limited unique contribution to assessing performance. Alternative measures could be developed to provide more useful information, with especial relevance for the symptoms of schizophrenia.

7.3 Verifying the modified model

The modified model was confirmed to both still behave as an order-preserving action selector, and provide a better fit for the original group of healthy controls' BG activation. However, some aspects of model performance were not fully explored: for example, whether or not the new model would scale appropriate with an increasing/decreasing number of channels. For the original model, capacity scaling was an emergent feature of the STN-GPe feedback loop. Given the addition of a feedback loop between striatum and GPe, it would be important to establish that such an anatomical modification does not prevent a function that is crucial to adaptive central action selection.

By fitting the modified model to a second, independent set of healthy controls, it was possible to examine whether it had been overfitted to the original data set. When considered in its entirety, the Gurney model has many parameters, and it is arguable that these could be "tweaked" to fit the data in a way that isn't particularly biologically-informative. However, the majority of these parameters remained static across simulations, with only two being nominated *a priori* as representing the neuropathology of schizophrenia. The most of the newly-introduced parameters were set according to mathematical analysis before simulations commenced: the rest were set following preliminary simulations, at a value that produced activation patterns resembling those seen in the first data set.

7.3.1 Gains and losses in fit following model modification

As for the original data set, the modified model provided a good fit for the striatal activation of the new second group of controls. However, the precise location of peak best fit shifted from left putamen to right anterior striatum. The anterior shift suggests that the modified model is

providing a better account for the motivational aspects of the task - evaluating the respective saliences of each competing action, rather than addressing them in purely motor terms. This is also in line with the left-to-right shift: an especially prominent role for right-sided cortical and subcortical structures has been found in studies of behavioural inhibition [32, 318], attention deficit/hyperactivity disorder [553].

The modified model did not provide a good account for pallidal activation, which was unfortunate, given the GPe's prominent theoretical role in regulating action selection. Examining the parameter estimates for each of the experimental conditions revealed that the pallidal responses were qualitatively different to those seen during the first study. The one major alteration made to the task itself was in the inclusion of trial-by-trial feedback. This decision was made in light of the overall weakness of striatum's activation during the original study, which could have been because the associated reward from which task salience was derived was delivered after the scan itself. This was perhaps too abstract, with the participant having minimal conscious appreciation for the reward during the scan, especially after getting into the routine of responding to a long stream of stimuli, with no indication as to performance. However, including feedback seems to have had unforeseen consequences on pallidal activation. It appears that the task may have developed an element of motor learning, although the intention was for this not to be the case, with all learning having happened beforehand during training. Closer inspection of the feedback received by participants revealed that there were occasional discrepancies: the feedback would sometimes report a response as being incorrect when the raw glove data suggested this wasn't actually the case. The means by which the glove data was assessed within the scanner was not ideal, and could certainly be technically improved. As pallidus is known to be dynamically modulated according to the value and probability of experiencing rewards [554, 555], as well as the experience of pleasure [556] (or in this case displeasure, given the frustration of receiving inaccurate feedback), then it is unsurprising that the pallidus is responding in a different way to that seen during the earlier study. The anterior cingulate cortex showed a robust response to Competition during the second study, but not during the first. Response conflict is known to engage anterior cingulate cortex [461], a structure believed to perform conflict monitoring [468]. Evidence suggests it plays a more evaluative than regulatory role [469], updating motor strategies and so improving subsequent performance [470]. The presence of cingulate activation during the second study does support the view that trial-by-trial feedback has introduced a further level of computation to the task. It can therefore be argued that the modified model is not overly-fitted to the original data set, but that the new data set is complicated by the new factor of motor learning. The

model makes no attempt to describe the processes underlying learning, as its main goal is action selection. This would be a desirable elaboration however, as the negative-disorganised spectrum focused on here provides no account for symptoms such as delusions, which are most likely the result of inappropriate associative learning. Given pallidus' diminished fit, it is unclear why striatal activation has been preserved: it could be that it is encoding values over a broader time window, with occasional perturbations being "absorbed within the average", making it less susceptible to these occasional feedback anomalies.

7.3.2 Estimating "Healthy" values for key parameters of interest

By modifying the model and so allowing it to provide a more accurate account for the activation seen in healthy controls, the intention was to determine how well the theory matched the evidence by varying the parameters describing cortical inefficiency and the effects of medication. Before this could take place, appropriate Healthy values for cortical efficiency (c_{ton}) and D2 antagonism (λ_e) had to be estimated, to establish a foundation from which systematic parametric variation could take place. When fitting to the healthy control data, a relatively constrained parameter space was explored. This was partly due to the limited computational resources available, making broader, finer-resolution sweeps too time consuming. Ideally, this would be expanded - not only in fitting to healthy control data, but also to that of patients. However, an examination of how fit improved with those parameters suggested linear trends: the model provided a superior account as cortical efficiency increased, and D2 efficacy decreased. Both these states are most compatible with the model's envisioning of "Healthy", as the theoretical dysfunction associated with schizophrenia proceeds in the opposite direction. This suggests the model has a certain validity, as well as a degree of stability, as the landscape drawn by these two parameters is simple, with no local minima.

7.4 Selection and schizophrenia

Having modified the model to provide a better account of striatal activation in healthy controls, isolated two parameters of clinical interest, and determined appropriate "Healthy" values for them, the neurobiology and symptomatology of schizophrenia were explored as a function of this parameter space. In analysing the patient data, two approaches were adopted: first the interactions between Competition/Selectivity and Instability/medication were assessed; then the data were regressed with simulated time courses produced by models with parameter sets

defining particular hypothetical illness states.

7.4.1 Cortical inefficiency and hypofrontality

Throughout this work, it has been argued that both cortical inefficiency and subcortical dopamine upregulation underlie schizophrenia, and that the former leads to the latter through a disinhibitory process. In this work, it is assumed that cortical activity is reduced overall, in reduced SNR is proportionally more related to a loss of signal rather than increased noise. This highlights the issue of how cortical inefficiency relates to ”hypofrontality” [557]: the idea that frontal activation is relatively increased or decreased from some ideal is obviously crude, and vulnerable to confounding factors such as performance. This could especially be the case for the patients scanned during this study, as they were specifically selected to have above-average IQs, to ensure they could perform the task. They are therefore not representative of the greater population. However, the control group was matched, and the ability to ensure each participant's colour stimuli were appropriately calibrated meant that ”difficulty” could be objectively determined.

On balance, hypofrontality is supported by several functional imaging studies [558]. Further to this, people expressing prodromal symptoms show abnormal prefrontal ERP responses, which correspond with decreased glutamate/glutamine levels in thalamus [559]. As thalamus is a key afferent projection of cortex, this suggests a dysfunction of cortex that results in deceased overall activity. Functional evidence of deficient frontal cortical and lenticular activity has been observed in people at high risk of developing schizophrenia [560]. As morphological changes require time to become manifest at a level appreciable using MRI, this implies an ongoing and relatively well-developed process has occurred.

The first healthy control study replicates the finding that IFG responds to increasing ambiguity, uncertainty or effort-demands [473, 474, 334, 475]. This gyrus is believed to contribute to planning by generating goal-oriented forward models [477, 476], therefore increased activity may reflect the complex goal landscape associated with ambiguous decisions [478, 480, 476]. Interestingly, withholding an initiated response - for example during ‘Stop’ tasks - also activates IFG [481, 32]. Forward models must presumably be rapidly updated following sudden goal alteration, perhaps contributing an augmenting suppressive influence along the “hyperdirect” corticosubthalamic pathway. In the second study, patients show a flattened response to Competition within IFG. The consequences of losing the integrity of this goal-related signal will be broad-ranging, and could underpin several symptoms of schizophrenia. Deficient forward

model generation is seen as a core feature of schizophrenia [490], and IFG has shown both abnormal activation [491] and diminished correlation with task performance in patient groups [561]. This is especially marked in patients with pronounced negative symptoms [493]. The distorted encoding of forward models would theoretically result in many mismatches between expectation and perception, perhaps leading to an altered and confusing sense of reality. If these mismatches occur in the context of learning, they may induce the phasic dopamine signals believed to encode prediction error, resulting in aberrant association formation.

7.4.2 Frontostriatal-mediated dopaminergic disinhibition

This work proposes that at a later stage, cortical inefficiency induces an upregulation of tonic subcortical dopamine release, and recent multimodal imaging supports such a possibility. Comparing people at high risk of developing the illness to healthy controls using both fMRI and $\left[^{18}F\right]$ fluorodopa positron emission tomography (PET), Fusar-Poli and colleagues produced several instructive findings in those at risk: reduced middle frontal performance-matched activation; elevated dopaminergic activity within associative striatum; and a negative association between middle frontal activation and striatal dopamine levels [562]. The middle frontal cortical area in question is the same as that which has been shown here to have a diminished SNR in patients - a signal that gets progressively weaker in patients having greater Instability. It is also the ”decision” region replicated in several other studies [335]. This frontostriatal association between deficient cortical activation, elevated striatal dopamine and an increasingly unstable/disorganised symptom profile ties in very well with the model described in this thesis. This work also found evidence of elevated dopaminergic activity in patients that became more exaggerated with Instability.

7.4.3 The progression from prodromal to frank schizophrenia

The hypothetical prodromal state of schizophrenia was described as involving moderate cortical inefficiency, but an as yet minimal dopaminergic compensation. Simulations of this state suggested an overly-rigid selection mechanism, with cortical representations of salience having a reduced SNR, leaving them less able to induce full disinhibition of their associated channels. More likely than not, this would result in dysfunction along a more negative/deficit dimension: indeed it has been suggested that negative symptomatology may be an early indication that someone has entered an elevated risk state, occurring significantly before the onset of the more positive symptomatology associated with the prodromal state [563]. There are also

significant relationships between the types of symptoms seen during the prodrome, and those expressed once the illness is established: negative prodromal symptoms can be especially predictive of a subsequent disorganised illness state, in particular poverty of content of speech, a lack of initiative/drive, blunted affect, marked withdrawal, impaired personal grooming and concentration [564]. This lends support to the novel Instability measure defined in this thesis, which contrasts symptoms suggestive of strict vs flexible selection mechanisms, broadly analogous to negative and disorganised symptoms. It also suggests that patients with a more pronounced negative prodrome (which would be theoretically associated with greater cortical inefficiency) are especially prone to degenerating into a disorganised state, which would be the case in the event of runaway dopaminergic compensation. The association between negative prodrome and disorganised illness was quite specific; while a positive illness state was associated with prodromal ideas of reference and odd beliefs [564], which are both theoretically more attributable to dysfunctional associative learning. Imaging studies of reinforcement learning in patients suggest a pronounced dysregulation of the expected phasic release of dopamine [247, 545], which contrasts to the hypothetical upregulation of tonic dopamine described throughout this thesis. However it is known that tonic dopamine acts to depress phasic release via D2 autoreceptors [565]. If tonic dopamine was pathologically elevated as a result of a cortical disinhibition of midbrain, this could lead to deficits in associative learning, as the required phasic dopamine signals are undermined. Interestingly, the depressive effects of tonic dopamine on midbrain neurons appear to undergo desensitisation over prolonged phasic bursts: although earlier spikes are inhibited, later spikes show a degree of potentiation [565]. The nature of dysfunctional associative learning in schizophrenia is not simply a reduced ability to learn relevant associations, but an increased tendency to learn irrelevant ones [247], which shows an association with dysfunctional midbrain activation [545]. Inappropriately potentiated spikes occurring significantly after a moment worth potentiating could lead to inappropriate associations developing between co-occurring but otherwise unconnected neutral stimuli and events. As phasic dopamine release is subject to a more enthusiastic reuptake mechanism, it is unclear how a phasic release problem could be primary: tonic levels would not be dramatically altered, therefore probably remaining unable to modulate selection as described here.

Abnormal associative learning is also seen in those at ultra-high risk of developing schizophrenia, but not in currently-healthy first-degree relatives [566], suggesting that this is a later manifestation of an ongoing pathological process. The modulatory interactions between tonic and phasic dopamine release are especially pronounced in the striatum's limbic section [567]. This potential vulnerability of the limbic system tallies with the recent observation that people at

high risk of schizophrenia show a greater tendency to attribute angry emotion to neutral faces, and that this is especially the case for those showing a loss of inhibitory control.

In summary, one could paint the following hypothetical picture of the underlying progressive process:

- The development of schizophrenia is initially characterised by negative symptoms, driven in part by a NMDAR-related loss of cortical efficiency.
- Through cortex's feedback loop with dopaminergic brain, an overall reduction in cortical activity results in an upregulation of subcortical dopamine release (and a concomitant decrease of dopamine release within cortex [15, 529]). This striatal dopamine facilitates selection, reducing negative symptoms.
- The primary pathological process affecting cortical efficiency continues over time, and coupled with the decrease in cortical dopamine leads to a situation requiring ever greater levels of subcortical dopaminergic compensation.
- Eventually tonic dopamine levels reach a point where selection is excessively facilitated, and a more disorganised symptom set begins to manifest. Not only is selection affected, but the background levels of dopamine are also able to activate D2 autoreceptors, resulting in the derailment of the phasic dopamine release underlying associative learning. This underlies the development of a broad range of delusional ideation.

7.4.4 Primacy

This is one possible explanation for how underlying neurobiological dysfunction can lead to runaway compensatory mechanisms and the expression of symptoms. However, cortico-BG and cortico-mesolimbic relationships are locked in multiple feedback loops. A fault along any point of these loops could induce the same overall effect, and the issue of primacy remains to be resolved. It is perfectly feasible that the initial dysfunction is dopaminergic, with subsequently reduced mesocortical activity resulting in diminished cortical SNR. However, elevated dopamine activity in people at risk of psychosis has been found to be intermediate between that of healthy controls and patients, becoming more elevated with increasing prodromal symptoms [568]. This suggests that dopamine levels are varying along a continuum of compensation, and that once a critical threshold is reached, a patient develops the illness proper. It could be that striatal D2 receptor upregulation is the primary problem: within the constraints of the modified Gurney model, and with all other things being equal, this could lead to a selection mechanism

with enhanced suppression, more pronounced negativity, and a recurrent reduction in cortical activity. However, PET studies of medication-naive first-episode patients have failed to find altered striatal D2 receptor binding [569, 570]. Alternatively, the primary problem may concern D2 autoreceptors within dopaminergic midbrain: this could dysregulate tonic and phasic dopamine release, again resulting finally in cortical inefficiency [571]. Lower D2/3 receptor binding has been observed in the midbrains of medication-naive patients with schizophrenia [572]. Further downstream, the initial insult may be within thalamus: a site where altered D2/3 receptor binding has been confirmed in medication-naive patients [569, 570]. Although presenting a complex puzzle to be solved, these recurrent loops also offer multiple potential therapeutic targets, both neurochemically and anatomically [489].

7.4.5 Instability

This work proposed that a reduction in cortical efficiency is the primary insult in schizophrenia, and that this would induce the following trajectory: first, a reduction in cortical SNR would be observed; second, the overall drop in cortical activity would disinhibit dopaminergic midbrain, leading to an upregulation of subcortical dopamine release; and third, this exaggerated dopamine response would destabilise selection, leading to symptoms broadly analogous to disorganisation, but summarised here as Instability. The novel measure of Instability, derived from PANSS subscores for symptoms believed to be especially indicative of a dysfunctional selection mechanism, was found to bear fruit. In summary, it related to three interesting results:

- Patients with elevated Instability showed a loss of SNR in middle frontal gyrus, a cortical area believed to encode chosen value during decision-making. Controls showed a trend for an appropriate Competition response in this region. This supports the hypothesis that greater selection dysfunction - Instability - will be associated with reduced cortical efficiency.
- Patients with elevated Instability showed an exaggerated response to Competition within dopaminergic midbrain (substantia nigra). Controls showed the same, albeit more restrained response to increasing Competition, in line past findings that tonic dopamine levels increase with decision uncertainty [573]. The model predicted this as comprising part of normal function, serving to facilitate selection in circumstances where it might otherwise be quite difficult to reach threshold. In patients, this response is exaggerated -

and increasing so for patients showing marked selection dysfunction (Instability). Combined with the previous finding, this supports the theory that cortical inefficiency leads to selection dysfunction through dopaminergic disinhibition.

- Patients with elevated Instability were more likely to display the inverted striatal Competition response predicted by ”Ill” simulations of the model, and that this was especially the case when selection demands were at their highest (i.e. with high Selectivity). This symptom-derived metric of selection dysfunction therefore seemed to modulate the degree to which the ”Ill” i.e. cortically inefficient model accounted for activation within this key selection nucleus.

This trifecta of results resembles a cohesive body of evidence supporting the idea that schizophrenia is characterised by both cortical inefficiency and an upregulation of subcortical dopamine activity. Instability - a proposed symptom-derived correlate of selection performance, and indicator of cortical inefficiency - is found to have a special bearing on selection-related BG function, above and beyond the influence of medication. Thankfully there was no significant correlation between the patients' medication dosage and Instability, allowing for a disambiguation of their effects.

The measure of Instability was determined *a priori* by examining each PANSS subscore in turn, and asking two questions:

- Is this symptom a likely consequence of a dysfunctional selection mechanism?
- Does it imply an overly-flexible or -rigid system?

This approach was chosen as it was seen as a more focused attempt to ask a specific question of the data, as opposed to techniques such as factor or cluster analysis which ask the data what the question should be [574, 575, 576, 577, 578, 579, 580, 581]. Factor analysis offers a potentially powerful means to detect trends and covariances that could otherwise go undetected. However, such techniques must necessarily include as much data as possible - usually from a heterogeneous population with multiple confounding factors. Levels of medication, comorbid drug use and abuse, duration of illness, comorbidity with more affective disorders: all will cloud the nature of ”pure” schizophrenia, possibly smothering evidence of the underlying neuropathology. Although such models do include these variables, there is a danger that complex interactions within and between the confounds and symptoms will not be adequately captured [582].

Although several models have been produced, having between three and five factors, they are generally variations on a theme of Liddle's original 3-factor model [310], which categorised Scale for the Assessment of Positive/Negative Symptoms (SANS/SAPS [583, 584]) items into the groups of psychomotor poverty, disorganisation and reality distortion. There is a strong overlap between the symptoms of psychomotor poverty, and the PANSS items chosen to reflect "rigidity" in this work. Likewise, Liddle's disorganised symptoms are similar to those described here as "Unstable". Therefore there is a degree of convergence both from the data-driven and *a priori* selection-specific approaches. As alluded to in section 7.4.4, the symptoms of Liddle's reality distortion group could result from a dysregulation of phasic dopamine release, and go beyond the remit of this particular model of action selection.

However, the *a priori* selection-specific approach was not perfectly aligned with Liddle's 3-factor model: for example "poverty of content of speech" was categorised here as a symptom of rigidity, whereas Liddle classifies it as "disorganised". This approach therefore traverses the clusters to some extent, and has produced functional results that are strongly suggestive of both selection dysfunction and the hypothetical neurobiology of schizophrenia. There is some evidence for negativity and disorganisation having a correlation with both behavioural and functional evidence of selection dysfunction: in people with schizotypal personality disorder, those with more pronounced negative symptoms show worse performance during the Wisconsin Card Sorting Task, whereas more disorganised people has preserved function [585]. Patients expressing formal thought disorder show elevated right caudate activation while describing Rorschach stimuli [586]. Behaviourally, failures of inhibition are more associated with the disorganised vs negative state [587]. However it is believed that the results described in this thesis are more informative, as a more focused question is being asked, working from a firm theoretical background, using a task that parametrically explored two selection dimensions. The Instability measure may be especially useful given that the disorganised component of schizophrenia is often neglected in research.

7.4.6 Medication

Obtaining a better understanding of how medication acts to ameliorate the symptoms of schizophrenia is crucial if we are to rationally develop better medication, or optimally tailor the current formulary to the needs of each patient. The BG model of action selection developed during this thesis is well-placed to be able to comment on the theoretical effects of D2 receptor blockade, both for predicting functional activation, and then to relate that to changes in symptoms. It is

hoped that this concept could be developed to provide an idea of ”optimal” dose, where the theoretical overcompensation of subcortical dopamine is brought back in line, but not to the point of exacerbating negative symptoms, or inducing Parkinsonism.

It was found that the ”Ill” model simulations were able to provide an account for the activation seen within patients' anterior striatum, and that the degree of fit increased as the D2 efficacy λ_e parameter decreased (i.e. as medication theoretically increased). Effectively a ”dose-dependent” was seen between the three different medicated simulations - the ”Unmedicated” provided the worst fit, whereas the ”Overmedicated” model did much better. To better determine whether this theoretical medication effect was actually related to the patients' dosages, the ”Ill-medicated” fit parameters were correlated to medication dose, and found to be positively related, i.e. the more medicated a patient actually was, the better the ”I'll-Medicated” model was able to describe their striatal responses, taking Instability into account. This suggests that D2 receptor blockade has a special role in mediating selection responses in striatum's more limbic/cognitive zone. This is reflected in genetic mouse models of D2R ablation within anterior-ventral striatum, which lead to exaggerated responses to reward [588]. If this finding could be replicated, it would be of great potential therapeutic use (see section 7.5.1).

The measure of medication used here was derived from the average dose the patient had received over the past two years, rather than a simple snapshot of the present. It was believed this would better capture the related effects of medication, as its use is clearly associated with longer-term structural alterations [589]. However, this difference has not been demonstrated, and it would be interesting to explicitly compare the two, to see whether the longer-term measure offers advantages of fit over the instant dose. If so, this approach may well be of benefit to future fMRI studies involving patients with schizophrenia.

It must be noted that disambiguating the negative symptoms induced by medication (Parkinsonism) from those of the disease itself can never be done in the patient group studied here, as there were all medicated. However, this work was unusual in that participants were explicitly and thoroughly assessed for signs of Parkinsonism using the UPDRS. Parkinsonism would represent a significant confound, and reflect a rather excessive dose of medication, therefore it was important to establish whether or not this was a concern for particular participants - this was found not to be the case.

7.4.7 Computational models and neuropsychiatric disorders

The complex and multifactoral aetiology of neuropsychiatric illnesses have allowed them to remain mysterious for quite some time. Pulling out the most informative themes, and trying to weave them into a cohesive argument requires formal frameworks that are both biologically-relevant, and mathematically tractable. This work represents an attempt along these lines, and it is a method that offers some strengths. As the hypothetical pathological processes leading to cortical inefficiency will likely be partly cancelled out by the dopaminergic compensation mechanism, and the further corrective influence of medication being loaded on top of that, producing and testing simplified predictions of behaviour and activation was not straightforward. Indeed, when the patients' data was addressed in simple terms, few significant differences were identified. It was only when the covariates of Instability and medication were included that informative results began to emerge. However this was superseded by the use of regressors derived from simulations of several "Ill" and "Medicated" states. This approach really tried to capture the most salient points of all that complexity, using just two parameters, and then produced very specific predictions of how the BG would then respond under varying selection demands. The level of fit that each of these simulations was able to provide for the striatum and pallidus could then be examined as a function of Instability and medication, providing a relatively valid and rigorous a means of testing the hypotheses concerning the neurobiology of schizophrenia.

The modified computational model suggested that a reduction in cortical efficiency c_{ton} would result in a marked inversion of striatum's response to Competition. This provided a useful, categorical prediction regarding what one would expect to see in the fMRI activation of patients performing the decision-making task. However, it could be that this inversion was merely an artifact of the parameter value chosen to represent the ill-state. The fMRI data did not support that theory however: the degree to which the Competition response became inverted depended both on the overall selection stress (i.e. it was more pronounced when Selectivity was high), and each patient's Instability - a measure believed to increase with cortical inefficiency. This triangulation of simulated activation, actual activation and the hypothetical symptom-defined consequences of cortical inefficiency argue for this being partially representative of the pathological neurobiological state of schizophrenia.

On the other hand, the modified model made no such categorical predictions concerning pallidal activation, and yet one was found: patients showed an inverted response to Selectivity. In general, it is believed that the inclusion of trial-by-trial feedback had altered the task beyond

the point where the model could account for activation within pallidus, as discussed in section 7.3.1, therefore this discrepancy is not seen as a failure of the model per se. The model was not intended to account for the processes underlying learning, and such an extension would be a large undertaking.

7.5 Proposed future work

7.5.1 Estimating patients' cortical efficiency and D2 efficacy parameters

A shortcoming of this work was that the simulations used to regress the data only used one parameter value of cortical inefficiency $c_{ton} = 0.05$ to represent the ill state. This was because the decision was taken to adopt a more principled approach *a priori*, defining a single value that was lower than that estimated as representing the Healthy state during the first study ($c_{ton} = 0.1$). It was not deemed to be appropriate to go "fishing around" for a value that happened to fit this patient data set fairly well. However, estimating the values of c_{ton} and λ_e that best fit each individual patient's data could provide some valuable insights, if these were to then be regressed by their Instability and medication dosage measures. It was observed that the degree to which the simulated data time courses fit each patient's data covaried with medication and Instability, therefore it would be useful to "pin down" the parameters describing the theoretical underpinnings, before relating them to medication and Instability, which are measures of the secondary consequences of the illness process. This could be especially useful for estimating the efficacy of each patient's medication level. Being able to find a simulation that gives a functional indication of D2 receptor blockade could help clinicians determine whether a patient's current symptoms are essentially primary in nature - for example, they have relatively uncompensated cortical inefficiency, and so are more negative - or are a result of excessive medication, in this case profound D2 antagonism may have removed the beneficial effects of the dopaminergic compensation. However, this work is very preliminary in nature, and the model's real predictive power remains to be tested. A more rigorous approach would be to conduct a case-control study, using a larger sample of patients, categorised as having low, medium or high doses of medication. The modified model could then be fit to each patient's fMRI responses, and the resultant D2 efficacy parameters compared between groups.

7.5.2 Explicitly modelling separate associative and motor loops

The Healthy modified model was able to provide a good account for putamen activation in patients, suggesting a relative preservation of function in this more straightforwardly motor region. However, only the Ill variants were able to fit activation seen in more anterior striatum, the implication being that the more motivational and cognitive processes are especially affected. This work has also highlighted how altered activation in both the inferior and middle frontal gyri are closely associated to Instability, a proposed symptom-derived correlate of selection dysfunction. In the first study, two task-responsive regions of pallidus were identified: one whose response to Competition satisfied the predictions of the original model, and a second that demonstrated an inverted response. Contrasts of Granger causality metrics describing the inputs these pallidal regions differentially received according to increasing Competition revealed that the inverted area received a specific input from IFG. The relationship between IFG haemodynamic time-to-peak and behavioral response time became ever closer as Competition increased, indicating ”cognitive” pallidus' increasingly dominant role over behavioural variance. This suggests that IFG's assessment of uncertainty is not post-hoc, but acts to guide action execution.

A useful elaboration of the model could therefore be the segregation of a frontal/associative/motivational loop from a simple motor loop, with each having different c_{ton} parameters to define a more localised dysfunction. Such broad-scale cortico-BG parallel loops are well-documented [269], although there is disagreement as to the degree of segregation [590] vs convergence [482] they undergo while projecting from cortex to striatum, and then to pallidus. The more likely situation is one of partial convergence [254, 291], especially in light of the spiralling nature of dopaminergic nigrostriatal projections [295]. A degree of segregation is necessary for the emergence of meaningful convergent subregions, encoding integrated variables along partially-open subloops [261]. Premotor cortex is believed to receive convergent projections [591], and has been shown to encode multiple action alternatives [463]. As sensory evidence accumulates, these options are progressively narrowed down, following a short delay. It could be that decisions pertaining to goals are encoded within frontal loops, and that these mediate this elimination process either by directly impinging on shared pallidus, or by feeding back to premotor cortex via thalamic loops, facilitating the isolation of optimal motor parameters [298]. Such an architecture would resemble a dual controller [467], which dichotomises cortical influences into those encoding habitual reward representations, and a hierarchical cognitive plan of short-term future response—reward associations. They propose that these systems operate over corticostriatal circuits, arbitrated according to their reward-predicting accuracy by anterior cingulate

cortex. A similar principle is seen in Haruno and Kawato's experimentally-supported heterarchical reinforcement-learning model [296]. Rather than selecting for action, this model uses spiralling nigrostriatal connectivity [295] to propagate limbic- and cognitive-derived dopamine reward predictions to motor areas. This refines the selection of rewarding stimulus-outcome associations. Their different but compatible implementation may have precedence during the learned acquisition of new associations, exploiting the temporal characteristics of phasic dopamine modulation. The framework described here perhaps predominates during selection within established contexts. Indeed, the BG are not the only nuclei implicated in centralised selection [484]: it remains arguable that several cortical and subcortical systems operate in parallel [485]. The BG may bias the refinement of detailed action parameters across multiple domains [298].

By teasing out the motivational/cognitive elements from more straightforward motor execution, the model may be better equipped to comment on how much medication is "too much": to determine that optimal point where the dopaminergic compensation is trimmed to a level where its presumed original function of facilitating appropriate selection is preserved, but without going so far as to induce cognitive rigidity (i.e. negativity) or Parkinsonism at the more extreme end of the motor spectrum.

7.5.3 Using Granger causality to probe cortico-midbrain interactions

A key question that this work has not been able to answer is that of primacy: is it cortical dysfunction that leads to dopaminergic upregulation, or *vice versa*? The results described here are compatible with the hypothesis that it is cortex which develops a problem first, however this is by no means the only, exclusive explanation. Isolating the true primary disturbance is essential for the developing our understanding of schizophrenia, providing insights not only into how it may be treated, but also into normal brain function. During the first study, Granger causality metrics were used to identify cortical regions that made a specific contribution to a cluster of activation within pallidus - a contribution that became more dominant with increasing Competition. It may be informative if this method was applied to the midbrain cluster found to show a Competition x Instability interaction in patients. Although the hypothetical disinhibition of midbrain would be operating at a temporal scale inaccessible to Granger causality, the specific exaggerated response to Competition is more likely to be varying on a trial-to-trial basis, and is therefore within the detectable range of Granger measures. Using temporal precedence, it may therefore be possible to make a statement as to the direction of information

flow between middle or IFG and the substantia nigra.

7.5.4 Developing a more physiologically-realistic model

The Gurney model uses rate-encoded leaky integrate-and-fire units to represent the activity of neural populations. In this work, these rates were assumed to be correlated with the BOLD signal observed using fMRI. However, such assumptions may be unsafe. Within the Gurney group, Humphries has developed a spiking version of the BG model [487] that offers improved physiological plausibility. However, it is unclear what such an elaboration would add to the implementation described here - were this a model of reinforcement learning, where spike timing is so important in initiating long term potentiation versus depression, then this would clearly be a great strength. One advantage of spiking models is there amenability to conversion to a BOLD signal, as has been done with SPM-based techniques such as psychophysiological interactions [592] and dynamic causal modelling [421]. Therefore moving to a spiking model could offer a means to make the model more compatible with the fMRI data it is intended to regress. However, unless the associated parameters were kept tightly constrained, this could balloon an already multidimensional model.

7.5.5 Adapting the decision-making task

The task made use of a specialist piece of hardware - the fibre optic data glove - to record high quality information concerning both finger flexion and extension. These are a series of relatively well-balanced movements that dramatically open up the limited behavioural repertoire available to participants taking part in fMRI studies. The rate at which data is sampled also allows for the speed and acceleration of movements to be determined. This could be of great value for looking for subtle drug-induced effects, such as the bradykinesia of Parkinsonism. As the rate of change of force is known to modulate pallidal activity in particular [402], a better fit may be established if acceleration is included as a nuisance covariate during first level analysis.

An alternative approach would be to adopt an entirely different motor response set - such as saccadic eye movements. They also allow movements along two very well-balanced axes, with saccades intermediate between horizontal and vertical shifts being detectable using eye-tracking equipment. These intermediate movements would arguably represent partial failures of suppression, and may provide ”odd-ball” moments of great relevance to BG function. Saccades are known to be dependent on the more anterior aspects of the BG, including caudate

nucleus and the substantia nigra pars reticulata [269]. There are therefore arguably better-suited to examining the specific parts of the BG that may be especially affects in people with schizophrenia - indeed, dysfunctional saccadic behaviour is a frequent observation in patients [593, 400]. This is especially pronounced in those with predominantly negative symptoms [594], has been related to deficits in selection such as perseveration [595], and is sensitive in detecting the effects of antipsychotic medication [596].

7.6 Conclusions

This work set out to examine the hypotheses that the primary dysfunction in schizophrenia can be summarised as ”cortical inefficiency”, resulting in part from NMDA receptor underactivity, and that this in turn induces an upregulation of subcortical dopamine activity by disinhibition. This has knock-on effects on the crucial decision-making stage of action selection, namely an increased tendency to disorganisation. To formalise these ideas, and create a means for deriving predictions concerning fMRI activation and behaviour, the BG computational model of Gurney and colleagues was iteratively developed to (a) provide a sufficient account for BG activation and (b) incorporate parameters describing cortical inefficiency and dopamine efficacy, to examine their theoretical consequences. These consequences were tested using a novel perceptual decision-making task, that used calibrated colour cues to evoke Competitions between responses which either were or were not mechanically compatible, varying the demand to suppress the competing loser, termed Selectivity. Following an initial validation study involving healthy controls only, the model underwent a series of biologically- informed revisions, improving the account it could provide. These were the inclusion of FSIs within striatum, an inhibition of those interneurons by GPe, and the explicit modelling of dopaminergic midbrain, which was influenced both by cortex and the BG circuitry. This greatly improved the account the model was able to provide for striatal activation in both the original data set, and within a second independent group of healthy controls. Having determined which parameter values for cortical efficiency and D2 efficacy best represented the controls’ activation, these parameters were then used to explore the proposed illness and medication states of the patients. It was found that ”Healthy” models provided a good account for patients’ motor putamen activation, suggesting preserved function, but only the ”Ill”, cortically-inefficient models could capture the activity seen in anterior striatum. This fit improved as the parameter describing medication dosage was increased. To examine how observed activation covaried with the proposed behavioural consequences of selection dysfunction, the new symptom-derived metric of Insta-

bility was created, which contrasted PANSS scores believed to reflect selection rigidity versus flexibility. This produced an interesting series of results: a loss of SNR in decision-encoding cortex was more pronounced in unstable patients; these patients also showed exaggerated dopaminergic responses within midbrain, and their striatum was more likely to resemble the patterns of activation seen in simulated ”Ill” models. This all lent some credence to the hypothesis that cortical inefficiency, dopaminergic disinhibition and behavioural Instability go hand in hand. This model may have therapeutic implications, as a means for assessing a person’s current illness state, and whether their medication dosage is optimal.

Bibliography

[1] Hoffman RE, McGlashan TH. Using a speech perception neural network computer simulation to contrast neuroanatomic versus neuromodulatory models of auditory hallucinations. Pharmacopsychiatry. 2006 Feb;39 Suppl 1:S54–64.

[2] McGlashan TH, Hoffman RE. Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Arch Gen Psychiatry. 2000 Jul;57(7):637–48.

[3] Wang XJ. Toward a prefrontal microcircuit model for cognitive deficits in schizophrenia. Pharmacopsychiatry. 2006 Feb;39 Suppl 1:S80–7.

[4] Arnsten AFT. Stress signalling pathways that impair prefrontal cortex structure and function. Nat Rev Neurosci. 2009 Jun;10(6):410–22.

[5] Brunel N, Wang XJ. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. Journal of computational neuroscience. 2001 Jan;11(1):63–85.

[6] Braver TS, Barch DM, Cohen JD. Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. Biol Psychiatry. 1999 Aug;46(3):312–28.

[7] Geva AB, Peled A. Simulation of cognitive disturbances by a dynamic threshold semantic neural network. Journal of the International Neuropsychological Society : JINS. 2000 Jul;6(5):608–19.

[8] Grossberg S. The imbalanced brain: from normal behavior to schizophrenia. Biol Psychiatry. 2000 Jul;48(2):81–98.

[9] Schmajuk N. Brain-behaviour relationships in latent inhibition: a computational model. Neuroscience and biobehavioral reviews. 2005 Jan;29(6):1001–20.

[10] Smith AJ, Becker S, Kapur S. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. Neural computation. 2005 Feb;17(2):361–95.

[11] Frank MJ, Claus ED. Anatomy of a Decision: Striato-Orbitofrontal Interactions in Reinforcement Learning, Decision Making, and Reversal. Psychological Review. 2006 Jan;113(2):300–326.

[12] Nakano K. Neural circuits and topographic organization of the basal ganglia and related regions. Brain Dev. 2000 Sep;22 Suppl 1:S5–16.

[13] Gurney. . . KN, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biol Cybern. 2001 Jun;84(6):401–10.

[14] Gurney. . . KN, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biol Cybern. 2001 Jun;84(6):411–23.

[15] Sesack S, Carr D. Selective prefrontal cortex inputs to dopamine cells: implications for schizophrenia. Physiology and Behaviour. 2002;77:513–517.

[16] Prescott TJ, Bryson JJ, Seth AK. Introduction. Modelling natural action selection. Philosophical Transactions of the Royal Society B: Biological Sciences. 2007 Apr;362(1485):1521–1529.

[17] Krain A, Wilson A, Arbuckle R, Castellanos F, Milham M. Abstract Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. Neuroimage. 2006;32:477–484.

[18] Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. J Neurosci. 2007 Aug;27(31):8161–5.

[19] Lee D, Rushworth MFS, Walton ME, Watanabe M, Sakagami M. Functional specialization of the primate frontal cortex during decision making. J Neurosci. 2007 Aug;27(31):8170–3.

[20] Snaith S, Holland O. An investigation of two mediation strategies suitable for behavioural control in animals and animats. In: Meyer JA, SW W, editors. From Animals to Animats 3: Proceedings of the First International Conference on the Simulation of Adaptive Behaviour. Cambridge, MA: MIT Press;. p. 255–262.

[21] Maes P. Modelling adaptive automonous agents. In: Langton E, editor. Artificial Life: An Overview. Cambridge, MA: MIT Press; 1995. p. 135–162.

[22] Gillies A, Arbuthnott G. Computational models of the basal ganglia. Mov Disord. 2000;15:762–770.

[23] McFarland D. Problems of Animal Behaviour. London: Longman Scientific and Technical; 1989.

[24] Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. Progress in Neurobiology. 1996 Nov;50(4):381–425.

[25] Redgrave P, Prescott TJ, Gurney K. The basal ganglia: a vertebrate solution to the selection problem? Neuroscience. 1999;89(4):1009–1023.

[26] Graybiel A, Kimura M. Adaptive neural models in the basal ganglia. In: Houk J, Davis D JL andBeiser, editors. Models of Information Processing in the Basal Ganglia. Cambridge MA: MIT Press; 1995. p. 103–116.

[27] Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci. 1986 Jan;9:357–81.

[28] Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biol Cybern. 2001;84:411–423.

[29] Rubia K, Russell T, Overmeyer S, Brammer MJ, Bullmore ET, Sharma T, et al. Mapping motor inhibition: conjunctive brain activations across different versions of go/no-go and stop tasks. Neuroimage. 2001 Feb;13(2):250–61.

[30] Rubia K, Russell T, Bullmore ET, Soni W, Brammer MJ, Simmons A, et al. An fMRI study of reduced left prefrontal activation in schizophrenia during normal inhibitory function. Schizophr Res. 2001 Oct;52(1-2):47–55.

[31] Kelly AMC, Hester R, Murphy K, Javitt DC, Foxe JJ, Garavan H. Prefrontal-subcortical dissociations underlying inhibitory control revealed by event-related fMRI. Eur J Neurosci. 2004 Jun;19(11):3105–12.

[32] Aron AR, Poldrack RA. Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus. J Neurosci. 2006 Mar;26(9):2424–33.

[33] van Os J, Kapur S. Schizophrenia. Lancet. 2009 Aug;374(9690):635–45.

[34] Kety SS. The syndrome of schizophrenia: unresolved questions and opportunities for research. The British Journal of Psychiatry. 1980 May;136:421–36.

[35] Hoenig J. The concept of Schizophrenia. Kraepelin-Bleuler-Schneider. The British Journal of Psychiatry. 1983 Jun;142:547–56.

[36] Andreasen NC. The evolving concept of schizophrenia: from Kraepelin to the present and future. Schizophr Res. 1997 Dec;28(2-3):105–9.

[37] Adityanjee, Aderibigbe YA, Theodoridis D, Vieweg VR. Dementia praecox to schizophrenia: the first 100 years. Psychiatry Clin Neurosci. 1999 Aug;53(4):437–48.

[38] Boyle M. Schizophrenia: A Scientific Delusion? Routledge, New York; 2002.

[39] Johnstone EC. 13: Schizophrenia. In: Johnstone EC, Freeman CPL, Zealley AK, editors. Companion to Psychiatric Studies. Churchill Livingston; 1998. .

[40] Cranefield P. A seventeenth century view of mental deficiency and schizophrenia: Thomas Willis on "Stupidity or foolishness". Bulletin of the History of Medicine. 1961;35:291–316.

[41] Jelliffe S. Dementia praecox: An historical summary. Int J Psycho-Anal. 1910;91:521–531.

[42] Bark NM. On the history of schizophrenia. Evidence of its existence before 1800. New York state journal of medicine. 1988 Jul;88(7):374–83.

[43] Wender P. Dementia praecox: the development of the concept. The American journal of psychiatry. 1963 Jun;119:1143–51.

[44] Sedler MJ. The legacy of Ewald Hecker: a new translation of "Die Hebephrenie". Translated by Marie-Louise Schoelly. The American journal of psychiatry. 1985 Nov;142(11):1265–71.

[45] Barnes MP, Saunders M, Walls TJ, Saunders I, Kirk CA. The syndrome of Karl Ludwig Kahlbaum. Journal of Neurology, Neurosurgery & Psychiatry. 1986 Sep;49(9):991–6.

[46] Odegård O. Changes in the prognosis of functional psychoses since the days of Kraepelin. The forty-first Maudsley Lecture, delivered before the Royal Medico-Psychological Association, 18th November, 1966. The British Journal of Psychiatry. 1967 Aug;113(501):813–22.

[47] Kraepelin E. Psychiatrie: Ein Lehrbuch fur Studirende und Aerzte. 4th ed. Barth, Leipzig; 1893.

[48] Kraepelin E. Dementia praecox and paraphrenia (Translated by RM Barclay). Livingstone, Edinburgh; 1919.

[49] Kraepelin E. One hundred years of psychiatry (Translated by W Baskin). Philosophical Library, New York; 1962.

[50] Kendler KS. Kraepelin and the differential diagnosis of dementia praecox and manic-depressive insanity. Comprehensive psychiatry. 1986 Jan;27(6):549–58.

[51] Bleuler E. Dementia praecox or the group of schizophrenias (Translated by J Zinkin). International University Press, New York; 1950.

[52] Mellor CS. First rank symptoms of schizophrenia. I. The frequnncy in schizophrenics on admission to hospital. II. Differences between individual first rank symptoms. The British Journal of Psychiatry. 1970 Jul;117(536):15–23.

[53] van Praag HM. About the impossible concept of schizophrenia. Comprehensive psychiatry. 1976 Jan;17(4):481–97.

[54] Schneider K. Wesen und Erfassung des Schizophrenen. Zeitschrift fűr die gesamte Neurololologie und Psychiatrie. 1925;99:542–547.

[55] Association AP. Diagnostic and Statistical Manual of Mental Disorders. DSM-IV. 4th ed. American Psychiatric Association, Washington (DC); 1994.

[56] Trimble MR. First-rank symptoms of Schneider. A new perspective? The British Journal of Psychiatry. 1990 Feb;156:195–200.

[57] Nordgaard J, Arnfred SM, Handest P, Parnas J. The diagnostic status of first-rank symptoms. Schizophrenia Bulletin. 2008 Jan;34(1):137–54.

[58] Griesinger W. Mental Pathology and Therapeutics. London; 1867.

[59] Alzheimer A. Beitraege zur pathologischen Anatomie der Hirnrinde und zur anatomischen Grundlage einiger Psychosen. [Contributions to the pathological anatomy of the cerebral cortex and the anatomical basis of some psychoses]. Monatsschrift fuer Psychiatrie und Neurologie. 1897;2(82).

[60] Alzheimer A. Beitraege zur pathologischen Anatomie der Dementia praecox. [Contributions on the pathological anatomy of the dementia praecox]. Zeitschrift fűr die gesamte Neurololologie und Psychiatrie. 1913;7.

[61] Spielmeyer W. Alzheimer's Lebenswerk. Zeitschrift fűr die gesamte Neurololologie und Psychiatrie. 1916;33:1–44.

[62] Corsellis J. Psychoses of Obscure Pathology. In: Blackwood W, Corsellis J, editors. Greenfield's Neuropathology. 3rd ed. Edward Arnold, London; 1976. .

[63] Cooper B. Nature, nurture and mental disorder: old concepts in the new millennium. The British journal of psychiatry Supplement. 2001 Apr;40:s91–101.

[64] Wiener N. Cybernetics. The MIT Press; 1948.

[65] Sakel M. The classical Sakel shock treatment: a reappraisal. Journal of clinical and experimental psychopathology. 1954 Jan;15(3):255–316.

[66] Fink M. Meduna and the origins of convulsive therapy. The American journal of psychiatry. 1984 Sep;141(9):1034–41.

[67] Owens DGC. Advances in psychopharmacology–schizophrenia. British Medical Bulletin. 1996 Jul;52(3):556–74.

[68] Delay J, Deniker P, Harl J. [Therapeutic method derived from hiberno-therapy in excitation and agitation states.]. Annales médico-psychologiques. 1952 Jul;110(2:2):267–73.

[69] Hollister L, TRAUB L, Prusmack J. Use of thioridazine for intensive treatment of schizophrenics refractory to other tranquilizing drugs. Journal of neuropsychiatry. 1960 Apr;1:200–4.

[70] Manceaux A, Bardenat C, Pelicier Y, Pascalis G. [Findings on pneumoencephalography in dementia praecox.]. Annales médico-psychologiques. 1957 May;115(5):828–36.

[71] Johnstone EC, Crow TJ, Frith CD, Husband J, Kreel L. Cerebral ventricular size and cognitive impairment in chronic schizophrenia. Lancet. 1976 Oct;2(7992):924–6.

[72] Daniel DG, Goldberg TE, Gibbons RD, Weinberger DR. Lack of a bimodal distribution of ventricular size in schizophrenia: a Gaussian mixture analysis of 1056 cases and controls. Biol Psychiatry. 1991 Nov;30(9):887–903.

[73] Horn JDV, McManus IC. Ventricular enlargement in schizophrenia. A meta-analysis of studies of the ventricle:brain ratio (VBR). The British Journal of Psychiatry. 1992 May;160:687–97.

[74] Lawrie SM, Abukmeil SS. Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. The British Journal of Psychiatry. 1998 Feb;172:110–20.

[75] Degreef G, Ashtari M, Bogerts B, Bilder RM, Jody DN, Alvir JM, et al. Volumes of ventricular system subdivisions measured from magnetic resonance images in first-episode schizophrenic patients. Arch Gen Psychiatry. 1992 Jul;49(7):531–7.

[76] Lim KO, Tew W, Kushner M, Chow K, Matsumoto B, Delisi LE. Cortical gray matter volume deficit in patients with first-episode schizophrenia. The American journal of psychiatry. 1996 Dec;153(12):1548–53.

[77] Gur RE, Cowell P, Turetsky BI, Gallacher F, Cannon T, Bilker W, et al. A follow-up magnetic resonance imaging study of schizophrenia. Relationship of neuroanatomical changes to clinical and neurobehavioral measures. Arch Gen Psychiatry. 1998 Feb;55(2):145–52.

[78] Whitworth AB, Honeder M, Kremser C, Kemmler G, Felber S, Hausmann A, et al. Hippocampal volume reduction in male schizophrenic patients. Schizophr Res. 1998 May;31(2-3):73–81.

[79] Zipursky RB, Lambe EK, Kapur S, Mikulis DJ. Cerebral gray matter volume deficits in first episode psychosis. Arch Gen Psychiatry. 1998 Jun;55(6):540–6.

[80] Lawrie SM, Whalley HC, Kestelman JN, Abukmeil SS, Byrne M, Hodges A, et al. Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. Lancet. 1999 Jan;353(9146):30–3.

[81] Suddath RL, Christison GW, Torrey EF, Casanova MF, Weinberger DR. Anatomical abnormalities in the brains of monozygotic twins discordant for schizophrenia. N Engl J Med. 1990 Mar;322(12):789–94.

[82] Noga JT, Bartley AJ, Jones DW, Torrey EF, Weinberger DR. Cortical gyral anatomy and gross brain dimensions in monozygotic twins discordant for schizophrenia. Schizophr Res. 1996 Oct;22(1):27–40.

[83] Seeman P. Dopamine receptors and the dopamine hypothesis of schizophrenia. Synapse. 1987 Jan;1(2):133–52.

[84] Angrist BM, Gershon S. The phenomenology of experimentally induced amphetamine psychosis–preliminary observations. Biol Psychiatry. 1970 Apr;2(2):95–107.

[85] Angrist B, Sathananthan G, Wilk S, Gershon S. Amphetamine psychosis: behavioral and biochemical aspects. Journal of Psychiatric Research. 1974 Jan;11:13–23.

[86] Carlsson A, Lindqvist M. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normethanephrine in mouse brain. Acta pharmacologica et toxicologica. 1963 Jan;20:140–4.

[87] Seeman P, Lee T, Chau-Wong M, Wong K. Antipsychotic drug doses and neuroleptic/dopamine receptors. Nature. 1976 Jun;261(5562):717–9.

[88] Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III–the final common pathway. Schizophrenia Bulletin. 2009 May;35(3):549–62.

[89] Cohen B, ROSENBAUM G, Luby E, Gottlieb J. Comparison of phencyclidine hydrochloride (Sernyl) with other drugs. Simulation of schizophrenic performance with phencyclidine hydrochloride (Sernyl), lysergic acid diethylamide (LSD-25), and amobarbital (Amytal) sodium; II. Symbolic and sequential thinking. Arch Gen Psychiatry. 1962 May;6:395–401.

[90] Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch Gen Psychiatry. 1994 Mar;51(3):199–214.

[91] Allen RM, Young SJ. Phencyclidine-induced psychosis. The American journal of psychiatry. 1978 Sep;135(9):1081–4.

[92] Domino E, Luby E. Abnormal mental states induced by phencyclidine as a model of schizophrenia. In: Domino E, editor. PCP (Phencyclidine): Historical and Current Perspectives). NPP Books, Ann Arbor, Mich.; 1981. p. 401–418.

[93] Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. Brain. 1999 Apr;122 (Pt 4):593–624.

[94] Jakob H, Beckmann H. Prenatal developmental disturbances in the limbic allocortex in schizophrenics. Journal of neural transmission (Vienna, Austria : 1996). 1986 Jan;65(3-4):303–26.

[95] Jakob H, Beckmann H. Gross and histological criteria for developmental disorders in brains of schizophrenics. Journal of the Royal Society of Medicine. 1989 Aug;82(8):466–9.

[96] Arnold SE, Hyman BT, Hoesen GWV, Damasio AR. Some cytoarchitectural abnormalities of the entorhinal cortex in schizophrenia. Arch Gen Psychiatry. 1991 Jul;48(7):625–32.

[97] Akil M, Lewis DA. Cytoarchitecture of the entorhinal cortex in schizophrenia. The American journal of psychiatry. 1997 Jul;154(7):1010–2.

[98] Krimer LS, Herman MM, Saunders RC, Boyd JC, Hyde TM, Carter JM, et al. A qualitative and quantitative analysis of the entorhinal cortex in schizophrenia. Cereb Cortex. 1997 Dec;7(8):732–9.

[99] Kovelman JA, Scheibel AB. A neurohistological correlate of schizophrenia. Biol Psychiatry. 1984 Dec;19(12):1601–21.

[100] Altshuler LL, Conrad A, Kovelman JA, Scheibel A. Hippocampal pyramidal cell orientation in schizophrenia. A controlled neurohistologic study of the Yakovlev collection. Arch Gen Psychiatry. 1987 Dec;44(12):1094–8.

[101] Conrad AJ, Abebe T, Austin R, Forsythe S, Scheibel AB. Hippocampal pyramidal cell disarray in schizophrenia as a bilateral phenomenon. Arch Gen Psychiatry. 1991 May;48(5):413–7.

[102] Jönsson SA, Luts A, Guldberg-Kjaer N, Brun A. Hippocampal pyramidal cell disarray correlates negatively to cell number: implications for the pathogenesis of schizophrenia. European archives of psychiatry and clinical neuroscience. 1997 Jan;247(3):120–7.

[103] Zaidel DW, Esiri MM, Harrison PJ. Size, shape, and orientation of neurons in the left and right hippocampus: investigation of normal asymmetries and alterations in schizophrenia. The American journal of psychiatry. 1997 Jun;154(6):812–8.

[104] Christison GW, Casanova MF, Weinberger DR, Rawlings R, Kleinman JE. A quantitative investigation of hippocampal pyramidal cell size, shape, and variability of orientation in schizophrenia. Arch Gen Psychiatry. 1989 Nov;46(11):1027–32.

[105] Benes FM, Sorensen I, Bird ED. Reduced neuronal size in posterior hippocampus of schizophrenic patients. Schizophrenia Bulletin. 1991 Jan;17(4):597–608.

[106] Arnold SE, Franz BR, Gur RC, Gur RE, Shapiro RM, Moberg PJ, et al. Smaller neuron size in schizophrenia in hippocampal subfields that mediate cortical-hippocampal interactions. The American journal of psychiatry. 1995 May;152(5):738–48.

[107] Jeste DV, Lohr JB. Hippocampal pathologic findings in schizophrenia. A morphometric study. Arch Gen Psychiatry. 1989 Nov;46(11):1019–24.

[108] Falkai P, Bogerts B. Cell loss in the hippocampus of schizophrenics. European archives of psychiatry and neurological sciences. 1986 Jan;236(3):154–61.

[109] Benes FM, McSparren J, Bird ED, SanGiovanni JP, Vincent SL. Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. Arch Gen Psychiatry. 1991 Nov;48(11):996–1001.

[110] Rajkowska G, Selemon LD, Goldman-Rakic PS. Neuronal and glial somal size in the prefrontal cortex: a postmortem morphometric study of schizophrenia and Huntington disease. Arch Gen Psychiatry. 1998 Mar;55(3):215–24.

[111] Pakkenberg B. Total nerve cell number in neocortex in chronic schizophrenics and controls estimated using optical disectors. Biol Psychiatry. 1993 Dec;34(11):768–72.

[112] Selemon LD, Rajkowska G, Goldman-Rakic PS. Abnormally high neuronal density in the schizophrenic cortex. A morphometric analysis of prefrontal area 9 and occipital area 17. Arch Gen Psychiatry. 1995 Oct;52(10):805–18; discussion 819–20.

[113] Selemon LD, Rajkowska G, Goldman-Rakic PS. Elevated neuronal density in prefrontal area 46 in brains from schizophrenic patients: application of a three-dimensional, stereologic counting method. J Comp Neurol. 1998 Mar;392(3):402–12.

[114] Schlaug G, Armstrong E, Schleicher A, Zilles K. Layer V pyramidal cells in the adult human cingulate cortex. A quantitative Golgi-study. Anat Embryol. 1993 Jun;187(6):515–22.

[115] Garey LJ, Ong WY, Patel TS, Kanani M, Davis A, Mortimer AM, et al. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. Journal of Neurology, Neurosurgery & Psychiatry. 1998 Oct;65(4):446–53.

[116] Glantz LA, Lewis DA. Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. Arch Gen Psychiatry. 2000 Jan;57(1):65–73.

[117] Kalus P, Müller TJ, Zuschratter W, Senitz D. The dendritic architecture of prefrontal pyramidal neurons in schizophrenic patients. Neuroreport. 2000 Nov;11(16):3621–5.

[118] Broadbelt K, Byne W, Jones LB. Evidence for a decrease in basilar dendrites of pyramidal cells in schizophrenic medial prefrontal cortex. Schizophrenia Research. 2002 Nov;58(1):75–81.

[119] Black JE, Kodish IM, Grossman AW, Klintsova AY, Orlovskaya D, Vostrikov V, et al. Pathology of layer V pyramidal neurons in the prefrontal cortex of patients with schizophrenia. The American journal of psychiatry. 2004 Apr;161(4):742–4.

[120] Kolluri N, Sun Z, Sampson AR, Lewis DA. Lamina-specific reductions in dendritic spine density in the prefrontal cortex of subjects with schizophrenia. The American journal of psychiatry. 2005 Jun;162(6):1200–2.

[121] Levitt JB, Lewis DA, Yoshioka T, Lund JS. Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 and 46). J Comp Neurol. 1993 Dec;338(3):360–76.

[122] Pucak ML, Levitt JB, Lund JS, Lewis DA. Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. J Comp Neurol. 1996 Dec;376(4):614–30.

[123] Erickson SL, Lewis DA. Cortical connections of the lateral mediodorsal thalamus in cynomolgus monkeys. J Comp Neurol. 2004 May;473(1):107–27.

[124] Arnold SE, Lee VM, Gur RE, Trojanowski JQ. Abnormal expression of two microtubule-associated proteins (MAP2 and MAP5) in specific subfields of the hippocampal formation in schizophrenia. Proc Natl Acad Sci USA. 1991 Dec;88(23):10850–4.

[125] Browning MD, Dudek EM, Rapier JL, Leonard S, Freedman R. Significant reductions in synapsin but not synaptophysin specific activity in the brains of some schizophrenics. Biol Psychiatry. 1993 Oct;34(8):529–35.

[126] Eastwood SL, Burnet PW, Harrison PJ. Altered synaptophysin expression as a marker of synaptic pathology in schizophrenia. Neuroscience. 1995 May;66(2):309–19.

[127] Eastwood SL, Harrison PJ. Decreased synaptophysin in the medial temporal lobe in schizophrenia demonstrated using immunoautoradiography. Neuroscience. 1995 Nov;69(2):339–43.

[128] Harrison PJ, Eastwood SL. Preferential involvement of excitatory neurons in medial temporal lobe in schizophrenia. Lancet. 1998 Nov;352(9141):1669–73.

[129] Young CE, Arima K, Xie J, Hu L, Beach TG, Falkai P, et al. SNAP-25 deficit and hippocampal connectivity in schizophrenia. Cereb Cortex. 1998 Jan;8(3):261–8.

[130] Kolomeets NS, Orlovskaya DD, Uranova NA. Decreased numerical density of CA3 hippocampal mossy fiber synapses in schizophrenia. Synapse. 2007 Aug;61(8):615–21.

[131] Perrone-Bizzozero NI, Sower AC, Bird ED, Benowitz LI, Ivins KJ, Neve RL. Levels of the growth-associated protein GAP-43 are selectively increased in association cortices in schizophrenia. Proc Natl Acad Sci USA. 1996 Nov;93(24):14182–7.

[132] Glantz LA, Lewis DA. Reduction of synaptophysin immunoreactivity in the prefrontal cortex of subjects with schizophrenia. Regional and diagnostic specificity. Arch Gen Psychiatry. 1997 Jul;54(7):660–9.

[133] Thompson PM, Sower AC, Perrone-Bizzozero NI. Altered levels of the synaptosomal associated protein SNAP-25 in schizophrenia. Biol Psychiatry. 1998 Feb;43(4):239–43.

[134] Woo TU, Whitehead RE, Melchitzky DS, Lewis DA. A subclass of prefrontal gamma-aminobutyric acid axon terminals are selectively altered in schizophrenia. Proc Natl Acad Sci USA. 1998 Apr;95(9):5341–6.

[135] Kerwin R, Patel S, Meldrum B. Quantitative autoradiographic analysis of glutamate binding sites in the hippocampal formation in normal and schizophrenic brain post mortem. Neuroscience. 1990 Jan;39(1):25–32.

[136] Eastwood SL, McDonald B, Burnet PW, Beckwith JP, Kerwin RW, Harrison PJ. Decreased expression of mRNAs encoding non-NMDA glutamate receptors GluR1 and GluR2 in medial temporal lobe neurons in schizophrenia. Brain Res Mol Brain Res. 1995 Apr;29(2):211–23.

[137] Eastwood SL, Kerwin RW, Harrison PJ. Immunoautoradiographic evidence for a loss of alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate-preferring non-N-methyl-D-aspartate glutamate receptors within the medial temporal lobe in schizophrenia. Biol Psychiatry. 1997 Mar;41(6):636–43.

[138] Porter RH, Eastwood SL, Harrison PJ. Distribution of kainate receptor subunit mRNAs in human hippocampus, neocortex and cerebellum, and bilateral reduction of hippocampal GluR6 and KA2 transcripts in schizophrenia. Brain Res. 1997 Mar;751(2):217–31.

[139] Kim JS, Kornhuber HH, Schmid-Burgk W, Holzmüller B. Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. Neuroscience Letters. 1980 Dec;20(3):379–82.

[140] Perry TL. Normal cerebrospinal fluid and brain glutamate levels in schizophrenia do not support the hypothesis of glutamatergic neuronal dysfunction. Neuroscience Letters. 1982 Jan;28(1):81–5.

[141] Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. The American journal of psychiatry. 1991 Oct;148(10):1301–8.

[142] Bauer D, Gupta D, Harotunian V, Meador-Woodruff JH, Mccullumsmith RE. Abnormal expression of glutamate transporter and transporter interacting molecules in prefrontal cortex in elderly patients with schizophrenia. Schizophrenia Research. 2008 Sep;104(1-3):108–20.

[143] Bauer D, Haroutunian V, Meador-Woodruff J, McCullumsmith R. Abnormal glycosylation of EAAT1 and EAAT2 in prefrontal cortex of elderly patients with schizophrenia. Schizophrenia Research. 2009 Aug;.

[144] Dracheva S, Marras SA, Elhakem SL, Kramer FR, Davis KL, Haroutunian V. N-methyl-D-aspartic acid receptor expression in the dorsolateral prefrontal cortex of elderly patients with schizophrenia. The American journal of psychiatry. 2001 Sep;158(9):1400–10.

[145] Akbarian S, Sucher NJ, Bradley D, Tafazzoli A, Trinh D, Hetrick WP, et al. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. J Neurosci. 1996 Jan;16(1):19–30.

[146] Beneyto M, Meador-Woodruff JH. Lamina-specific abnormalities of NMDA receptor-associated postsynaptic protein transcripts in the prefrontal cortex in schizophrenia and bipolar disorder. Neuropsychopharmacology. 2008 Aug;33(9):2175–86.

[147] Dracheva S, Byne W, Chin B, Haroutunian V. Ionotropic glutamate receptor mRNA expression in the human thalamus: absence of change in schizophrenia. Brain Research. 2008 Jun;1214:23–34.

[148] Kristiansen L, Huerta I, Beneyto M, Meadorwoodruff J. NMDA receptors and schizophrenia. Current Opinion in Pharmacology. 2007 Feb;7(1):48–55.

[149] Lewis DA, González-Burgos G. Neuroplasticity of neocortical circuits in schizophrenia. Neuropsychopharmacology. 2008 Jan;33(1):141–65.

[150] Olney JW, Labruyere J, Price MT. Pathological changes induced in cerebrocortical neurons by phencyclidine and related drugs. Science. 1989 Jun;244(4910):1360–2.

[151] Hargreaves RJ, Rigby M, Smith D, Hill RG, Iversen LL. Competitive as well as uncompetitive N-methyl-D-aspartate receptor antagonists affect cortical neuronal morphology and cerebral glucose metabolism. Neurochem Res. 1993 Dec;18(12):1263–9.

[152] Allen HL, Iversen LL. Phencyclidine, dizocilpine, and cerebrocortical neurons. Science. 1990 Jan;247(4939):221.

[153] Fix AS, Horn JW, Wightman KA, Johnson CA, Long GG, Storts RW, et al. Neuronal vacuolization and necrosis induced by the noncompetitive N-methyl-D-aspartate (NMDA) antagonist MK(+)801 (dizocilpine maleate): a light and electron microscopic evaluation of the rat retrosplenial cortex. Exp Neurol. 1993 Oct;123(2):204–15.

[154] Olney JW, Labruyere J, Wang G, Wozniak DF, Price MT, Sesma MA. NMDA antagonist neurotoxicity: mechanism and prevention. Science. 1991 Dec;254(5037):1515–8.

[155] Olney JW, Farber NB. Glutamate receptor dysfunction and schizophrenia. Arch Gen Psychiatry. 1995 Dec;52(12):998–1007.

[156] van Elst LT, Valerius G, Büchert M, Thiel T, Rüsch N, Bubl E, et al. Increased prefrontal and hippocampal glutamate concentration in schizophrenia: evidence from a magnetic resonance spectroscopy study. Biol Psychiatry. 2005 Nov;58(9):724–30.

[157] Stone J, Day F, Tsagaraki H, Valli I, McLean M, Lythgoe D, et al. Glutamate Dysfunction in People with Prodromal Symptoms of Psychosis: Relationship to Gray Matter Volume. Biol Psychiatry. 2009 Jun;.

[158] Akbarian S, Kim JJ, Potkin SG, Hagman JO, Tafazzoli A, Bunney WE, et al. Gene expression for glutamic acid decarboxylase is reduced without loss of neurons in prefrontal cortex of schizophrenics. Arch Gen Psychiatry. 1995 Apr;52(4):258–66.

[159] Benes FM, Lim B, Matzilevich D, Walsh JP, Subburaju S, Minns M. Regulation of the GABA cell phenotype in hippocampus of schizophrenics and bipolars. Proc Natl Acad Sci USA. 2007 Jun;104(24):10164–9.

[160] Durstewitz D, Seamans J. Beyond bistability: Biophysics and temporal dynamics of working memory. Neuroscience. 2006 Apr;139(1):119–133.

[161] Goldman-Rakic PS. Cellular basis of working memory. Neuron. 1995 Mar;14(3):477–85.

[162] Callicott JH, Mattay VS, Verchinski BA, Marenco S, Egan MF, Weinberger DR. Complexity of prefrontal cortical dysfunction in schizophrenia: more than up or down. The American journal of psychiatry. 2003 Dec;160(12):2209–15.

[163] Manoach DS. Prefrontal cortex dysfunction during working memory performance in schizophrenia: reconciling discrepant findings. Schizophrenia Research. 2003 Apr;60(2-3):285–98.

[164] Peters A, Proskauer CC, Ribak CE. Chandelier cells in rat visual cortex. J Comp Neurol. 1982 Apr;206(4):397–416.

[165] Klausberger T, Magill PJ, Márton LF, Roberts JDB, Cobden PM, Buzsáki G, et al. Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo. Nature. 2003 Feb;421(6925):844–8.

[166] Whittington MA, Traub RD. Interneuron diversity series: inhibitory interneurons and network oscillations in vitro. Trends in Neurosciences. 2003 Dec;26(12):676–82.

[167] Howard SK, Gaba DM, Smith BE, Weinger MB, Herndon C, Keshavacharya S, et al. Simulation study of rested versus sleep-deprived anesthesiologists. Anesthesiology. 2003 Jun;98(6):1345–55; discussion 5A.

[168] Cho RY, Konecky RO, Carter CS. Impairments in frontal cortical gamma synchrony and cognitive control in schizophrenia. Proc Natl Acad Sci USA. 2006 Dec;103(52):19878–83.

[169] Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. Nat Rev Neurosci. 2005 Apr;6(4):312–24.

[170] Stephan KE, Friston KJ, Frith CD. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. Schizophrenia Bulletin. 2009 May;35(3):509–27.

[171] Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry. 1987 Jul;44(7):660–9.

[172] Farber NB, Wozniak DF, Price MT, Labruyere J, Huss J, Peter HS, et al. Age-specific neurotoxicity in the rat associated with NMDA receptor blockade: potential relevance to schizophrenia? Biol Psychiatry. 1995 Dec;38(12):788–96.

[173] Karp HN, Kaufman ND, Anand SK. Phencyclidine poisoning in young children. J Pediatr. 1980 Dec;97(6):1006–9.

[174] Welch MJ, Correa GA. PCP intoxication in young children and infants. Clinical pediatrics. 1980 Aug;19(8):510–4.

[175] Marshall B, Longnecker D. General Anesthetics. In: Goodman L, Gilman A, Rall T, Nies A, Taylor P, editors. The Pharmacological Basis of Therapeutics. Pergamon Press Inc, Elmsform, NY; 1990. p. 285–310.

[176] Homayoun H, Moghaddam B. NMDA receptor hypofunction produces opposite effects on prefrontal cortex interneurons and pyramidal neurons. J Neurosci. 2007 Oct;27(43):11496–500.

[177] Reynolds GP, Czudek C, Andrews HB. Deficit and hemispheric asymmetry of GABA uptake sites in the hippocampus in schizophrenia. Biol Psychiatry. 1990 May;27(9):1038–44.

[178] Benes F. Altered glutamatergic and GABAergic mechanisms in the cingulate cortex of the schizophrenic brain. Arch Gen Psych. 1995;52:1015–1018.

[179] Durstewitz D, Seamans JK. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol Psychiatry. 2008 Nov;64(9):739–49.

[180] Carlsson A, Waters N, Carlsson ML. Neurotransmitter interactions in schizophrenia–therapeutic implications. Biol Psychiatry. 1999 Nov;46(10):1388–95.

[181] Flores C, Coyle JT. Regulation of glutamate carboxypeptidase II function in corticolimbic regions of rat brain by phencyclidine, haloperidol, and clozapine. Neuropsychopharmacology. 2003 Jul;28(7):1227–34.

[182] Kehrer C, Maziashvili N, Dugladze T, Gloveli T. Altered Excitatory-Inhibitory Balance in the NMDA-Hypofunction Model of Schizophrenia. Frontiers in molecular neuroscience. 2008 Jan;1:6.

[183] Campana A, Duci A, Gambini O, Scarone S. An artificial neural network that uses eye-tracking performance to identify patients with schizophrenia. Schizophrenia Bulletin. 1999 Jan;25(4):789–99.

[184] Trillenberg P, Lencer R, Heide W. Eye movements and psychiatric disease. Curr Opin Neurol. 2004 Feb;17(1):43–7.

[185] Lan TH, Loh EW, Wu MS, Hu TM, Chou P, Lan TY, et al. Performance of a neuro-fuzzy model in predicting weight changes of chronic schizophrenic patients exposed to antipsychotics. Mol Psychiatry. 2008 Dec;13(12):1129–37.

[186] Guo Y, Bowman FD, Kilts C. Predicting the brain response to treatment using a Bayesian hierarchical model with application to a study of schizophrenia. Hum Brain Mapp. 2008 Sep;29(9):1092–109.

[187] Fu W, Shen J, Luo X, Zhu W, Cheng J, Yu K, et al. Dopamine D1 Receptor Agonist and D2 Receptor Antagonist Effects of the Natural Product ( ) Stepholidine: Molecular Modeling and Dynamics Simulations. Biophysical Journal. 2007 May;93(5):1431–1441.

[188] Katritzky AR, Dobchev DA, Stoyanova-Slavova IB, Kuanar M, Bespalov MM, Karelson M, et al. Novel computational models for predicting dopamine interactions. Exp Neurol. 2008 May;211(1):150–71.

[189] Hsu PC, Yang UC, Shih KH, Liu CM, Liu YL, Hwu HG. A protein interaction based model for schizophrenia study. BMC Bioinformatics. 2008 Jan;9 Suppl 12:S23.

[190] Heeg B, Buskens E, Botteman M, Caleo S, Ingham M, Damen J, et al. The cost-effectiveness of atypicals in the UK. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2008 Dec;11(7):1007–21.

[191] Serretti A, Mandelli L, Bajo E, Cevenini N, Papili P, Mori E, et al. The socio-economical burden of schizophrenia: a simulation of cost-offset of early intervention program in Italy. Eur Psychiatry. 2009 Jan;24(1):11–6.

[192] Rolls ET, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. Nat Rev Neurosci. 2008 Sep;9(9):696–709.

[193] Hoffman RE. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. Arch Gen Psychiatry. 1987 Feb;44(2):178–88.

[194] Hoffman RE, Dobscha SK. Cortical pruning and the development of schizophrenia: a computer model. Schizophrenia Bulletin. 1989 Jan;15(3):477–90.

[195] Hoffman RE, McGlashan TH. Parallel distributed processing and the emergence of schizophrenic symptoms. Schizophrenia Bulletin. 1993 Jan;19(1):119–40.

[196] Hoffman RE, McGlashan TH. Synaptic elimination, neurodevelopment, and the mechanism of hallucinated "voices" in schizophrenia. The American journal of psychiatry. 1997 Dec;154(12):1683–9.

[197] Hoffman RE, Quinlan DM, Mazure CM, McGlashan TM. Cortical instability and the mechanism of mania: a neural network simulation and perceptual test. Biol Psychiatry. 2001 Mar;49(6):500–9.

[198] Hoffman RE, McGlashan TH. Neural network models of schizophrenia. The Neuroscientist. 2001 Oct;7(5):441–54.

[199] Siekmeier PJ, Hoffman RE. Enhanced semantic priming in schizophrenia: a computer model based on excessive pruning of local connections in association cortex. The British Journal of Psychiatry. 2002 Apr;180:345–50.

[200] Feinberg I. Schizophrenia: caused by a fault in programmed synaptic elimination during adolescence? Journal of Psychiatric Research. 1982 Jan;17(4):319–34.

[201] Miikkulainen R. Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory. Cambridge, Mass, MIT Press; 1993.

[202] Han SD, Nestor PG, Shenton ME, Niznikiewicz M, Hannah G, McCarley RW. Associative memory in chronic schizophrenia: a computational model. Schizophr Res. 2003 Jun;61(2-3):255–63.

[203] Horn D, Ruppin E. Compensatory mechanisms in an attractor neural network model of schizophrenia. Neural computation. 1995 Jan;7(1):182–205.

[204] Ruppin E. Neural modelling of psychiatric disorders. Network: Computation in Neural Systems. 1995;6(4):635–656.

[205] Ruppin E, Reggia JA, Horn D. Pathogenesis of schizophrenic delusions and hallucinations: a neural model. Schizophrenia Bulletin. 1996 Jan;22(1):105–23.

[206] Reggia J, Ruppin E, Berndt R. Computer Models: A New Approach to the Investigation of Disease. MD Computing: Computers in Medical Practice. 1997;14:160–168.

[207] Ruppin E, Reggia J. Seeking order in disorder: computational studies of neurologic and psychiatric diseases. Artificial Intelligence In Medicine. 1998;13(1-2):1–12.

[208] Greenstein-Messica A, Ruppin E. Synaptic runaway in associative networks and the pathogenesis of schizophrenia. Neural computation. 1998 Feb;10(2):451–65.

[209] Hall J, Romaniuk L, Mcintosh A, Steele JD, Johnstone EC, Lawrie S. Associative learning and the genetics of schizophrenia. Trends in Neurosciences. 2009 May;.

[210] Harris JM, Moorhead TWJ, Miller P, McIntosh AM, Bonnici HM, Owens DGC, et al. Increased prefrontal gyrification in a large high-risk cohort characterizes those who develop schizophrenia and reflects abnormal prefrontal development. Biol Psychiatry. 2007 Oct;62(7):722–9.

[211] Chen EY. A neural network model of cortical information processing in schizophrenia. I: Interaction between biological and social factors in symptom formation. Canadian journal of psychiatry Revue canadienne de psychiatrie. 1994 Oct;39(8):362–7.

[212] Chen EY. A neural network model of cortical information processing in schizophrenia. II–Role of hippocampal-cortical interaction: a review and a model. Canadian journal of psychiatry Revue canadienne de psychiatrie. 1995 Feb;40(1):21–6.

[213] Friston K. Theoretical neurobiology and schizophrenia. British Medical Bulletin. 1996;52(3):644–655.

[214] Koukkou M, Lehmann D, Wackermann J, Dvorak I, Henggeler B. Dimensional complexity of EEG brain mechanisms in untreated schizophrenia. Biol Psychiatry. 1993 Mar;33(6):397–407.

[215] Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb Cortex. 2000 Sep;10(9):910–23.

[216] Constantinidis C, Wang XJ. A neural circuit basis for spatial working memory. The Neuroscientist. 2004 Dec;10(6):553–65.

[217] Wang XJ, Tegnér J, Constantinidis C, Goldman-Rakic PS. Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. Proc Natl Acad Sci USA. 2004 Feb;101(5):1368–73.

[218] Wolf JA, Moyer JT, Lazarewicz MT, Contreras D, Benoit-Marand M, O'Donnell P, et al. NMDA/AMPA ratio impacts state transitions and entrainment to oscillations in a computational model of the nucleus accumbens medium spiny projection neuron. J Neurosci. 2005 Oct;25(40):9080–95.

[219] Loh M, Rolls ET, Deco G. A dynamical systems hypothesis of schizophrenia. PLoS Comput Biol. 2007 Nov;3(11):e228.

[220] Siekmeier PJ, Hasselmo ME, Howard MW, Coyle J. Modeling of context-dependent retrieval in hippocampal region CA1: implications for cognitive function in schizophrenia. Schizophr Res. 2007 Jan;89(1-3):177–90.

[221] Siekmeier PJ. Evidence of multistability in a realistic computer simulation of hippocampus subfield CA1. Behav Brain Res. 2009 Jun;200(1):220–31.

[222] Spencer KM. The functional consequences of cortical circuit abnormalities on gamma oscillations in schizophrenia: insights from computational modeling. Frontiers in human neuroscience. 2009 Jan;3:33.

[223] Singer W. Neuronal synchrony: a versatile code for the definition of relations? Neuron. 1999 Sep;24(1):49–65, 111–25.

[224] Bartos M, Vida I, Jonas P. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. Nat Rev Neurosci. 2007 Jan;8(1):45–56.

[225] Desimone R. Neuropsychology. Is dopamine a missing link? Nature. 1995 Aug;376(6541):549–50.

[226] Williams GV, Goldman-Rakic PS. Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. Nature. 1995 Aug;376(6541):572–5.

[227] Tanaka S. Dopaminergic control of working memory and its relevance to schizophrenia: a circuit dynamics perspective. Neuroscience. 2006 Apr;139(1):153–71.

[228] Tanaka S. Dysfunctional GABAergic inhibition in the prefrontal cortex leading to "psychotic" hyperactivation. BMC neuroscience. 2008 Jan;9:41.

[229] Lavigne F, Darmon N. Dopaminergic neuromodulation of semantic priming in a cortical network model. Neuropsychologia. 2008 Jul;.

[230] Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, et al. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. Proc Natl Acad Sci USA. 2001 Jun;98(12):6917–22.

[231] Winterer G, Musso F, Vucurevic G, Stoeter P, Konrad A, Seker B, et al. COMT genotype predicts BOLD signal and noise characteristics in prefrontal circuits. Neuroimage. 2006 Oct;32(4):1722–32.

[232] Servan-Schreiber D, Printz H, Cohen JD. A network model of catecholamine effects: gain, signal-to-noise ratio, and behavior. Science. 1990 Aug;249(4971):892–5.

[233] Cohen JD, Servan-Schreiber D. Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychological Review. 1992 Jan;99(1):45–77.

[234] Cohen JD, Braver TS, O'Reilly RC. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. Philos Trans R Soc Lond, B, Biol Sci. 1996 Oct;351(1346):1515–27.

[235] Servan-Schreiber D, Bruno RM, Carter CS, Cohen JD. Dopamine and the mechanisms of cognition: Part I. A neural network model predicting dopamine effects on selective attention. Biol Psychiatry. 1998 May;43(10):713–22.

[236] Servan-Schreiber D, Carter CS, Bruno RM, Cohen JD. Dopamine and the mechanisms of cognition: Part II. D-amphetamine effects in human subjects performing a selective attention task. Biol Psychiatry. 1998 May;43(10):723–9.

[237] Amos A. A computational model of information processing in the frontal cortex and basal ganglia. Journal of cognitive neuroscience. 2000 May;12(3):505–19.

[238] Monchi O, Taylor JG, Dagher A. A neural model of working memory processes in normal subjects, Parkinson's disease and schizophrenia for fMRI design and predictions. Neural networks : the official journal of the International Neural Network Society. 2000 Jan;13(8-9):953–73.

[239] Moxon KA, Gerhardt GA, Gulinello M, Adler LE. Inhibitory control of sensory gating in a computer model of the CA3 region of the hippocampus. Biol Cybern. 2003 Apr;88(4):247–64.

[240] Peled A, Geva AB. Brain organization and psychodynamics. The Journal of psychotherapy practice and research. 1999 Jan;8(1):24–39.

[241] Peled A, Geva AB. The perception of rorschach inkblots in schizophrenia: a neural network model. Int J Neurosci. 2000 Jan;104(1-4):49–61.

[242] Moore SC, Sellen JL. Jumping to conclusions: a network model predicts schizophrenic patients' performance on a probabilistic reasoning task. Cognitive, affective & behavioral neuroscience. 2006 Dec;6(4):261–9.

[243] Smith AJ, Li M, Becker S, Kapur S. Linking animal models of psychosis to computational models of dopamine function. Neuropsychopharmacology. 2007 Jan;32(1):54–66.

[244] Sutton RS, Barto AG. Toward a modern theory of adaptive networks: expectation and prediction. Psychological Review. 1981 Mar;88(2):135–70.

[245] Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997 Mar;275(5306):1593–9.

[246] Moore H, West AR, Grace AA. The regulation of forebrain dopamine transmission: relevance to the pathophysiology and psychopathology of schizophrenia. Biol Psychiatry. 1999 Jul;46(1):40–55.

[247] Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey GD, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry. 2008 Mar;13(3):239, 267–76.

[248] Frank MJ, Loughry B, O'Reilly RC. Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cognitive, affective & behavioral neuroscience. 2001 Jun;1(2):137–60.

[249] Frank M. Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. Neural Networks. 2006 Oct;19(8):1120–1136.

[250] Frank MJ, Scheres A, Sherman SJ. Understanding decision-making deficits in neurological conditions: insights from models of natural action selection. Philosophical Transactions of the Royal Society B: Biological Sciences. 2007 Apr;362(1485):1641–1654.

[251] Frank MJ. Schizophrenia: a computational reinforcement learning perspective. Schizophrenia Bulletin. 2008 Nov;34(6):1008–11.

[252] Parent A, Hazrati LN. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res Brain Res Rev. 1995 Jan;20(1):91–127.

[253] Parent A, Hazrati LN. Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. Brain Res Brain Res Rev. 1995 Jan;20(1):128–54.

[254] Haber SN. The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy. 2003 Dec;26(4):317–30.

[255] Haber SN, McFarland NR. The place of the thalamus in frontal cortical-basal ganglia circuits. The Neuroscientist. 2001 Aug;7(4):315–24.

[256] Gerfen CR. The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. Nature. 1984 Jan;311(5985):461–4.

[257] Graybiel AM. Neurotransmitters and neuromodulators in the basal ganglia. Trends in Neurosciences. 1990 Jul;13(7):244–54.

[258] Canales JJ. Stimulant-induced adaptations in neostriatal matrix and striosome systems: transiting from instrumental responding to habitual behavior in drug addiction. Neurobiol Learn Mem. 2005 Mar;83(2):93–103.

[259] Kemp JM, Powell TP. The structure of the caudate nucleus of the cat: light and electron microscopy. Philos Trans R Soc Lond, B, Biol Sci. 1971 Sep;262(845):383–401.

[260] Tepper JM, Koós T, Wilson CJ. GABAergic microcircuits in the neostriatum. Trends in Neurosciences. 2004 Nov;27(11):662–9.

[261] Romanelli P, Esposito V, Schaal D, Heit G. Somatotopy in the basal ganglia: experimental and clinical evidence for segregated sensorimotor channels. Brain Research Reviews. 2005 Feb;48(1):112–128.

[262] Gerfen CR, Engber TM, Mahan LC, Susel Z, Chase TN, Monsma FJ, et al. D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. Science. 1990 Dec;250(4986):1429–32.

[263] Akkal D, Burbaud P, Audin J, Bioulac B. Responses of substantia nigra pars reticulata neurons to intrastriatal D1 and D2 dopaminergic agonist injections in the rat. Neurosci Lett. 1996 Jul;213(1):66–70.

[264] Monakow KH, Akert K, Künzle H. Projections of the precentral motor cortex and other cortical areas of the frontal lobe to the subthalamic nucleus in the monkey. Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale. 1978 Nov;33(3-4):395–403.

[265] Carpenter MB, Carleton SC, Keller JT, Conte P. Connections of the subthalamic nucleus in the monkey. Brain Research. 1981 Nov;224(1):1–29.

[266] Nambu A, Tokuno H, Takada M. Functional significance of the cortico-subthalamopallidal 'hyperdirect' pathway. Neurosci Res. 2002 Jun;43(2):111–7.

[267] Bar-Gad I, Morris G, Bergman H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Progress in Neurobiology. 2003 Dec;71(6):439–73.

[268] Grillner S, Hellgren J, Ménard A, Saitoh K, Wikström MA. Mechanisms for selection of basic motor programs–roles for the striatum and pallidum. Trends in Neurosciences. 2005 Jul;28(7):364–70.

[269] Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci. 1986;9(0147-006X (Print)):357–81.

[270] Yelnik J. Modeling the organization of the basal ganglia. Rev Neurol (Paris). 2008 Dec;164(12):969–76.

[271] Doyon J. Motor sequence learning and movement disorders. Curr Opin Neurol. 2008 Aug;21(4):478–83.

[272] Kotz S, Schwartze M, Schmidt-Kassow M. Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. Cortex; a journal devoted to the study of the nervous system and behavior. 2009 Mar;.

[273] Grahn JA, Parkinson JA, Owen AM. The role of the basal ganglia in learning and memory: neuropsychological studies. Behav Brain Res. 2009 Apr;199(1):53–60.

[274] Grillner S, Wallén P, Saitoh K, Kozlov A, Robertson B. Neural bases of goal-directed locomotion in vertebrates–an overview. Brain Research Reviews. 2008 Jan;57(1):2–12.

[275] Prodoehl J, Corcos DM, Vaillancourt DE. Basal ganglia mechanisms underlying precision grip force control. Neuroscience and biobehavioral reviews. 2009 Jun;33(6):900–8.

[276] Dormont JF, Condé H, Cheruel F, Farin D. Correlations between activity of pallidal neurons and motor parameters. Somatosens Mot Res. 1997 Jan;14(4):281–94.

[277] Grahn JA. The role of the basal ganglia in beat perception: neuroimaging and neuropsychological investigations. Ann N Y Acad Sci. 2009 Jul;1169:35–45.

[278] Meck WH, Penney TB, Pouthas V. Cortico-striatal representation of time in animals and humans. Current Opinion in Neurobiology. 2008 Apr;18(2):145–52.

[279] Schultz W, Romo R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. Journal of Neurophysiology. 1990;63(3):607–624.

[280] Sehlmeyer C, Schöning S, Zwitserlood P, Pfleiderer B, Kircher T, Arolt V, et al. Human fear conditioning and extinction in neuroimaging: a systematic review. PLoS ONE. 2009 Jan;4(6):e5865.

[281] Mishkin M, Petri H. Memories and habits: Some implications for the analysis of learning and retention. In: Squire L, Butters N, editors. Neuropsychology of memory. New York: Guilford Press; 1984. p. 287–296.

[282] Horvitz JC. Stimulus-response and response-outcome learning mechanisms in the striatum. Behav Brain Res. 2009 Apr;199(1):129–40.

[283] Yacubian J, Sommer T, Schroeder K, Gläscher J, Braus DF, Büchel C. Subregions of the ventral striatum show preferential coding of reward magnitude and probability. NeuroImage. 2007 Nov;38(3):557–63.

[284] Rolls ET, McCabe C, Redoute J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb Cortex. 2008 Mar;18(3):652–63.

[285] Brown RG, Marsden CD. Internal versus external cues and the control of attention in Parkinson's disease. Brain. 1988 Apr;111 (Pt 2):323–45.

[286] Frith CD, Bloxham CA, Carpenter KN. Impairments in the learning and performance of a new manual skill in patients with Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry. 1986 Jun;49(6):661–8.

[287] Saint-Cyr JA. Frontal-striatal circuit functions: context, sequence, and consequence. Journal of the International Neuropsychological Society : JINS. 2003 Jan;9(1):103–27.

[288] Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neuroscience and biobehavioral reviews. 2008 Jan;32(2):265–78.

[289] Robbins TW. Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. Philos Trans R Soc Lond, B, Biol Sci. 2007 May;362(1481):917–32.

[290] Dalley JW, Mar AC, Economidou D, Robbins TW. Neurobehavioral mechanisms of impulsivity: fronto-striatal systems and functional neurochemistry. Pharmacol Biochem Behav. 2008 Aug;90(2):250–60.

[291] Samejima K, Doya K. Multiple representations of belief states and action values in corticobasal ganglia loops. Annals of the New York Academy of Sciences. 2007 May;1104:213–28.

[292] Daw ND, Doya K. The computational neurobiology of learning and reward. Current Opinion in Neurobiology. 2006 Apr;16(2):199–204.

[293] Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. Psychological Review. 1998 Jul;105(3):442–81.

[294] Nakahara H, Doya K, Hikosaka O. Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuomotor sequences - a computational approach. Journal of cognitive neuroscience. 2001 Jul;13(5):626–47.

[295] Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci. 2000 Mar;20(6):2369–82.

[296] Haruno M, Kawato M. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. Neural networks : the official journal of the International Neural Network Society. 2006 Oct;19(8):1242–54.

[297] Shah A, Barto AG. Effect on movement selection of an evolving sensory representation: a multiple controller model of skill acquisition. Brain Research. 2009 Nov;1299:55–73.

[298] Houk JC, Bastianen C, Fansler D, Fishbach A, Fraser D, Reber PJ, et al. Action selection and refinement in subcortical loops through basal ganglia and cerebellum. Philos Trans R Soc Lond, B, Biol Sci. 2007 Sep;362(1485):1573–83.

[299] Lo CC, Wang XJ. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci. 2006 Jul;9(7):956–63.

[300] Rubchinsky LL, Kopell N, Sigvardt KA. Modeling facilitation and inhibition of competing motor programs in basal ganglia subthalamic nucleus-pallidal circuits. Proc Natl Acad Sci USA. 2003 Nov;100(24):14427–32.

[301] Brown JW, Bullock D, Grossberg S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. Neural networks : the official journal of the International Neural Network Society. 2004 May;17(4):471–510.

[302] Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. Journal of cognitive neuroscience. 2005 Jan;17(1):51–72.

[303] Hazy TE, Frank MJ, O'reilly RC. Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. Philos Trans R Soc Lond, B, Biol Sci. 2007 Sep;362(1485):1601–13.

[304] Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. Behav Brain Res. 2009 Apr;199(1):141–56.

[305] Molina-Vilaplana J, Contreras-Vidal J, Herrero-Ezquerro M, Lopez-Coronado J. A model for altered neural network dynamics related to prehension movements in Parkinson disease. Biological cybernetics. 2009 Feb;.

[306] Humphries MD, Gurney... KN. A pulsed neural network model of bursting in the basal ganglia. Neural networks : the official journal of the International Neural Network Society. 2001 Jan;14(6-7):845–63.

[307] Gurney... KN, Humphries MD, Wood R, Prescott TJ, Redgrave P. Testing computational hypotheses of brain systems function: a case study with the basal ganglia. Network: Computation in Neural Systems. 2004 Nov;15(4):263–290.

[308] Bogacz R, Gurney... KN. The basal ganglia and cortex implement optimal decision making between alternative actions. Neural computation. 2007 Feb;19(2):442–77.

[309] Redgrave P, Prescott TJ, Gurney... KN. The basal ganglia: a vertebrate solution to the selection problem? Neuroscience. 1999 Jan;89(4):1009–23.

[310] Liddle PF. The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy. The British Journal of Psychiatry. 1987 Aug;151:145–51.

[311] Bleuler E. Handbuch der Psychiatrie. Franz Deuticke, Leipzig und Wien; 1911.

[312] Buckley PF, Stahl SM. Pharmacological treatment of negative symptoms of schizophrenia: therapeutic opportunity or cul-de-sac? Acta Psychiatr Scand. 2007;115:93–100.

[313] Stratta P, Mancini F, Mattei P, Daneluzzo E, Casacchia M, Rossi A. Association between striatal reduction and poor Wisconsin card sorting test performancein patients with schizophrenia. Biol Psychiatry. 1997;42:816–820.

[314] Menon V, Anagnoson RT, Glover GH, Pfefferbaum A. Functional magnetic resonance imaging evidence for disrupted basal ganglia function in schizophrenia. Am J Psychiatry. 2001 Apr;158(4):646–649.

[315] von Neumann J, Morgerstern O. Theory of games and economic behavior. Princeton, NJ: Princeton UP; 1944.

[316] Hester R, Fassbender C, Garavan H. Individual differences in error processing: a review and reanalysis of three event-related fMRI studies using the GO/NOGO task. Cereb Cortex. 2004;14(9):986–94.

[317] Hester R. Individual Differences in Error Processing: A Review and Reanalysis of Three Event-related fMRI Studies Using the GO/NOGO Task. Cerebral Cortex. 2004 Apr;14(9):986–994.

[318] Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J Neurosci. 2007 Apr;27(14):3743–52.

[319] Stroop J. Studies of interference in serial verbal reactions. J Exp Psychol. 1935;18:643–662.

[320] Simon JR. Reactions toward the source of stimulation. J Exp Psychol. 1969 Jul;81(1):174–6.

[321] Eriksen B, Eriksen C. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception and Psychophysics. 1974;25:249–263.

[322] Casey BJ, Thomas KM, Welsh TF, Badgaiyan RD, Eccard CH, Jennings JR, et al. Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. Proc Natl Acad Sci USA. 2000 Jul;97(15):8728–33.

[323] Bunge S, Hazeltine E, Scanlon M, Rosen A, Gabrieli J. Dissociable Contributions of Prefrontal and Parietal Cortices to Response Selection. Neuroimage. 2002 Nov;17(3):1562–1571.

[324] Peterson BS, Kane MJ, Alexander GM, Lacadie C, Skudlarski P, Leung HC, et al. An event-related functional MRI study comparing interference effects in the Simon and Stroop tasks. Brain research Cognitive brain research. 2002 May;13(3):427–40.

[325] Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG. A general mechanism for perceptual decision-making in the human brain. Nature. 2004 Oct;431(7010):859–62.

[326] Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci. 2005 May;6(5):363–75.

[327] Gold JI, Shadlen MN. The neural basis of decision making. Annu Rev Neurosci. 2007 Jan;30:535–74.

[328] Romo R, Hernández A, Salinas E, Brody CD, Zainos A, Lemus L, et al. From sensation to action. Behav Brain Res. 2002 Sep;135(1-2):105–18.

[329] Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. Journal of Neurophysiology. 2001 Oct;86(4):1916–36.

[330] Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG. Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. Proc Natl Acad Sci USA. 2006 Jun;103(26):10023–8.

[331] Philiastides MG, Sajda P. Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb Cortex. 2006 Apr;16(4):509–18.

[332] Kaiser J, Lennert T, Lutzenberger W. Dynamics of oscillatory activity during auditory decision making. Cereb Cortex. 2007 Oct;17(10):2258–67.

[333] Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. Neural correlates of sensory and decision processes in auditory object identification. Nat Neurosci. 2004 Mar;7(3):295–301.

[334] Grinband J, Hirsch J, Ferrera V. A Neural Representation of Categorization Uncertainty in the Human Brain. Neuron. 2006 Mar;49(5):757–763.

[335] Heekeren HR, Marrett S, Ungerleider LG. The neural systems that mediate human perceptual decision making. Nat Rev Neurosci. 2008 Jun;9(6):467–79.

[336] Kim JN, Shadlen MN. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci. 1999 Feb;2(2):176–85.

[337] Ditterich J, Mazurek ME, Shadlen MN. Microstimulation of visual cortex affects the speed of perceptual decisions. Nat Neurosci. 2003 Aug;6(8):891–898.

[338] Rustichini A. Neuroeconomics: what have we found, and what should we search for. Current Opinion in Neurobiology. 2009 Dec;19(6):672–7.

[339] Krain A, Wilson A, Arbuckle R, Castellanos F, Milham M. Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. Neuroimage. 2006 Aug;32(1):477–484.

[340] Addington J, Addington D. Facial affect recognition and information processing in schizophrenia and bipolar disorder. Schizophr Res. 1998 Aug;32(3):171–81.

[341] Baudouin JY, Vernet M, Franck N. Second-order facial information processing in schizophrenia. Neuropsychology. 2008 May;22(3):313–20.

[342] Butler PD, Tambini A, Yovel G, Jalbrzikowski M, Ziwich R, Silipo G, et al. What's in a face? Effects of stimulus duration and inversion on face processing in schizophrenia. Schizophr Res. 2008 Aug;103(1-3):283–92.

[343] Livingstone MS, Hubel DH. Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. J Neurosci. 1987 Nov;7(11):3416–3468.

[344] Schwartz BD, McGinn T, Winstead DK. Disordered spatiotemporal processing in schizophrenics. Biol Psychiatry. 1987 Jun;22(6):688–698.

[345] Cadenhead KS, Serper Y, Braff DL. Transient versus sustained visual channels in the visual backward masking deficits of schizophrenia patients. Biol Psychiatry. 1998 Jan;43(2):132–138.

[346] Bedwell JS, Brown JM, Miller LS. The magnocellular visual system and schizophrenia: what can the color red tell us? Schizophr Res. 2003 Oct;63(3):273–284.

[347] Butler PD, Zemon V, Schechter I, Saperstein AM, Hoptman MJ, Lim KO, et al. Early-stage visual processing and cortical amplification deficits in schizophrenia. Arch Gen Psychiatry. 2005 May;62(5):495–504.

[348] Keri S, Antal A, Szekeres G, Benedek G, Janka Z. Spatiotemporal visual processing in schizophrenia. J Neuropsychiatry Clin Neurosci. 2002 Spring;14(2):190–196.

[349] Keri S, Kiss I, Kelemen O, Benedek G, Janka Z. Anomalous visual experiences, negative symptoms, perceptual organization and the magnocellular pathway in schizophrenia: a shared construct? Psychol Med. 2005 Oct;35(10):1445–1455.

[350] Keri S, Kelemen O, Benedek G, Janka Z. Vernier threshold in patients with schizophrenia and in their unaffected siblings. Neuropsychology. 2004 Jul;18(3):537–542.

[351] Braus DF, Weber-Fahr W, Tost H, Ruf M, Henn FA. Sensory information processing in neuroleptic-naive first-episode schizophrenic patients: a functional magnetic resonance imaging study. Arch Gen Psychiatry. 2002 Aug;59(8):696–701.

[352] Slaghuis WL. Contrast sensitivity for stationary and drifting spatial frequency gratings in positive- and negative-symptom schizophrenia. J Abnorm Psychol. 1998 Feb;107(1):49–62.

[353] Heim M, Morgner J. [Disturbed color vision in endogenous psychoses]. Psychiatr Prax. 2001 Sep;28(6):284–286.

[354] Shuwairi SM, Cronin-Golomb A, McCarley RW, O'Donnell BF. Color discrimination in schizophrenia. Schizophr Res. 2002 May;55(1-2):197–204.

[355] Brenner CA, Lysaker PH, Wilt MA, O'Donnell BF. Visual processing and neuropsychological function in schizophrenia and schizoaffective disorder. Psychiatry Res. 2002 Aug;111(2-3):125–136.

[356] Keri S, Kelemen O, Janka Z, Benedek G. Visual-perceptual dysfunctions are possible endophenotypes of schizophrenia: evidence from the psychophysical investigation of magnocellular and parvocellular pathways. Neuropsychology. 2005 Sep;19(5):649–656.

[357] Witkovsky P. Dopamine and retinal function. Doc Ophthalmol. 2004 Jan;108(1):17–40.

[358] Dacey DM. The dopaminergic amacrine cell. J Comp Neurol. 1990 Nov;301(3):461–489.

[359] Jensen RJ, Daw NW. Effects of dopamine and its agonists and antagonists on the receptive field properties of ganglion cells in the rabbit retina. Neuroscience. 1986 Mar;17(3):837–855.

[360] Witkovsky P, Shi XP. Slow light and dark adaptation of horizontal cells in the Xenopus retina: a role for endogenous dopamine. Vis Neurosci. 1990 Oct;5(4):405–413.

[361] DeVries SH, Schwartz EA. Modulation of an electrical synapse between solitary pairs of catfish horizontal cells by dopamine and second messengers. J Physiol. 1989 Jul;414:351–375.

[362] McMahon DG, Knapp AG, Dowling JE. Horizontal cell gap junctions: single-channel conductance and modulation by dopamine. Proc Natl Acad Sci U S A. 1989 Oct;86(19):7639–7643.

[363] Xia XB, Mills SL. Gap junctional regulatory mechanisms in the AII amacrine cell of the rabbit retina. Vis Neurosci. 2004 Sep;21(5):791–805.

[364] Cahill GM, Grace MS, Besharse JC. Rhythmic regulation of retinal melatonin: metabolic pathways, neurochemical mechanisms, and the ocular circadian clock. Cell Mol Neurobiol. 1991 Oct;11(5):529–560.

[365] Muresan Z, Besharse JC. D2-like dopamine receptors in amphibian retina: localization with fluorescent ligands. J Comp Neurol. 1993 May;331(2):149–160.

[366] Patel S, Chapman KL, Marston D, Hutson PH, Ragan CI. Pharmacological and functional characterisation of dopamine D4 receptors in the rat retina. Neuropharmacology. 2003 Jun;44(8):1038–1046.

[367] Rivera A, Trias S, Penafiel A, Angel Narvaez J, Diaz-Cabiale Z, Moratalla R, et al. Expression of D4 dopamine receptors in striatonigral and striatopallidal neurons in the rat striatum. Brain Res. 2003 Oct;989(1):35–41.

[368] Kapur S, Agid O, Mizrahi R, Li M. How antipsychotics work-from receptors to reality. NeuroRx. 2006 Jan;3(1):10–21.

[369] Veruki ML, Wassle H. Immunohistochemical localization of dopamine D1 receptors in rat retina. Eur J Neurosci. 1996 Nov;8(11):2286–2297.

[370] Nguyen-Legros J, Simon A, Caille I, Bloch B. Immunocytochemical localization of dopamine D1 receptors in the retina of mammals. Vis Neurosci. 1997 May;14(3):545–551.

[371] Hare WA, Owen WG. Similar effects of carbachol and dopamine on neurons in the distal retina of the tiger salamander. Vis Neurosci. 1995 May;12(3):443–455.

[372] Yazulla S, Kleinschmidt J. Dopamine blocks carrier-mediated release of GABA from retinal horizontal cells. Brain Res. 1982 Feb;233(1):211–215.

[373] Mangel SC. Analysis of the horizontal cell contribution to the receptive field surround of ganglion cells in the rabbit retina. J Physiol. 1991 Oct;442:211–234.

[374] Dong CJ, McReynolds JS. The relationship between light, dopamine release and horizontal cell coupling in the mudpuppy retina. J Physiol. 1991;440:291–309.

[375] Buttner T, Kuhn W, Muller T, Patzold T, Heidbrink K, Przuntek H. Distorted color discrimination in 'de novo' parkinsonian patients. Neurology. 1995 Feb;45(2):386–387.

[376] Muller T, Kuhn W, Buttner T, Przuntek H. Distorted colour discrimination in Parkinson's disease is related to severity of the disease. Acta Neurol Scand. 1997 Nov;96(5):293–296.

[377] Desai P, Roy M, Roy A, Brown S, Smelson D. Impaired color vision in cocaine-withdrawn patients. Arch Gen Psychiatry. 1997 Aug;54(8):696–699.

[378] Silva MF, Faria P, Regateiro FS, Forjaz V, Januario C, Freire A, et al. Independent patterns of damage within magno-, parvo- and koniocellular pathways in Parkinson's disease. Brain. 2005 Oct;128(Pt 10):2260–2271.

[379] Buttner T, Kuhn W, Patzold T, Przuntek H. L-Dopa improves colour vision in Parkinson's disease. J Neural Transm Park Dis Dement Sect. 1994;7(1):13–19.

[380] Witkovsky P, Nicholson C, Rice ME, Bohmaker K, Meller E. Extracellular dopamine concentration in the retina of the clawed frog, Xenopus laevis. Proc Natl Acad Sci U S A. 1993 Jun;90(12):5667–5671.

[381] Muller T, Meisel M, Russ H, Przuntek H. Motor impairment influences Farnsworth-Munsell 100 Hue test error scores in Parkinson's disease patients. J Neurol Sci. 2003 Sep;213(1-2):61–65.

[382] Chen Y, Levy DL, Sheremata S, Nakayama K, Matthysse S, Holzman PS. Effects of typical, atypical, and no antipsychotic drugs on visual contrast detection in schizophrenia. Am J Psychiatry. 2003 Oct;160(10):1795–1801.

[383] Harris JP, Calvert JE, Leendertz JA, Phillipson OT. The influence of dopamine on spatial vision. Eye. 1990;4 (Pt 6):806–812.

[384] Masson G, Mestre D, Blin O. Dopaminergic modulation of visual sensitivity in man. Fundam Clin Pharmacol. 1993;7(8):449–463.

[385] Stanzione P, Traversa R, Pierantozzi M, Semprini R, Marciani MG, Bernardi G. An electrophysiological study of D2 dopaminergic actions in normal human retina: a tool in Parkinson's disease. Neurosci Lett. 1992 Jun;140(1):125–128.

[386] Stanzione P, Pierantozzi M, Semprini R, Tagliati M, Traversa R, Peppe A, et al. Increasing doses of l-sulpiride reveal dose- and spatial frequency-dependent effects of D2 selective blockade in the human electroretinogram. Vision Res. 1995 Sep;35(18):2659–2664.

[387] Tagliati M, Bodis-Wollner I, Kovanecz I, Stanzione P. Spatial frequency tuning of the monkey pattern ERG depends on D2 receptor-linked action of dopamine. Vision Res. 1994 Aug;34(16):2051–2057.

[388] Bodis-Wollner I, Tzelepi A. The push-pull action of dopamine on spatial tuning of the monkey retina: the effects of dopaminergic deficiency and selective D1 and D2 receptor ligands on the pattern electroretinogram. Vision Res. 1998 May;38(10):1479–1487.

[389] Buttner T, Muller T, Kuhn W. Effects of apomorphine on visual functions in Parkinson's disease. J Neural Transm. 2000;107(1):87–94.

[390] Mora-Ferrer C, Neumeyer C. Reduction of red-green discrimination by dopamine D1 receptor antagonists and retinal dopamine depletion. Vision Res. 1996 Dec;36(24):4035–4044.

[391] Papadopoulos GC, Parnavelas JG. Distribution and synaptic organization of dopaminergic axons in the lateral geniculate nucleus of the rat. J Comp Neurol. 1990 Apr;294(3):356–361.

[392] Zhao Y, Kerscher N, Eysel U, Funke K. D1 and D2 receptor-mediated dopaminergic modulation of visual responses in cat dorsal lateral geniculate nucleus. J Physiol. 2002 Feb;539(Pt 1):223–238.

[393] Roth A, Lanthony P. Vision des couleurs. JF R, editor. Masson, Paris; 1999.

[394] Krainik A, Lehericy S, Hennel F, de Moortele PFV, Marsault C, Bihan DL. Comparison of cortical and basal ganglia activation at 1.5T and 3T. NeuroImage. 2000 Jun;11(5):S552. Notes.

[395] Scholz VH, Flaherty AW, Kraft E, Keltner JR, Kwong KK, Chen YI, et al. Laterality, somatotopy and reproducibility of the basal ganglia and motor cortex during motor tasks. Brain Research. 2000 Oct;879(1-2):204–15.

[396] Moritz CH, Meyerand ME, Cordes D, Haughton VM. Functional MR imaging activation after finger tapping has a shorter duration in the basal ganglia than in the sensorimotor cortex. AJNR American journal of neuroradiology. 2000 Aug;21(7):1228–34.

[397] Drake RL, Vogl AW, Mitchell AWM. 8: Head and Neck. In: Gray's Anatomy for Students. 2nd ed. Churchill Livingston Elsevier, Philadelphia, PA; 2010. p. 887–891.

[398] Gagnon D, O'Driscoll GA, Petrides M, Pike GB. The effect of spatial and temporal information on saccades and neural activity in oculomotor structures. Brain. 2002 Jan;125(Pt 1):123–39.

[399] Camchong J, Dyckman KA, Chapman CE, Yanasak NE, McDowell JE. Basal ganglia-thalamocortical circuitry disruptions in schizophrenia during delayed response tasks. Biol Psychiatry. 2006 Aug;60(3):235–41.

[400] Thaker GK. Neurophysiological endophenotypes across bipolar and schizophrenia psychosis. Schizophrenia Bulletin. 2008 Jul;34(4):760–73.

[401] Menon V, Glover GH, Pfefferbaum A. Differential activation of dorsal basal ganglia during externally and self paced sequences of arm movements. Neuroreport. 1998 May;9(7):1567–73.

[402] Vaillancourt D. Subthalamic nucleus and internal globus pallidus scale with the rate of change of force production in humans. Neuroimage. 2004 Sep;23(1):175–186.

[403] Taniwaki T, Okayama A, Yoshiura T, Nakamura Y, Goto Y, ichi Kira J, et al. Reappraisal of the motor role of basal ganglia: a functional magnetic resonance image study. J Neurosci. 2003 Apr;23(8):3432–8.

[404] Wolf R, Walter H. Evaluation of a novel event-related parametric fMRI paradigm investigating prefrontal function. Psychiatry Research: Neuroimaging. 2005 Oct;140(1):73–83.

[405] Wager TD, Nichols TE. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage. 2003 Feb;18(2):293–309.

[406] Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS. Human striatal response to salient nonrewarding stimuli. J Neurosci. 2003 Sep;23(22):8092–7.

[407] Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS. Human striatal responses to monetary reward depend on saliency. Neuron. 2004;42:509–517.

[408] Zink C, Pagnoni G, Chappelow J, Martin-Skurski M, Berns G. Human striatal activation reflects degree of stimulus saliency. Neuroimage. 2005 Sep;.

[409] Taniwaki T, Okayama A, Yoshiura T, Togao O, Nakamura Y, Yamasaki T, et al. Functional network of the basal ganglia and cerebellar motor loops in vivo: Different activation patterns between self-initiated and externally triggered movements. Neuroimage. 2006 Jun;31(2):745–753.

[410] Monchi O, Petrides M, Strafella AP, Worsley KJ, Doyon J. Functional role of the basal ganglia in the planning and execution of actions. Ann Neurol. 2006 Feb;59(2):257–264.

[411] Pinckers A, Cruysberg JR. FM 100 Hue test and lightness discrimination test. Documenta ophthalmologica Advances in ophthalmology. 1986 Dec;64(1):19–22.

[412] International Commission on Illumination (CIE). Joint ISO/CIE Standard ISO 11664-5: 2009(E)/CIE S014-5/E:2008, Colorimetry - Part 5: CIE 1976 L*u*v* Colour Space and u', v' Uniform Chromaticity Scale Diagram; 2009.

[413] Anstis S, Cavanagh P. A minimum motion technique for judging equiluminance. Mollon JD, Sharpe LT, editors. Academic Press, London; 1983.

[414] Regan BC, Reffin JP, Mollon JD. Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. Vision Res. 1994 May;34(10):1279–99.

[415] Treutwein B. Adaptive psychophysical procedures. Vision Research. 1995 Sep;35(17):2503–22.

[416] Taylor M, Creelman C. PEST: Efficient estimates on probability functions. Journal of the Acoustical Society of America. 1967;41:782–787.

[417] Büchel C, Friston K. Assessing interactions among neuronal systems using functional neuroimaging. Neural networks : the official journal of the International Neural Network Society. 2000 Jan;13(8-9):871–82.

[418] Rogers BP, Morgan VL, Newton AT, Gore JC. Assessing functional connectivity in the human brain by fMRI. Magnetic Resonance Imaging. 2007 Dec;25(10):1347–57.

[419] Büchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. Cereb Cortex. 1997 Dec;7(8):768–78.

[420] Steele JD, Meyer M, Ebmeier KP. Neural predictive error signal correlates with depressive illness severity in a game paradigm. Neuroimage. 2004 Sep;23(1):269–80.

[421] Friston K. Dynamic causal modelling. Neuroimage. 2003 Aug;19(4):1273–1302.

[422] Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage. 1997 Oct;6(3):218–29.

[423] Gitelman DR, Penny WD, Ashburner J, Friston KJ. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage. 2003 May;19(1):200–7.

[424] Granger C. Investigating causal relations by econometric models and cross-spectral methods. Econometrica. 1969;37(3):424–438.

[425] Goebel R, Roebroeck A, Kim DS, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magnetic Resonance Imaging. 2003 Dec;21(10):1251–61.

[426] Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage. 2005 Mar;25(1):230–242.

[427] Abler B, Roebroeck A, Goebel R, Höse A, Schönfeldt-Lecuona C, Hole G, et al. Investigating directed influences between activated brain areas in a motor-response task using fMRI. Magnetic Resonance Imaging. 2006 Feb;24(2):181–5.

[428] Londei A, Dausilio A, Basso D, Sestieri C, Delgratta C, Romani G, et al. Brain network for passive word listening as evaluated with ICA and Granger causality. Brain Research Bulletin. 2007 May;72(4-6):284–292.

[429] Stilla R, Deshpande G, Laconte S, Hu X, Sathian K. Posteromedial Parietal Cortical Activity and Inputs Predict Tactile Spatial Acuity. J Neurosci. 2007 Oct;27(41):11091–11102.

[430] Bullmore E, Brammer M, Williams SC, Rabe-Hesketh S, Janot N, David A, et al. Statistical methods of estimation and inference for functional MR image analysis. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine. 1996 Feb;35(2):261–77.

[431] Weisskoff R, Baker J, Belliveau J, Davis T, Kwong K, Cohen M, et al. Power spectrum analysis of functionally weighted MR data: what's in the noise. Proc SMRM (New York). 1993;12:7.

[432] Lund T, Hanson L. Physiological noise correction in fMRI using vessel time-series as covariates in a general linear model. Proceedings of the 9th Annual Meeting of ISMRM. 2001;p. 22.

[433] Geweke J. Measurement of Linear Dependence and Feedback Between Multiple Time Series. J AM STAT ASSOC. 1982;77(378):304–313.

[434] Geweke J. Measures of conditional linear dependence and feedback between time series. J AM STAT ASSOC. 1984;79(388):907–915.

[435] Genovese C. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. Neuroimage. 2002 Apr;15(4):870–878.

[436] Abler B, Roebroeck A, Goebel R, Hose A, Schonfeldtlecuona C, Hole G, et al. Investigating directed influences between activated brain areas in a motor-response task using fMRI. Magnetic Resonance Imaging. 2006 Feb;24(2):181–185.

[437] Stouffer SA, Suchman EA, DeVinney LC, Star SA, Williams RMJ. he American Soldier, Vol. 1: Adjustment during Army Life. Princeton University Press, Princeton; 1949.

[438] Glenn N. Replications, Significance Tests and Confidence in Findings in Survey Research. Public Opinion Quarterly. 1983 Jul;p. 261–269.

[439] Whitlock MC. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. J Evolution Biol. 2005 Sep;18(5):1368–73.

[440] Bartels A, Zeki S. The architecture of the colour centre in the human visual brain: new results and a review. Eur J Neurosci. 2000 Jan;12(1):172–93.

[441] Beauchamp MS, Haxby JV, Jennings JE, DeYoe EA. An fMRI version of the Farnsworth-Munsell 100-Hue test reveals multiple color-selective areas in human ventral occipitotemporal cortex. Cereb Cortex. 1999 Jan;9(3):257–63.

[442] Claeys KG, Dupont P, Cornette L, Sunaert S, Hecke PV, Schutter ED, et al. Color discrimination involves ventral and dorsal stream visual areas. Cereb Cortex. 2004 Jul;14(7):803–22.

[443] Conway BR. Colour vision: a clue to hue in v2. Curr Biol. 2003 Apr;13(8):R308–10.

[444] Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, et al. fMRI of human visual cortex. Nature. 1994 Jun;369(6481):525.

[445] Engel S, Zhang X, Wandell B. Colour tuning in human visual cortex measured with functional magnetic resonance imaging. Nature. 1997 Jul;388(6637):68–71.

[446] Gegenfurtner KR. Cortical mechanisms of colour vision. Nat Rev Neurosci. 2003 Jul;4(7):563–72.

[447] Hadjikhani N, Liu AK, Dale AM, Cavanagh P, Tootell RB. Retinotopy and color sensitivity in human visual cortical area V8. Nat Neurosci. 1998 Jul;1(3):235–41.

[448] Komatsu H. Mechanisms of central color vision. Current Opinion in Neurobiology. 1998 Aug;8(4):503–8.

[449] Pulvermüller F, Hauk O. Category-specific conceptual processing of color and form in left fronto-temporal cortex. Cereb Cortex. 2006 Aug;16(8):1193–201.

[450] Simmons WK, Ramjee V, Beauchamp MS, McRae K, Martin A, Barsalou LW. A common neural substrate for perceiving and knowing about color. Neuropsychologia. 2007 Sep;45(12):2802–10.

[451] Thulborn KR, Chang SY, Shen GX, Voyvodic JT. High-resolution echo-planar fMRI of human visual cortex at 3.0 tesla. NMR in biomedicine. 1997 Jan;10(4-5):183–90.

[452] Wachtler T, Sejnowski TJ, Albright TD. Representation of color stimuli in awake macaque primary visual cortex. Neuron. 2003 Feb;37(4):681–91.

[453] Wade AR, Brewer AA, Rieger JW, Wandell BA. Functional measurements of human ventral occipital cortex: retinotopy and colour. Philos Trans R Soc Lond, B, Biol Sci. 2002 Aug;357(1424):963–73.

[454] Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, et al. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 2000 Jul;10(3):120–31.

[455] Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 2003 Jul;19(3):1233–9.

[456] Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF. Neural systems responding to degrees of uncertainty in human decision-making. Science. 2005 Dec;310:1680–1683.

[457] Brett M, Anton JL, Valabregue R, Poline JB. Region of interest analysis using an SPM toolbox. In: 8th International Conference on Functional Mapping of the Human Brain; 2002. .

[458] O'doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron. 2003 Apr;38(2):329–37.

[459] Brett M, Penny W, Kiebel S. SPM2 manual: Chapter 14 - An introduction to random field theory. 2003 Mar;p. 23.

[460] Carpenter MB, Baton RRr, Carleton SC, Keller JT. Interconnections and organization of pallidal and subthalamic nucleus neurons in the monkey. J Comp Neurol. 1981;197(4):579–603.

[461] Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature. 1999 Nov;402(6758):179–81.

[462] Vanveen V, Cohen JD, Botvinick M, Stenger V, Carter C. Anterior Cingulate Cortex, Conflict Monitoring, and Levels of Processing. Neuroimage. 2001 Dec;14(6):1302–1308.

[463] Cisek P, Kalaska JF. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron. 2005 Mar;45(5):801–14.

[464] Leh S, Ptito A, Chakravarty M, Strafella A. Fronto-striatal connections in the human brain: A probabilistic diffusion tractography study. Neuroscience Letters. 2007 May;419(2):113–118.

[465] Bogacz R. Optimal decision-making theories: linking neurobiology with behaviour. Trends in Cognitive Sciences. 2007 Mar;11(3):118–25.

[466] Pasquereau B, Nadjar A, Arkadir D, Bezard E, Goillandeau M, Bioulac B, et al. Shaping of motor responses by incentive values through the basal ganglia. J Neurosci. 2007;27(5):1176–83.

[467] Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci. 2005 Dec;8(12):1704–1711.

[468] Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. Trends in Cognitive Sciences. 2004 Dec;8(12):539–46.

[469] Carter CS, Macdonald AM, Botvinick M, Ross LL, Stenger VA, Noll D, et al. Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. Proc Natl Acad Sci USA. 2000 Feb;97(4):1944–8.

[470] Kerns JG. Anterior Cingulate Conflict Monitoring and Adjustments in Control. Science. 2004 Feb;303(5660):1023–1026.

[471] Wager T, Sylvester C, Lacey S, Nee D, Franklin M, Jonides J. Common and unique components of response inhibition revealed by fMRI. Neuroimage. 2005 Aug;27(2):323–340.

[472] de Jong BM, Paans AMJ. Medial versus lateral prefrontal dissociation in movement selection and inhibitory control. Brain Res. 2007;1132(1):139–47.

[473] Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, et al. Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. J Neurosci. 1999 Oct;19(20):9029–38.

[474] Critchley HD, Mathias CJ, Dolan RJ. Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron. 2001 Feb;29(2):537–45.

[475] Tobler P, O'Doherty J, Dolan R, Schultz W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J Neurophysiol. 2007;97:1621–1632.

[476] Kilner JM, Friston KJ, Frith CD. Predictive coding: an account of the mirror neuron system. Cogn Process. 2007 Aug;8(3):159–166.

[477] Wolpert DM, Miall RC. Forward Models for Physiological Motor Control. Neural networks : the official journal of the International Neural Network Society. 1996 Nov;9(8):1265–1279.

[478] Koski L, Wohlschläger A, Bekkering H, Woods RP, Dubeau MC, Mazziotta JC, et al. Modulation of motor and premotor activity during imitation of target-directed actions. Cereb Cortex. 2002 Aug;12(8):847–55.

[479] Heiser M, Iacoboni M, Maeda F, Marcus J, Mazziotta JC. The essential role of Broca's area in imitation. Eur J Neurosci. 2003 Mar;17(5):1123–8.

[480] Volz K, Schubotz R, Cramon D. Variants of uncertainty in decision-making and their neural correlates. Brain Research Bulletin. 2005 Nov;67(5):403–412.

[481] Rubia K. Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. Neuroimage. 2003 Sep;20(1):351–358.

[482] Bar-Gad I, Bergman H. Stepping out of the box: information processing in the neural networks of the basal ganglia. Curr Opin Neurobiol. 2001 Dec;11(6):689–95.

[483] Haruno M, Kawato M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. Journal of Neurophysiology. 2006 Feb;95(2):948–59.

[484] Humphries MD, Gurney... KN, Prescott TJ. Is there a brainstem substrate for action selection? Philos Trans R Soc Lond, B, Biol Sci. 2007 Sep;362(1485):1627–39.

[485] Cisek P. Integrated neural processes for defining potential actions and deciding between them: a computational model. J Neurosci. 2006 Sep;26(38):9761–70.

[486] Contreras-Vidal JL, Schultz W. A predictive reinforcement model of dopamine neurons for learning approach behavior. Journal of computational neuroscience. 1999 Jan;6(3):191–214.

[487] Humphries MD, Stewart RD, Gurney... KN. A Physiologically Plausible Model of Action Selection and Oscillatory Activity in the Basal Ganglia. J Neurosci. 2006 Dec;26(50):12921–12942.

[488] Robbins TW. The case of frontostriatal dysfunction in schizophrenia. Schizophr Bull. 1990;16:391–402.

[489] Lewis DA, Gonzalez-Burgos G. Pathophysiologically based treatment interventions in schizophrenia. Nat Med. 2006 Sep;12(9):1016–22.

[490] Frith CD, Blakemore S, Wolpert DM. Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. Brain Res Brain Res Rev. 2000 Mar;31(2-3):357–63.

[491] Eyler LT, Olsen RK, Jeste DV, Brown GG. Abnormal brain response of chronic schizophrenia patients despite normal performance during a visual vigilance task. Psychiatry Res. 2004;130:245–257.

[492] Weiss I, Feldon J. Environmental animal models for sensorimotor gating deficiencies in schizophrenia: a review. Psychopharmacology. 2001 Jul;156(2-3):305–326.

[493] Gonul AS, Kula M, Eşel E, Tutuş A, Sofuoglu S. A Tc-99m HMPAO SPECT study of regional cerebral blood flow in drug-free schizophrenic patients with deficit and non-deficit syndrome. Psychiatry Res. 2003;123:199–205.

[494] Brown LL, Sharp FR. Metabolic mapping of rat striatum: somatotopic organization of sensorimotor activity. Brain Res. 1995 Jul;686(2):207–22.

[495] Feldman J, Ballard D. Connectionist models and their properties. Cognitive Sc: A Multidisciplinary J. 1982 Jan;6:205–254.

[496] Yuille AL, Geiger D. Winner-Take-All Mechanisms. The Handbook of Brain Theory and Neural Networks. 1998 Aug;p. 1–16.

[497] Groves PM. A theory of the functional organization of the neostriatum and the neostriatal control of voluntary movement. Brain Res. 1983 Mar;286(2):109–32.

[498] Beiser DG, Houk JC. Model of cortical-basal ganglionic processing: encoding the serial order of sensory events. Journal of Neurophysiology. 1998 Jun;79(6):3168–88.

[499] Wickens J, Oorshchot D. Neuronal dynamics and surround inhibition in the neostriatum: a possible connection. In: Miller R, Wickens J, editors. Brain Dynamics and the Striatal Complex. Harwood; 2000. p. 141–150.

[500] Kawaguchi Y, Wilson CJ, Emson PC. Projection subtypes of rat neostriatal matrix cells revealed by intracellular injection of biocytin. J Neurosci. 1990 Oct;10(10):3421–38.

[501] Bolam JP, Somogyi P, Takagi H, Fodor I, Smith AD. Localization of substance P-like immunoreactivity in neurons and nerve terminals in the neostriatum of the rat: a correlated light and electron microscopic study. J Neurocytol. 1983 Apr;12(2):325–44.

[502] Jaeger D, Kita H, Wilson CJ. Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. Journal of Neurophysiology. 1994 Nov;72(5):2555–8.

[503] Kita H. GABAergic circuits of the striatum. Prog Brain Res. 1993 Jan;99:51–72.

[504] Zheng T, Wilson CJ. Corticostriatal combinatorics: the implications of corticostriatal axonal arborizations. Journal of Neurophysiology. 2002 Feb;87(2):1007–17.

[505] Wilson CJ. GABAergic inhibition in the neostriatum. Prog Brain Res. 2007 Jan;160:91–110.

[506] Tepper JM, Wilson CJ, Koós T. Feedforward and feedback inhibition in neostriatal GABAergic spiny neurons. Brain Research Reviews. 2008 Aug;58(2):272–81.

[507] Kawaguchi Y. Physiological, morphological, and histochemical characterization of three classes of interneurons in rat neostriatum. J Neurosci. 1993 Nov;13(11):4908–23.

[508] Kubota Y, Kawaguchi Y. Three classes of GABAergic interneurons in neocortex and neostriatum. Jpn J Physiol. 1994 Jan;44 Suppl 2:S145–8.

[509] Kawaguchi Y, Wilson CJ, Augood SJ, Emson PC. Striatal interneurones: chemical, physiological and morphological characterization. Trends in Neurosciences. 1995 Dec;18(12):527–35.

[510] Koós T, Tepper JM. Inhibitory control of neostriatal projection neurons by GABAergic interneurons. Nat Neurosci. 1999 May;2(5):467–72.

[511] Kita H, Kosaka T, Heizmann CW. Parvalbumin-immunoreactive neurons in the rat neostriatum: a light and electron microscopic study. Brain Res. 1990 Dec;536(1-2):1–15.

[512] Bennett BD, Bolam JP. Synaptic input and output of parvalbumin-immunoreactive neurons in the neostriatum of the rat. Neuroscience. 1994 Oct;62(3):707–19.

[513] Ramanathan S, Hanley JJ, Deniau JM, Bolam JP. Synaptic convergence of motor and somatosensory cortical afferents onto GABAergic interneurons in the rat striatum. J Neurosci. 2002 Sep;22(18):8158–69.

[514] Plotkin JL, Wu N, Chesselet MF, Levine MS. Functional and molecular development of striatal fast-spiking GABAergic interneurons and their cortical inputs. Eur J Neurosci. 2005 Sep;22(5):1097–108.

[515] Taverna S, Canciani B, Pennartz CMA. Membrane properties and synaptic connectivity of fast-spiking interneurons in rat ventral striatum. Brain Res. 2007 Jun;1152:49–56.

[516] Yuille A, Grzywacz N. A winner-take-all mechanism based on presynaptic inhibition feedback. Neural computation. 1989 Jan;1:334–347.

[517] Yu AJ, Giese MA, Poggio TA. Biophysiologically plausible implementations of the maximum operation. Neural computation. 2002 Dec;14(12):2857–81.

[518] Humphries MD, Wood R, Gurney. . . KN. Dopamine-modulated dynamic cell assemblies generated by the GABAergic striatal microcircuit. Neural networks : the official journal of the International Neural Network Society. 2009 Jul;.

[519] Gittis AH, Nelson AB, Thwin MT, Palop JJ, Kreitzer AC. Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. J Neurosci. 2010 Feb;30(6):2223–34.

[520] Kita H. Globus pallidus external segment. Prog Brain Res. 2007 Jan;160:111–33.

[521] Kita H, Tokuno H, Nambu A. Monkey globus pallidus external segment neurons projecting to the neostriatum. Neuroreport. 1999 May;10(7):1467–72.

[522] Kita H, Kita T. Number, origins, and chemical types of rat pallidostriatal projection neurons. J Comp Neurol. 2001 Sep;437(4):438–48.

[523] Bevan MD, Booth PA, Eaton SA, Bolam JP. Selective innervation of neostriatal interneurons by a subclass of neuron in the globus pallidus of the rat. J Neurosci. 1998 Nov;18(22):9438–52.

[524] Rajakumar N, Elisevich K, Flumerfelt BA. The pallidostriatal projection in the rat: a recurrent inhibitory loop? Brain Res. 1994 Jul;651(1-2):332–6.

[525] Bracci E, Centonze D, Bernardi G, Calabresi P. Dopamine excites fast-spiking interneurons in the striatum. Journal of Neurophysiology. 2002 Apr;87(4):2190–4.

[526] Frankle WG, Laruelle M, Haber SN. Prefrontal cortical projections to the midbrain in primates: evidence for a sparse connection. Neuropsychopharmacology. 2006 Aug;31(8):1627–36.

[527] Celada P, Paladini CA, Tepper JM. GABAergic control of rat substantia nigra dopaminergic neurons: role of globus pallidus and substantia nigra pars reticulata. Neuroscience. 1999 Mar;89(3):813–25.

[528] Tepper JM, Lee CR. GABAergic control of substantia nigra dopaminergic neurons. Prog Brain Res. 2007 Jan;160:189–208.

[529] Sesack SR, Carr DB, Omelchenko N, Pinto A. Anatomical substrates for glutamate-dopamine interactions: evidence for specificity of connections and extrasynaptic actions. Annals of the New York Academy of Sciences. 2003 Nov;1003:36–52.

[530] Plenz D, Kitai ST. Up and down states in striatal medium spiny neurons simultaneously recorded with spontaneous activity in fast-spiking interneurons studied in cortex-striatum-substantia nigra organotypic cultures. J Neurosci. 1998 Jan;18(1):266–83.

[531] Sesack SR, Grace A. Cortico-Basal Ganglia Reward Network: Microcircuitry. Neuropsychopharmacology. 2009 Aug;.

[532] Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry. 2003 Jan;160(1):13–23.

[533] Tan H, Callicott J, Weinberger D. Dysfunctional and compensatory prefrontal cortical systems, genes and the pathogenesis of schizophrenia. Cereb Cortex. 2007;17:171–181.

[534] Farnsworth D. The Farnsworth-Munsell 100-hue test for the examinaion of color discrimination. 1957 Nov;p. 9.

[535] Wade D. Measurement in neurological rehabilitation. Oxford: Oxford University Press; 1992.

[536] Fahn S, Elton RS, of the UPDRS Development Committee M. Unified Parkinson's Disease Rating Scale. Fahn S, Marsden CD, Calne DB, Goldstein M, editors. NJ: Macmillan, Florham Park; 1987.

[537] Kay SR. Positive-negative symptom assessment in schizophrenia: psychometric issues and scale comparison. The Psychiatric quarterly. 1990 Jan;61(3):163–78.

[538] Nelson H. National Adult Reading Test (NART). Windsor: NFER; 1982.

[539] Wechsler D. Wechsler Abbreviated Scale of Intelligence (WASI). San Antonio, TX: Harcourt Assessment.; 1999.

[540] Woods SW. Chlorpromazine equivalent doses for the newer atypical antipsychotics. The Journal of clinical psychiatry. 2003 Jun;64(6):663–7.

[541] Schlagenhauf F, Sterzer P, Schmack K, Ballmaier M, Rapp M, Wrase J, et al. Reward Feedback Alterations in Unmedicated Schizophrenia Patients: Relevance for Delusions. Biol Psychiatry. 2009 Feb;.

[542] Juckel G, Schlagenhauf F, Koslowski M, Wustenberg T, Villringer A, Knutson B, et al. Dysfunction of ventral striatal reward prediction in schizophrenia. Neuroimage. 2006 Jan;29(2):409–416.

[543] Lisman J, Coyle J, Green R, Javitt D, Benes F, Heckers S, et al. Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. TINS. 2008;In Press.

[544] Corlett PR, Murray GK, Honey GD, Aitken MRF, Shanks DR, Robbins TW, et al. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain. 2007 Sep;130(Pt 9):2387–400.

[545] Romaniuk L, Honey GD, King JR, Whalley HC, McIntosh AM, Levita L, et al. Midbrain Activation During Pavlovian Conditioning and Delusional Symptoms in Schizophrenia. Arch Gen Psychiatry. 2010 Nov;67(12):1246–1254.

[546] Hall J, Whalley HC, McKirdy JW, Romaniuk L, McGonigle D, McIntosh AM, et al. Overactivation of fear systems to neutral faces in schizophrenia. Biol Psychiatry. 2008 Jul;64(1):70–3. Notes Superb.

[547] Miwa H, Fuwa T, Nishi K, Kondo T. Subthalamo-pallido-striatal axis: a feedback system in the basal ganglia. Neuroreport. 2001 Dec;12(17):3795–8.

[548] Burkhardt JM, Constantinidis C, Anstrom KK, Roberts DCS, Woodward DJ. Synchronous oscillations and phase reorganization in the basal ganglia during akinesia induced by high-dose haloperidol. Eur J Neurosci. 2007 Oct;26(7):1912–24.

[549] Gage GJ, Stoetzner CR, Wiltschko AB, Berke JD. Selective activation of striatal fast-spiking interneurons during choice execution. Neuron. 2010 Aug;67(3):466–79.

[550] Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science. 2003 Mar;299(5614):1898–902. Available from: `http://www.sciencemag.org/cgi/content/full/299/5614/1898`.

[551] Lynch MR. Schizophrenia and the D1 receptor: focus on negative symptoms. Prog Neuropsychopharmacol Biol Psychiatry. 1992 Jan;16(6):797–832.

[552] Laruelle M, Kegeles LS, Abi-Dargham A. Glutamate, dopamine, and schizophrenia: from pathophysiology to treatment. Annals of the New York Academy of Sciences. 2003 Nov;1003:138–58.

[553] Stefanatos GA, Wasserstein J. Attention deficit/hyperactivity disorder as a right hemisphere syndrome. Selective literature review and detailed neuropsychological case studies. Annals of the New York Academy of Sciences. 2001 Jun;931:172–95.

[554] Joshua M, Adler A, Rosin B, Vaadia E, Bergman H. Encoding of probabilistic rewarding and aversive events by pallidal and nigral neurons. Journal of Neurophysiology. 2009 Feb;101(2):758–72.

[555] Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O. A pallidus-habenula-dopamine pathway signals inferred stimulus values. Journal of Neurophysiology. 2010 Aug;104(2):1068–76.

[556] Osuch EA, Bluhm RL, Williamson PC, Théberge J, Densmore M, Neufeld RWJ. Brain activation to favorite music in healthy controls and depressed patients. Neuroreport. 2009 Aug;20(13):1204–8.

[557] Andreasen NC, O'Leary DS, Flaum M, Nopoulos P, Watkins GL, Ponto LLB, et al. Hypofrontality in schizophrenia: distributed dysfunctional circuits in neuroleptic-naïve patients. Lancet. 1997 Jun;349(9067):1730–4.

[558] Lawrie SM, McIntosh AM, Hall J, Owens DGC, Johnstone EC. Brain structure and function changes during the development of schizophrenia: the evidence from studies of subjects at increased genetic risk. Schizophrenia Bulletin. 2008 Mar;34(2):330–40.

[559] Stone JM, Bramon E, Pauls A, Sumich A, McGuire PK. Thalamic neurochemical abnormalities in individuals with prodromal symptoms of schizophrenia - relationship to auditory event-related potentials. Psychiatry Res. 2010 Aug;183(2):174–6.

[560] Broome MR, Matthiasson P, Fusar-Poli P, Woolley JB, Johns LC, Tabraham P, et al. Neural correlates of executive function and working memory in the 'at-risk mental state'. Br J Psychiatry. 2009 Jan;194(1):25–33.

[561] Weiss EM, Golaszewski S, Mottaghy FM, Hofer A, Hausmann A, Kemmler G, et al. Brain activation patterns during a selective attention test-a functional MRI study in healthy volunteers and patients with schizophrenia. Psychiatry Res. 2003 May;123(1):1–15.

[562] Fusar-Poli P, Howes OD, Allen P, Broome M, Valli I, Asselin MC, et al. Abnormal frontostriatal interactions in people with prodromal signs of psychosis: a multimodal imaging study. Archives of General Psychiatry. 2010 Jul;67(7):683–91.

[563] Keshavan MS, Delisi LE, Seidman LJ. Early and broadly defined psychosis risk mental states. Schizophr Res. 2010 Nov;.

[564] Gourzis P, Katrivanou A, Beratis S. Symptomatology of the initial prodromal phase in schizophrenia. Schizophr Bull. 2002 Jan;28(3):415–29.

[565] Kita JM, Parker LE, Phillips PEM, Garris PA, Wightman RM. Paradoxical modulation of short-term facilitation of dopamine release by dopamine autoreceptors. J Neurochem. 2007 Aug;102(4):1115–24.

[566] Orosz AT, Feldon J, Simon AE, Hilti LM, Gruber K, Yee BK, et al. Learned Irrelevance and Associative Learning Is Attenuated in Individuals at Risk for Psychosis but not in Asymptomatic First-Degree Relatives of Schizophrenia Patients: Translational State Markers of Psychosis? Schizophr Bull. 2010 Jan;.

[567] Zhang L, Doyon WM, Clark JJ, Phillips PEM, Dani JA. Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. Mol Pharmacol. 2009 Aug;76(2):396–404.

[568] Howes OD, Montgomery AJ, Asselin MC, Murray RM, Valli I, Tabraham P, et al. Elevated striatal dopamine function linked to prodromal signs of schizophrenia. Archives of General Psychiatry. 2009 Jan;66(1):13–20.

[569] Glenthoj BY, Mackeprang T, Svarer C, Rasmussen H, Pinborg LH, Friberg L, et al. Frontal dopamine D(2/3) receptor binding in drug-naive first-episode schizophrenic patients correlates with positive psychotic symptoms and gender. Biol Psychiatry. 2006 Sep;60(6):621–9.

[570] Talvik M, Nordström AL, Okubo Y, Olsson H, Borg J, Halldin C, et al. Dopamine D2 receptor binding in drug-naïve patients with schizophrenia examined with raclopride-C11 and positron emission tomography. Psychiatry Res. 2006 Dec;148(2-3):165–73.

[571] Winterer G, Weinberger D. Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends in Neurosciences. 2004 Nov;27(11):683–690.

[572] Tuppurainen H, Kuikka JT, Laakso MP, Viinamäki H, Husso M, Tiihonen J. Midbrain dopamine D2/3 receptor binding in schizophrenia. Eur Arch Psychiatry Clin Neurosci. 2006 Sep;256(6):382–7.

[573] Floresco SB, West AR, Ash B, Moore H, Grace AA. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. Nat Neurosci. 2003 Sep;6(9):968–973.

[574] Bassett AS, Bury A, Honer WG. Testing Liddle's three-syndrome model in families with schizophrenia. Schizophrenia Research. 1994 Jun;12(3):213–21.

[575] Lindenmayer JP, Bernstein-Hyman R, Grochowski S. A new five factor model of schizophrenia. The Psychiatric quarterly. 1994 Jan;65(4):299–322.

[576] Lindenmayer JP, Grochowski S, Hyman RB. Five factor model of schizophrenia: replication across samples. Schizophrenia Research. 1995 Feb;14(3):229–34.

[577] Dollfus S, Everitt B, Ribeyre JM, Assouly-Besse F, Sharp C, Petit M. Identifying subtypes of schizophrenia by cluster analyses. Schizophr Bull. 1996 Jan;22(3):545–55.

[578] Dollfus S, Everitt B. Symptom structure in schizophrenia: two-, three- or four-factor models? Psychopathology. 1998 Jan;31(3):120–30.

[579] Smith DA, Mar CM, Turoff BK. The structure of schizophrenic symptoms: a meta-analytic confirmatory factor analysis. Schizophrenia Research. 1998 May;31(1):57–70.

[580] Lykouras L, Oulis P, Daskalopoulou E, Psarros K, Christodoulou GN. Clinical subtypes of schizophrenic disorders: a cluster analytic study. Psychopathology. 2001 Jan;34(1):23–8.

[581] Levine SZ, Rabinowitz J. Revisiting the 5 dimensions of the Positive and Negative Syndrome Scale. Journal of Clinical Psychopharmacology. 2007 Oct;27(5):431–6.

[582] Peralta V, Cuesta MJ. How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. Schizophrenia Research. 2001 Apr;49(3):269–85.

[583] Andreasen N. Scale for the Assessment of Negative Symptoms. Iowa City, Iowa; 1984.

[584] Andreasen N. Scale for the Assessment of Positive Symptoms. Iowa City, Iowa; 1984.

[585] Suhr JA, Spitznagel MB. Factor versus cluster models of schizotypal traits. II: relation to neuropsychological impairment. Schizophr Res. 2001 Dec;52(3):241–50.

[586] Kircher TT, Liddle PF, Brammer MJ, Williams SC, Murray RM, Mcguire P. Neural correlates of formal thought disorder in schizophrenia: preliminary findings from a functional magnetic resonance imaging study. Arch Gen Psychiatry. 2001 Aug;58(8):769–74.

[587] Williams LM. Cognitive inhibition and schizophrenic symptom subgroups. Schizophr Bull. 1996 Jan;22(1):139–51.

[588] Durieux PF, Bearzatto B, Guiducci S, Buch T, Waisman A, Zoli M, et al. D2R striatopallidal neurons inhibit both locomotor and drug reward processes. Nat Neurosci. 2009 Apr;12(4):393–5.

[589] Tomelleri L, Jogia J, Perlini C, Bellani M, Ferro A, Rambaldelli G, et al. Brain structural changes associated with chronicity and antipsychotic treatment in schizophrenia. European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology. 2009 Dec;19(12):835–40.

[590] Middleton FA, Strick PL. Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res Brain Res Rev. 2000 Mar;31(2-3):236–50.

[591] Middleton FA, Strick PL. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. Brain and Cognition. 2000 Mar;42(2):183–200.

[592] Büchel C, Wise RJ, Mummery CJ, Poline JB, Friston KJ. Nonlinear regression in parametric activation studies. Neuroimage. 1996 Aug;4(1):60–6.

[593] Barton JJS, Pandita M, Thakkar K, Goff DC, Manoach DS. The relation between antisaccade errors, fixation stability and prosaccade errors in schizophrenia. Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale. 2008 Mar;186(2):273–82.

[594] Winograd-Gurvich C, Fitzgerald PB, Georgiou-Karistianis N, Millist L, White O. Inhibitory control and spatial working memory: a saccadic eye movement study of negative symptoms in schizophrenia. Psychiatry Res. 2008 Jan;157(1-3):9–19.

[595] Franke C, Arndt D, Ploner CJ, Heinz A, Reuter B. Saccade generation and suppression in schizophrenia: effects of response switching and perseveration. Psychophysiology. 2008 Sep;45(5):698–704.

[596] Hill SK, Reilly JL, Harris MSH, Khine T, Sweeney JA. Oculomotor and neuropsychological effects of antipsychotic treatment for schizophrenia. Schizophrenia Bulletin. 2008 May;34(3):494–506.