MATR. AIDR02/18802

CAMPUS BIOMEDICO UNIVERSITY OF ROME

DEPARTMENT OF ENGINEERING

UNIVERSITY OF BARI ALDO MORO

DEPARTMENT OF TRANSLATIONAL BIOMEDICINE AND NEUROSCIENCE

Italian National Ph.D. in Artificial Intelligence
Health and Life Sciences

XXXVIII Cycle

Implementing machine-learning techniques to study inter-individual variability in fMRI-related brain cognitive functions

Supervisors

Prof. Giulio Pergola

Prof. Linda A. Antonucci

Prof. Loredana Bellantuono

Candidate

Antonella Lupo

January, 2026

Contents

List of figures

List of tables

Acronyms

ACC anterior cingulate cortex

ADHD attention-deficit/hyperactivity disorder

AI artificial intelligence

AN anterior nucleus

BOLD blood-oxygen-level-dependent

CBF cerebral blood flow

CBV cerebral blood volume

CGT Cambridge Guessing Task

$CMRO_2$ cerebral metabolic rate of oxygen consumption

CPT Continuous Performance Test

CS conditioned stimulus

DA Dopamine

DMN default mode network

FC functional connectivity

FEP first episode of psychosis

fMRI Functional magnetic resonance imaging

FPN frontoparietal network

GLM general linear model

HC healthy controls

ICA Independent Component Analysis

INT intermediate

LD lateral dorsal nucleus

LME linear mixed-effects model

MD mediodorsal nucleus

MID Monetary Incentive Delay

ML machine learning

MNI Montreal Neurological Institute

MTL medial temporal lobe

MTT multiple trace theory

NAc Nucleus accumbens

NBR negative BOLD response

OFC Orbitofrontal cortex

PALP Passive Avoidance Learning Paradigm

PBR positive BOLD response

PCA Pavlovian conditioned approach

PET positron-emission tomography

PFC prefrontal cortex

PIT Pavlovian-to-instrumental transfer

ROI region of interest

rs-fMRI Resting-state fMRI

SCZ schizophrenia

SEM structural equation modeling

SMA supplementary motor area

ST/GT Sign-Tracker/Goal-Tracker

SURPS Substance Use Risk Profile Scale

TCI-R Temperament and Character Inventory-Revised

UC unconditioned stimulus

VAS Visual Analog Scales

VMAC value-modulated attentional capture

vmPFC ventromedial prefrontal cortex

VStr Ventral striatum

VTA ventral tegmental area

WCST Wisconsin Card Sorting test

Abstract

In the field of cognitive neuroscience, key discoveries that form the foundation of introductory textbooks—such as the idea of brain localization—initially stemmed from N-of-1 studies. Notable examples include Phineas Gage, whose accident offered insights into brain function; H.M., who highlighted the important role of the hippocampus in memory; and the discovery of Broca's Area's involvement in language. As the field grew, neuroimaging techniques became essential for most studies that used experimental paradigms and statistical methods that averaged data across individuals. Although this approach has helped develop general theories about which brain regions or systems are involved in specific cognitive functions, it assumes that all humans process information in the same way—an assumption that may overlook the rich variety of human cognitive differences.

Therefore, one of the primary challenges in modern cognitive neuroscience is to understand how brain mechanisms contribute to individual differences in human behavior. Evidence consistently shows significant variability between individuals in functional brain pathways, reflecting stable and meaningful neural fingerprints that are linked to behavior, personality traits, and vulnerability to psychiatric disorders. Building on this foundation, the current work employs multiple experimental and statistical methods, combining task-based and resting-state fMRI with machine learning, to investigate individual differences in memory (Study 1) and reward processing (Study 2).

In the first study, I focused on interindividual variability in cortico-thalamic engagement impacting learning and memory in healthy individuals. I hypothesized that differences in thalamic recruitment within the frontoparietal networks influence memory performance and that thalamic activity could alter cortical configurations during resting states, aiding learning. I analyzed two independent groups of healthy adults (N1 = 29, N2 = 108) who underwent a multi-scan fMRI protocol, comprising a resting-state scan and an associative memory task. Through personalized mapping, structural equation modeling, and clustering of brain responses, I found that associative memory encoding relies on specific corticothalamic pathways, involving the recruitment of the medial thalamic subdivision and the suppression of the anterior one, to facilitate the successful encoding of new memories.

Once this methodology was validated, I used a clustering algorithm in the second Study to translate the preclinical model of Sign-Tracker/Goal-Tracker (ST/GT) to humans, as it evidenced individual differences in cue- versus outcome-driven responses. I collected fMRI data from 1,203 individuals across four cohorts of participants performing a monetary reward task (N1 = 890, N2 =

245, N3 = 34, N4 = 34). Clustering of reward-related BOLD activity consistently identified ST- and GT-like profiles, replicated across healthy cohorts with over 78% accuracy. I examined their relationship with responses to antipsychotics and psychopathology, finding that D_2 antagonists selectively reduced anticipatory striatal responses in ST but not GT. Finally, in a clinical cohort of schizophrenia patients, I observed that GT-like patients exhibited greater severity of negative symptoms, supporting a link between reduced ventral striatal activity and negative symptoms, and a positive correlation between D_2 receptor affinity and negative symptom severity.

Together, these studies show that significant interindividual variability exists in brain function associated with behavior, thereby providing a potential translational framework for clinical applications and patient stratification. The methodological and conceptual framework created in this thesis provides the foundation for fMRI-based clinical predictions and more tailored treatments.

Chapter 1

1.1. General Introduction

Individuals exhibit a variety of behaviors in response to environmental stimuli, thereby solving cognitive challenges in everyday life. For example, fleeing from a building on fire is an adaptive behavior that sustains survival. A less tragic example may be found in the cognitive processes involved in the reading of the current thesis, including visual processing, lexical decoding, grammatical and semantic comprehension, working memory, long-term memory, meaning building, attention, and critical thinking evaluation. Indeed, behaviors have neural correlates, meaning that every action, thought, or emotion people experience has a physical and biological basis in the brain, reflecting the varying engagement of diverse brain circuits that often interact with each other. Neuroscience, particularly the cognitive branch, examines this relationship, analyzing how the different areas of the brain and their interactions influence the way we act, think, and feel.

Previous research relying on group-average neuroimaging analyses in neuroscience has significantly advanced our understanding of core brain regions and canonical functional networks in different cognitive and affective domains (Cole et al., 2016; Biswal et al., 1995; Biswal et al., 2010; Fox et al., 2005; Yeo et al., 2011; Smith et al., 2009; Ranganath et al., 2003; Eichenbaum, Yonelinas & Ranganath, 2007; Knutson & Greer, 2008; Dolcos, Iordan & Dolcos, 2011; Crossley et al., 2013; Dosenbach et al., 2007). However, these approaches inherently assume homogeneity across individuals, potentially obscuring meaningful interindividual differences in neural organization and behavior (Seghier & Price, 2018; Gordon et al., 2017; Hermosillo et al., 2024; Du et al., 2024; Mattoni et al., 2025). By averaging brain patterns across subjects, group-level methods can mask idiosyncratic activation patterns, variable network configurations, and subject-specific strategies used during cognitive tasks. This limitation is particularly relevant when investigating complex cognitive and

affective processes that vary significantly between individuals, especially when considering the impact of potential impairments in these domains in clinical populations.

1.1. Beyond average group brain: interindividual variability in neuroscience

Over the past four decades, cognitive and affective neuroscience have witnessed significant advancements, thanks in part to neuroimaging, particularly functional magnetic resonance imaging (fMRI). This technique offers high spatial resolution and has become one of the preferred methods, alongside electroencephalography, for conducting in vivo neuroscience studies in humans. These studies are based on the blood-oxygen-level-dependent (BOLD) signal, which arises from increases in cerebral blood flow and local hemoglobin oxygenation in response to neuronal activity (Logothetis & Wandell, 2004), first studied by Seiji Ogawa (Ogawa et al., 1990). This method has widely been used in cognitive neuroscience to map brain activity associated with sensory, motor, language, memory, and emotional processing through task-based paradigms, providing insight into the neural substrates of these functions, with subsequent studies incorporating resting-state fMRI (rs-fMRI; Biswal et al., 1995; Fox et al., 2014; Rosazza & Minati, 2011). During the resting state, the participant refrains from performing specific tasks and is instructed to stay still, with eyes either closed or open and fixating on a cross. This setup allows recording of spontaneous brain blood flow fluctuations that indirectly indicate neuronal activity (Logothetis et al., 2001). In fact, it is well known that the brain under normal physiological conditions—such as when at rest—is never idle but always remains neuro-electrically and metabolically active. Functional connectivity (FC) refers to the temporal coherence in activation patterns of anatomically separated brain regions, i.e., their synchronized activity over time (Friston et al., 1993; Friston, 1994; Van den Heuvel and Hulshoff Pol, 2010), and it can be studied during both the performance of active tasks, such as finger tapping or visual stimulation, as well as during resting state. The rs-fMRI approach has revolutionized the study of neural architecture, as it allows us to

study functional interaction between distributed brain regions independent of neurocognitive tasks, overcoming the task-based limitation of studying one process at a time (Cole et al. 2016, Passiatore et al., 2021), with experimental tasks that sometimes may be very simple and with little ecological validity.

Recordings of spontaneous fluctuations in activation patterns during rest (Biswal et al., 1997; Greicius et al., 2003; Lowe et al., 2000) have been used to understand how the brain coordinates its various functions, leading to substantial progress in understanding human brain functioning and organization. Rs-fMRI shows moderate heritability (heritability estimates typically around 20-40%; Barber et al., 2021; Adhikari et al., 2018; Glahn et al. 2010; Yang et al. 2016; Ge et al. 2017) and may predict individual differences in behavior (Hampson et al. 2006; van den Heuvel et al. 2009; Finn et al. 2015; Smith et al. 2015), as it found its application in the study of a wide range of cognitive and behavioral processes, such as attention, memory, language, and emotions (Biswal et al., 1995; 1997; Cordes et al., 2000; Damoiseaux et al., 2006; Fox and Raichle, 2007; Greicius et al., 2003; Lowe et al., 2000; Van den Heuvel and Hulshoff Pol, 2010). Consequently, rs-fMRI has been widely utilized to estimate population-average functional brain networks by averaging data across multiple subjects (Beckmann et al. 2005; Damoiseaux et al. 2006; Fox et al. 2006; Dosenbach et al. 2007; Margulies et al. 2007; Power et al. 2011; Yeo et al. 2011; Lee et al. 2012), such as the 'default mode' network (DMN). The DMN is particularly active during rest and preferentially activates when individuals focus on internal tasks such as daydreaming, envisioning the future, retrieving memories, and taking others' perspectives (Raichle, 2015; Greicius et al., 2003; 2004). Moreover, rs-fMRI has increasingly emerged as a neuroimaging technique in clinical populations, as it is not demanding for participants, particularly those whose task execution may be challenging (Fox and Greicius, 2010). It has been used to identify unique patterns of FC in individuals, potentially predicting individual differences in cognitive abilities and clinical symptoms (Liu et al., 2018; Reineberg et al., 2015; Keller et al., 2023; Cole et al., 2011; Finn & Constable, 2016). For example, in patients with psychosis, it is possible to observe both a

disconnection (reduced connectivity) between brain regions that should be connected and hyperconnectivity (increased connectivity) between regions that normally should not be strongly connected (Li et al., 2019; Sapienza et al., 2023).

Although the rs-fMRI approach offers valuable insights into brain organization, there are significant limitations in measuring FC with it (Fox et al., 2012). First, because participants are not performing a specific task, there is no clear measure of performance or validation that participants are in a particular cognitive and/or emotional state. Another limitation is the correlational nature of the data, which restricts the conclusions that can be drawn about the processes studied. Additionally, rs-fMRI studies are susceptible to physiological noise, including heart rate, respiration, and motion-related artifacts (Doubovikov & Aksenov, 2025; Buckner et al., 2013; Drew et al., 2020; Yoshikawa et al., 2020). Finally, the low signal-to-noise ratio in rs-fMRI, especially with short scan durations (Elliot et al., 2020), can affect reliability and limit the ability to identify FC associations with traits (Bennett & Miller, 2013; Marek et al., 2022; Choi & Ogawa, 2022). To overcome these limitations, integrating rs-fMRI and task-based analyses allows researchers to assess how individual brain regions and networks are functionally connected both at rest and during cognitive or motor tasks, providing an innovative understanding of the brain-behavior relationships and potential biomarkers for neurological and psychiatric conditions (Rabini et al., 2023; Passiatore et al., 2023; Harrewijn et al., 2020; Lee & Shimony, 2013). For example, combining both approaches may help identify which networks are most relevant to specific tasks and how they contribute to task performance (Passiatore et al., 2021; Wagner et al., 2019). Combining both rest and task fMRI data has been used for studying group differences in FC between patients with schizophrenia (SCZ) and controls (Arbabshirani & Calhoun, 2011; Cetin et al., 2014), as well as for single-subject prediction (Rashid et al., 2016).

Although combining the study of intrinsic brain networks with their engagement during specific tasks has the advantage of better understanding the brain-behavior relationship, this approach

still identifies group-level central tendencies that aim to generalize the findings across individuals, focusing on the evaluation of the group-average brain. Although the group-averaging approach has demonstrated many fundamental principles of functional brain organization, it obscures subject-specific features. Additionally, a wealth of scientifically and clinically relevant information is often hidden and potentially invalidated by averaging data across subjects. In fact, two human brains are never identical. Detailed characterization of individual brains has so far been limited (Laumann et al., 2015; Poldrack et al., 2015).

Moreover, when different individuals are asked to perform the same task, they show markedly different patterns of brain activity. Often, interindividual variability in brain activity resulting from an experimental task is attributed to two primary factors. First, it may be due to differences in overall brain structure. The vast majority of brain imaging studies rely on the spatial alignment of different brains to a standard brain template (registration) to account for between-subject discrepancies in gross anatomy (Jenkinson & Smith, 2001). Second, in conventional neuroimaging studies, it is hypothesized that tasks are performed in a similar manner or via a single cognitive strategy (Hedge et al., 2018). However, it should be noted that many tasks are unconstrained, allowing subjects to adopt their own strategies, which in turn may involve different brain circuits or different engagement of the same brain circuits. These individual variations are seen in all behavioral domains (Besle et al., 2013; McNab & Klingberg, 2008; Mukai et al., 2007; Tom et al., 2007; Wig et al., 2008).

This variability is often treated as 'noise' and discarded through the process of averaging individual responses to find central tendencies, i.e., mean and median (Kanai et al., 2011). Importantly, it is thought that such individual differences are mainly explained by unstable factors such as task strategy or compliance related to the behavior (Tavor et al., 2016).

Figure I.1 offers a visual representation of the problem of averaging individual activation maps that vary in their characteristics (Thirion et al., 2007) when subjects adopt different strategies. The

resulting average in this toy example depicts a hybrid image that differs from that encoded in each of the original images. This average image contains a multitude of false negatives (where features vary between images) and false positives (where feature combinations create new features).

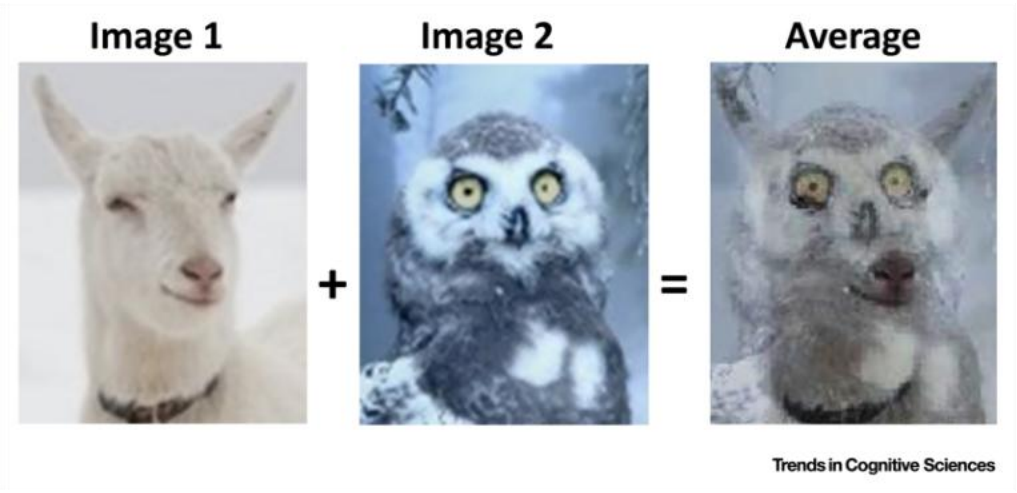

Figure I. 1 Averaging images with variable features. Figure from Seghier & Price, 2018.

The importance of treating variance as meaningful data rather than noise has started to be recognized by the neuroimaging community, facing the challenges of adapting and rethinking the most widely used analysis software packages originally built to provide estimates of group effects, such as Statistical Parametric Mapping (SPM) version 12 (http://www.fil.ion.ucl.ac.uk/spm), FSL (https://fsl.fmrib.ox.ac.uk/fsl/docs/#/), Analysis of Functional NeuroImages (AFNI, https://afni.nimh.nih.gov/), and CONN Toolbox (https://web.conn-toolbox.org/), which have progressively incorporated features for subject-level modeling, reliability assessment, and individualized connectivity analysis (https://www.statsinimaging.org/software_neuro/).

1.2. Neuroimaging-based individualized approaches: a promise for translational neuroscience

The following paragraph will examine some precision functional mapping techniques, including personalized brain networks and stratification methods, which have been developed as an alternative to using group-level atlases. These techniques are used to derive individually defined networks that capture each unique pattern of brain functional topography (Gordon et al., 2017; Keller et al., 2023).

By spatially registering and combining small amounts of data from many individuals using univariate analytical techniques, neuroscientists have been able to develop theories and identify common trends of both task-based (Sylvester et al., 2003; Rugg and Vilberg, 2013, Wang, Smith & Delgrado, 2016) and resting-state activity (Beckmann et al., 2005, Power et al., 2011, Smith et al., 2009, Yeo et al., 2011), either detecting brain features showing significant group differences between healthy controls (HC) and patients with brain disorders using statistical inferences, or establishing the brain-behavior relationship using correlational analysis (Sui et al., 2020). However, most neuroimaging findings from published studies are not easily translated into practical tools. Partly because of a lack of out-of-sample validation, group-level inferences often overfit both the signal and noise in a particular dataset. Therefore, the generalizability of such group-level findings to new individuals is still unknown (Arbabshirani et al., 2017). Additionally, traditional univariate research concentrates on explaining the neural basis of observed behavior rather than predicting future behavior from neuroimaging signatures (Gao, Calhoun, & Sui, 2018). Prediction is a crucial step toward a translational neuroscience era (Yarkoni & Westfall, 2017).

Although population-average networks provided important insights into the functional organization of the human brain (Buckner et al. 2013; Wig 2017), neural pathways vary between individuals, and population-average networks might obscure individual-specific network organizations (Beckmann et al. 2009; Bellec et al. 2010; Hacker et al. 2013; Wig et al. 2014; Chong et al. 2017).

One of the most promising directions in this domain has been the development of personalized brain networks, which are constructed by estimating functional or structural connectivity at the individual level rather than relying on group templates. Such personalized functional brain networks have been found to be highly stable within individuals and to predict an individual's spatial pattern of activation on fMRI tasks (Gordon et al., 2017; Glasser et al., 2016; Tavor et al., 2016; Cui et al., 2022). Numerous studies have demonstrated that individual-specific network topography (i.e., location and spatial arrangement) is behaviorally relevant. For instance, Kong and colleagues (2018) demonstrated that the spatial topography of individual brain networks can serve as a universal fingerprint of human behavior across cognition, personality, and emotion. In a recent study, Keller et al. (2023) found that functional network topography heterogeneity in 9–10-year-old children is associated with individual differences in cognition before the transition into adolescence. Moreover, Rasero et al. (2021) demonstrated that individual differences in brain structure and function significantly improve the prediction of cognitive traits, suggesting that interindividual variability reflects meaningful neurobiological mechanisms rather than random noise. The variability in the topography of functional brain networks, particularly in the associative cortex, is also implicated in a broad range of psychiatric disorders (Cui et al., 2022).

An approach in personalized translational neuroscience is stratifying individuals by neurobiological features, identifying subgroups with similar brain patterns using imaging data like FC or task activity to differentiate or predict clinical outcomes. Numerous neuroimaging studies have shown significant differences in brain patterns between subjects who relied upon different cognitive strategies during the execution of the same task despite similar behavioral performances (Miller et al., 2018; Miller et al., 2002; Sanfratello et al., 2014; Speer et al., 2003; Seghier & Price, 2009; Kherif et al., 2009; Mattoni et al., 2025). For instance, Miller et al. (2012) found that individual differences in cognitive style and encoding strategy were significant factors in explaining variability in individual

patterns of brain activity during a memory retrieval task. The greater the differences were in individuals' visualizations of word stimuli, the greater the differences were in their brain activity patterns across the brain.

The stratification of healthy individuals is especially important for establishing a baseline to understand how the brain functions in healthy people and to shed light on its significance in clinical conditions (Owen et al., 2014; Marquand et al., 2016). Indeed, an increasing number of studies used data-driven methods to stratify many disorders, including SCZ, bipolar disorder, major depression, Huntington's disease, Alzheimer's diseases (Khan et al., 2022), attention-deficit/hyperactivity disorder (ADHD), and autism based on many types of data, including symptoms, neuropsychologic scores, and neuroimaging measures (Polosecki et al., 2020; Cheng et al., 2013; Warren et al., 2024; Karalunas et al., 2014; Brodersen et al., 2014; Fair et al., 2012; gates et al., 2014; Dias et al., 2015; Van Loo et al., 2012; Pattyn et al., 2015; Sacco et al., 2012; Wang et al., 2020). For instance, Ma et al. (2025) analyzed a large sample of multimodal data, including genetic, environmental, neuroimaging, and symptom data, and discovered a caudate pathway linking the environmental stress to the psychotic symptoms in a subgroup characterized by smaller putamen, but not when SCZ was considered as a whole group. Their findings underscore the significance of neuroimaging-based stratification in enhancing the identification of mediators of gene-by-environment interactions in the etiology of SCZ (Ma et al., 2025).

Neuroimaging-based subtyping is increasingly used to explain the high phenotypic and neural heterogeneity underlying psychiatric disorders (Zhang et al., 2023; Ma et al., 2025; Vaskinn et al., 2020; Wolfers et al., 2018; Carroll & Owen, 2009). A recent study by Shao et al. (2022) identified neurobiological features between cognitive subtypes in the early course of SCZ using latent profile analysis and the amplitude of low-frequency fluctuations to measure brain spontaneous neural activity. Their results suggested that differences in the spontaneous neural activity of the precentral gyrus

(PreCG) and posterior cingulate cortex/precuneus (PCC/PCu) might be the potential neurobiological features of the cognitive subtypes in the early course of SCZ, which may deepen our understanding of the role of those brain areas in the pathogenesis of cognitive impairment in SCZ.

To summarize, this section has outlined how precision functional mapping and neuroimaging-based stratification serve as powerful alternatives to standard group-level atlases. By characterizing unique functional topographies and identifying biologically meaningful subgroups, these methods provide a robust framework to uncover biologically meaningful subtypes that may respond differently to treatment or exhibit distinct cognitive profiles (Marquand et al., 2016; Owen et al., 2014; Schumann et al., 2014; Zhang et al., 2023; Chand et al., 2020; Shao et al., 2022; Wen et al., 2025; Brucar et al., 2023).

1.3. Methodological Approaches to Studying Interindividual Variability in Neuroimaging

Explaining and predicting individual behavioral differences in brain patterns induced by clinical and social factors is one of the most promising applications of neuroimaging. In recent years, the use of traditional univariate brain mapping techniques for biomarker discovery has been replaced by multivariate predictive models for individuals (Woo et al., 2017). The large amount of data and the multidimensional nature of the phenomena have driven the need for approaches such as machine learning (ML) and artificial intelligence (AI) to systematically investigate interindividual variability in neuroimaging data. Unlike conventional approaches, ML–based methods can establish integrated brain models by considering the multivariate nature of diverse brain function or structure measurements across the entire brain (Kambeitz et al., 2015), thereby raising the opportunity to analyze neuroimaging data at the single-subject level (Osuch et al., 2018). In line with Breiman (2001) and Arlot & Celisse (2010), the model that extracts patterns is known as a training algorithm, and the pattern it extracts is referred to as a prediction model.

Supervised learning has been successful for predicting diagnosis or outcome from neuroimaging data in research settings (Orru et al., 2012; Wolfers et al., 2015; Klöppel et al., 2012) but is fundamentally limited by the quality of the clinical labels and the heterogeneity within disease cohorts (Wolfers et al., 2015) and cannot, by definition, inform on the validity of the labels. Therefore, there has been an emerging trend in using regression-based ML approaches to discover individual differences in cognitive functioning or disease severity from brain imaging data: that is, the individualized prediction of human behaviors (Finn et al., 2015). Instead of determining group membership through binary classification, behavior prediction may consider the problem of quantitatively estimating the specific scores for a continuous behavioral measure over the whole range of the variable (Shen et al., 2017). Given the high-dimensional nature of neuroimaging data, these approaches are commonly accompanied by a feature selection step to obtain low-dimensional representations (Bermingham et al., 2015).

Typically, predictive modeling of FC is one approach that combines simple linear regression and feature selection to predict individual differences in traits and behavior from connectivity data (Finn et al., 2015; Shen et al., 2017; Rosenberg et al., 2016). This procedure has several advantages, including straightforward interpretation, fast computation, and robust generalization, and it has been successfully employed for predicting multiple human behaviors (Rosenberg et al., 2016; Greene et al., 2018; Lake et al., 2019; Hsu et al., 2018). For instance, Finn et al. (2015), by using Human Connectome Project data from 126 subjects, each scanned during six separate sessions over 2 days, demonstrated that intrinsic FC architecture of individuals act as a 'fingerprint' that can accurately identify subjects from a large group, and that is both reliable enough across sessions and distinct enough from that of other individuals to identify him or her from the group regardless of how the brain is engaged during imaging. Furthermore, they demonstrated the relevance of these connectivity profiles to behavior by showing, through a fully cross-validated analysis, that FC profiles can predict the core cognitive trait

of fluid intelligence in individuals. Finn and colleagues (2015) laid the groundwork for studying neuroscience by examining how personal functional brain organization connects to various behavioral phenotypes.

When studying brain-behavior relationships, one limitation to consider is that there is often no direct causal relationship between a predictor and the outcome. Instead, it may very well depend on factors such as environmental, biological, psychological, or social elements. Building on regression frameworks, mediation models are particularly valuable in neuroimaging because they enable the identification of neural pathways that may underlie behavioral or clinical phenomena, thus investigating the mechanisms through which individual differences exert their effects. Specifically, structural equation modeling (SEM) analyzes complex relationships between multiple observed and unobserved (latent) variables simultaneously estimating a system of linked regression equations to assess how well a proposed model, such as a mediation model, fits the data (Gunzler et al., 2013). For instance, mediation models have been used to demonstrate that the effect of early-life stress on emotional regulation is mediated by amygdala reactivity and prefrontal connectivity (Guadagno et al., 2021). Modern implementations of mediation analysis often incorporate bootstrapping techniques to enhance statistical robustness, particularly in neuroimaging datasets where sample sizes are limited and data dimensionality is high. Moreover, serial and parallel mediation models allow for the exploration of more complex causal chains, providing a nuanced understanding of how multiple brain regions interact to mediate psychological processes.

So far, regression models have assumed that the dependent variable is given, which is the case in supervised ML techniques. If the dependent variable cannot (easily) be defined and measured, unsupervised learning, including clustering and dimensionality reduction techniques (e.g., principal component analysis, PCA), may provide some insight (Janssen et al., 2018). Particularly, clustering algorithms group individuals based on similarities in their neuroimaging profiles, such as patterns of

FC, regional activation, or multimodal features. This technique is particularly useful for identifying subtypes within heterogeneous populations, such as patients with psychiatric or neurological disorders, or for uncovering distinct cognitive strategies among healthy individuals. For instance, in a recent study, Nakuci and colleagues (2025) identified distinct and stable clusters in the leveraged trial-to-trial fluctuations of the brain-wide fMRI signal across three perceptual decision-making experiments using modularity-maximization, a data-driven classification method. Their results suggest that the same task can be accomplished in the presence of widely varying brain activation patterns.

Moreover, the application of ML to neuroimaging data in healthy populations has paved the way for applying these methods to psychiatric disorders. For instance, clustering has been successfully applied to stratify individuals with autism spectrum disorder based on connectivity patterns (Buch et al., 2024), to identify subgroups of depression with distinct neural signatures (Wang et al., 2021), and to differentiate learning styles in reinforcement learning tasks (Alkam et al., 2025; Sarakam, 2025). Recently, Schnack et al. (2019) reviewed the possible contributions that ML techniques could make to explore the heterogeneous nature of psychiatric disorders with a focus on SCZ. Dickinson et al. (2017) applied a 2-step clustering algorithm to divide a large sample of SCZ patients (N = 549) into "deficit", "distress", and "low symptom" subgroups, based on two composite scores of a clinical scale, i.e., PANSS. The patients in the "deficit" and "distress" groups showed both shared and distinct deviations in clinical and behavioral variables when compared to the "low symptom" group. A subset (N = 182) of the patients underwent fMRI scanning with the N-back working memory paradigm. The "deficit" subgroup showed a different activation profile than the two other groups, indicating a possible different etiology. Clustering techniques have been applied to subtype a large sample of patients across the psychosis spectrum (SCZ, schizoaffective, bipolar with psychosis), their relatives, and controls (Clementz et al., 2016; Dwyer et al., 2017).

The recent ML applications in neuroimaging-based single-subject prediction of brain disorders represent an important advance toward the development of biomarkers for psychiatric diseases and the stratification of heterogeneous populations to find different subtypes of individuals. Nevertheless, these studies present some important limitations that should be carefully considered. For example, most studies rely on relatively small samples, which, combined with publication bias—i.e., reporting the model with the highest accuracy in the training dataset—may result in overly optimistic estimates of generalization accuracy (Varoquaux, 2018; Mendelson et al., 2017). Typically, when testing the model's best generalizability in an independent test sample, a drop in accuracy occurs. Moreover, it is unknown whether the models can maintain high accuracy when applied to brain images acquired using different scanners and in different populations. Therefore, these studies need to be validated in external datasets to test the generalizability of the prediction models. In their review, Janssen and colleagues (2018) reported that only three studies (Drysdale et al., 2017; Koutsouleris et al., 2015; Koutsouleris et al., 2018) used independent test samples (or holdout test subjects from the training sample), including one multicenter study that applied leave-site-out CV (Koutsouleris et al., 2018).

In summary, studying interindividual variability in neuroscience will be essential for determining whether differences in brain organization are related to behavior, depend on disease, or are incidental (Gordon et al., 2017). Advanced analysis of fMRI data, along with ML implementation, offers a powerful framework for capturing complex and high-dimensional patterns in data, enabling researchers to identify individual "brain fingerprints", such as the frontoparietal network (Finn 2015), and predict behavioral traits or cognitive profiles from neuroimaging data (Rosenberg et al., 2016; Scheinost et al., 2019). This change of perspective is essential for identifying distinct brain configurations, which are reflected in different cognitive strategies that support behavior in real-life contexts.

1.4. Two case studies in cognitive neuroscience examining interindividual variability: memory and reward processing

To explore the potential of neuroimaging individualized approaches, this Ph.D. thesis examines interindividual variability in the brain-behavior relationship across two interconnected yet distinct cognitive functions that support the ability of individuals to learn and respond to external stimuli, namely memory and reward processing. These domains provide context for validating the shift from population-based averages to individual-specific characterizations, as they involve complex networks that differ significantly in both healthy and clinical populations. The following sections provide a descriptive overview of the specific mechanisms examined within these domains.

1.4.1. Memory systems in cognitive neuroscience

Memory encompasses multiple systems, including declarative (explicit) and non-declarative (implicit) memory (Tulving, 1972; Squire & Zola, 1996; Squire, 1992). Declarative memory refers to memories that can be consciously accessed, such as facts (semantic memory) and events (episodic memory), while non-declarative memory includes skills, priming, and conditioning, which often operate outside of conscious awareness and rely on different brain structures (Squire, 2004).

In human memory research, the discovery of multiple memory systems came from two major breakthroughs in the study of patients with pervasive, "global" amnesia. Perhaps the best-known case is that of patient H.M., who underwent bilateral resection of the hippocampus and adjacent medial temporal lobe (MTL) structures, resulting in profound anterograde amnesia (Scoville & Milner, 1957). This patient had the MTL area removed to alleviate his severe epileptic attacks. H.M. consequently suffered a nearly complete loss of the ability to form new permanent memories, and this impairment extended to all categories of learning materials. Yet, the early studies also showed that a few exceptions to the otherwise global deficit were apparent. H.M. was able to learn new motor skills (Corkin, 1968),

and he showed a facilitation of perceptual identification resulting from prior exposure to objects or words (an effect that later came to be understood as reflective of a preserved 'priming'; Warrington & Weiskrantz, 1968).

The second breakthrough came in 1980 when Cohen and Squire proposed that these 'exceptions' to amnesia were examples from a large domain of preserved learning capacities in amnesia. Their conclusion was based on the observation of complete preservation of the acquisition and retention of a perceptual skill (reading mirror-reversed words) in amnesic patients. These patients displayed fully intact skilled performance, yet they were significantly impaired in recognizing the specific words they were trained on and in recalling their training experiences. Cohen and Squire argued for the existence of functionally distinct memory systems: a procedural memory system dedicated to the tuning and modification of brain circuits that support skilled performance, and a declarative memory system that supports the encoding, storage, and retrieval on demand of memories for specific facts and events. These functionally distinct memory systems were tied to separate brain systems, with declarative memory seen as critically dependent on the MTL structures damaged in amnesia. Procedural memory was viewed as supported by several distinct brain systems specialized for particular types of skilled performance (Ferbinteanu, 2019; Eichenbaum & Cohen, 2008).

Over the decades, the original framework has been expanded through neuroimaging, electrophysiological studies, and computational research. Eichenbaum (2010) proposed that each memory system involves the cerebral cortex, particularly the perceptual and motor processing regions and the cortical association areas, which project to cortical regions (Figure I.2). The cerebral cortex thus provides major inputs to each of the three main pathways of processing related to distinct memory functions. First, there is a cortical-hippocampal circuit via the parahippocampal region, supporting declarative memory. The primary output of hippocampal and parahippocampal processing returns to the same cortical areas that provided input to the hippocampus, suggesting that the

hippocampus's role is to influence the persistence and organization of cortical representations. Second, procedural memory subsystems involve a cortical-striatal circuit that mediates habit formation and a parallel pathway that mediates different aspects of sensorimotor adaptations, which utilize sensory and motor pathways through the cerebellum. Lastly, the other major system involves the amygdala as a nodal stage in the association of exteroceptive sensory inputs to emotional outputs effected via the hypothalamic-pituitary axis and autonomic nervous system, as well as emotional influences over widespread brain areas (Eichenbaum, 2010, 2017). The putative involvement of this pathway in such processing functions has led many to consider this system as specialized for "emotional memory".

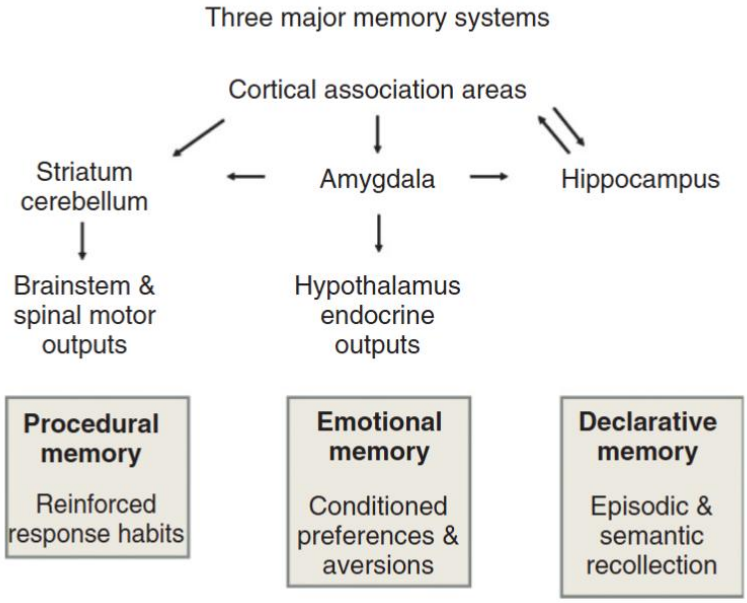

Figure I. 2 A schematic diagram of three prominent memory systems of the brain. The origins of major inputs to each of these systems involve widespread areas of the neocortex, and in particular, the association areas. Outputs of these cortical areas via three main routes. One route is through the parahippocampal region and into the hippocampus. The main outputs of hippocampal and parahippocampal processing are back to the same cortical areas that provide the main inputs. These pathways mediate declarative memory. Another route involves projections into two main subsystems via the striatum and cerebellum that mediate different aspects of procedural memory. These pathways involve both projections back to the cortex and outputs to brainstem motor nuclei. The third main route from the cortex is to the amygdala. The outputs of the amygdala project in several directions to hormonal and autonomic systems. This system mediates the expression of emotional memories. Amygdala outputs also return to the cortex and to the other memory systems to modulate the consolidation of other types of memory processing. Figure from Eichenbaum, 2010

1.4.2. The reward system in humans: anticipation and outcome phases

Individuals exhibit marked differences in their responses to cues that predict rewarding outcomes. Preclinical models in rodents have focused on observable behavior using Pavlovian conditioning paradigms to explore the reward system, identifying two distinct behavioral phenotypes: Sign-Trackers (ST), who assign salience to the reward-predictive cues (i.e., the lever) and approach and interact with it, and Goal-Trackers (GT), who focus on the reward itself (i.e., food; Boakes, 1977; Hearst & Jenkins, 1974). Several pharmacological and optogenetic studies demonstrated that the ST behavior is dependent on striatal dopamine (DA), as intra-accumbens injection of a DA antagonist substance significantly altered sign-tracking behavior, but not goal-tracking (Darvas et al., 2014; Saunders & Robinson, 2012; Saunders & Robinson, 2012; Flagel et al., 2011; Iglesias et al., 2023). However, in real-world contexts, humans primarily interact with cues and stimuli that already carry reward associations (Versace et al., 2016)—whether acquired through direct experience, well-established learning, or explicit instructions—rather than relying solely on newly formed associations.

Over the past 20 years, neuroimaging has significantly contributed to understanding the neural basis of reward processing in humans, primarily using the Monetary Incentive Delay (MID) task (Knutson et al., 2000), which is considered a robust measure of incentive motivation (Knutson & Greer, 2008; Knutson et al., 2000). This *instrumental-reward* task is inspired by literature on Pavlovian conditioning (Pavlov, 1927) and DA (Shultz, 1998), aligning with recent theories of reward processing (Berridge, 2007; 2009; Berridge & Kringelbach, 2015). In fact, the primary advantage of using this task is that it has distinct phases, namely anticipation and outcome, with separable activation patterns. For instance, the first stage during cue presentation, i.e., the *anticipation phas*e, may be modeled as a "wanting" phase (e.g., the lever presentation in the Pavlovian paradigm), eliciting motivation, and it should evoke robust activation in striatal regions through projections from the DA-rich ventral

tegmental area (VTA). Conversely, when modeling the *outcome phase* (e.g., the food receipt in Pavlovian conditioning), one would expect less activation of the ventral striatum (VStr) in response to reward, as only ~ 10% of neurons in the nucleus accumbens (NAc) facilitate pleasure.

Researchers have conducted meta-analyses of the MID task mostly using the 'Reward vs. Neutral' contrasts encompassing both phases, identifying similar patterns of neural activity in healthy adults (Jauhar et al., 2021; Knutson & Greer, 2008; Oldham et al., 2018; Wilson et al., 2018; Chen et al., 2022). These studies have shown activation of a frontostriatal-thalamic network, including bilateral anterior insula, striatum, thalamus, supplementary motor area (SMA), and precentral gyrus during the anticipation phase, while medial orbitofrontal cortex (OFC), ventromedial prefrontal cortex (vmPFC), dorsal anterior cingulate cortex (ACC), and occipital cortex were recruited during outcome (Chen et al 2022, Oldham et al 2018; Wilson et al., 2018).

Delineating the neural correlates of these component processes has facilitated empirical research on dysfunctional approach and avoidance in psychopathology, mapping functional differences in clinical populations (Whitton et al., 2015). For example, as compared to the control group, patients with major depressive disorder showed lower activation in the VStr during outcome (Carl et al., 2016; Whitton et al., 2015). Instead, patients with SCZ (compared to controls) exhibited hypoactivation in the VStr, ACC, median cingulate cortex, amygdala, precentral gyrus, and superior temporal gyrus during the presentation of a reward-indicating cue (Zeng et al., 2022; Juckel et al., 2006; Strauss et al., 2014).

These findings, based on group-average contrasts, have greatly influenced our understanding of reward processing and dysfunction across both normal and abnormal conditions. However, they assume a level of neural uniformity that may ignore significant variability between individuals in how rewards are processed.

1.5. Research proposal

In this thesis, I present the application of a multi-level pipeline to cognitive neuroscience problems, integrating neuroimaging, especially task-based and rs-fMRI, and machine learning algorithms to investigate the functional architecture of the brain.

My goal is to understand the brain functioning underlying the multiplicity of individual behaviors, moving beyond traditional group-level analyses, which often obscure meaningful differences between subjects through a "one-size-fits-all" approach assuming a homogenous functional anatomy of the brain (Laumann et al., 2015; Power et al., 2011; Seghier, M.L. and C.J. Price, 2018; Hermosillo, 2024), and instead treat interindividual variability—differences in cognitive performance, emotional responses, learning strategies, neural architecture, and FC—as a source of potentially valuable information into cognitive diversity, behavioral traits, and clinical heterogeneity (Dubois & Adolphs, 2016; Kanai & Rees, 2011).

I built my work on developing analytical models on HC to be applied to clinical populations presenting behavioral alterations in such cognitive domains, such as SCZ. Indeed, the expected impact is to validate and provide a novel methodology by using individualized approaches with fMRI in combination with unsupervised machine learning techniques, such as clustering. To do so, I conducted two case studies across different cognitive domains, specifically memory (Study 1) and reward processing (Study 2), to demonstrate the main hypothesis that different individuals utilize their brains differently. These studies provide the foundation for clinical predictions using fMRI and for developing more personalized treatment approaches.

Chapter 2 focuses on interindividual variability related to memory and learning in two independent samples of HC through a multi-scan fMRI experiment, encompassing both resting-state and memory task performance. Here, I hypothesized that brain resources may be differentially recruited within subgroups of individuals, underlying variability in memory performance. Specifically,

I tested whether interindividual differences in learning are contingent upon frontoparietal network configurations mediated by thalamic activity. First, I investigated the hypothesis that individual-level analytical approaches may gain greater predictive power of individual performances, by analyzing interindividual variability in thalamic recruitment before, during, and after memory encoding using a native individual-space approach, and by comparing it with group-level analyses using normalized individual FC maps. Second, I hypothesized differential involvement of thalamic subdivisions across memory phases. Thus, the use of an individualized approach may provide a better understanding of how thalamic task activity influences cortical configurations during the resting state, thereby supporting learning. To investigate the relationships between the individual associative memory performance and the individual-level cortico-thalamic recruitment across baseline resting-state FC, the activity during the encoding and the retrieval phases, serial path analyses within an SEM framework were implemented. Finally, to test the hypothesis that individuals could be grouped based on specific brain functional patterns that contribute to successful (or unsuccessful) information retention, I implemented an unsupervised ML algorithm, i.e., k-means clustering, that allows the stratification of individuals, by including the individual task-related activity during i) the encoding and ii) the retrieval phases during the associative memory task.

Furthermore, Chapter 3 examines inter-individual differences in reward processing across four independent cohorts, analyzing brain activity patterns during a reward task using fMRI. Building on the preclinical Sign-Tracker/Goal-Tracker (ST/GT) animal model (Flagel et al., 2007, 2009)—recently translated in humans (Schad et al., 2020; Colaizzi et al., 2023)—the core hypothesis is that different individuals may show distinct brain activations during two separate phases of a reward task, namely anticipation and outcome, offering a stratification model of reward processing. To test this hypothesis, I first applied a clustering algorithm in the three independent cohorts of healthy individuals, ensuring the generalizability of the model. Then, I applied the model to a fourth cohort that included patients

with SCZ and patients at their first episode of psychosis (FEP), assuming that, despite the expected overall attenuation in ventral striatal activation—at least partly due to antipsychotic treatment (Hawkins et al., 2021; Hawkins et al., 2018; Radua et al., 2015; Pessiglione et al., 2006)—meaningful variability between individuals in cue- versus outcome-driven neural responses similar to imaging profiles could still be detected. Furthermore, to determine whether sustained incentive value attribution to reward-anticipation cues differentiates individuals, I also leveraged the large numbers of trials per condition in the replication dataset to conduct a trial-by-trial analysis of ventral-striatal responses. In preclinical models, intra-accumbens injection of a DA antagonist substance significantly altered sign-tracking behavior, but not goal-tracking (Saunders & Robinson, 2012; Flagel et al., 2011; Iglesias et al., 2023), showing that only sign-tracking behavior depends on striatal DA. Once the imaging phenotypes have been reliably identified, linear-mixed effects models have been applied in a third independent cohort of healthy adults with a double-blind, placebo-controlled, crossover design to test the hypothesis that reward-related clusters of individuals may predict differential responses to dopaminergic agents. This cue-driven, DA-sensitive behavior provides a mechanistic parallel to the hypothesis of aberrant salience in psychosis, where dysregulated DA signaling leads to inappropriate attribution of significance to otherwise neutral cues. Thus, in the clinical cohort of SCZ and FEP, I hypothesized that these neurocognitive phenotypes have predictive utility in stratifying patients by examining their associations with current antipsychotic treatment type and psychopathological profiles.

Studying the neurobehavioral processes that differ among individuals may help in stratifying individuals based on their brain functioning, with the potential to bridge the gap between inferential statistics and precision psychiatry. An integrative approach exploiting the potential of machine learning on the *in vivo* measurement of the brain functioning may advance our understanding of neuropathological processes in clinical conditions, enabling more effective targeted treatments.

Chapter 2

Study 1: Interindividual variability in memory processes is related to cortico-thalamic networks during memory encoding and retrieval

The content of this chapter has been published as**: Interindividual Variability In Memory Performance Is Related To Cortico-Thalamic Networks During Memory Encoding And Retrieval**. *The Journal of Neuroscience*, 2025;45(19). DOI: 10.1523/JNEUROSCI.0975-24.2025.2.1

2.1. Introduction

Every person is unique in their learning process and related brain functional organization. Learning is the process by which an individual acquires new knowledge, which, by definition, depends on the ability to store and retrieve that knowledge through memory.

Memory, one of the most extensively studied cognitive processes in psychology and neuroscience, is a crucial cognitive function for environmental adaptation, enabling organisms to survive and thrive in favorable conditions while avoiding adverse ones. It is specifically defined as the process involved in the acquisition, storage, and retrieval of information (Squire, 2009; Zlotnik and Vansintjan, 2019). Memory formation transforms transient representations into durable engrams (Dudai, 2004; Squire et al., 2015) and typically involve three fundamental stages: (1) encoding, which refers to the initial registration and integration of incoming information; (2) storage, the process of maintaining information over time; and (3) retrieval, the ability to access stored information when needed (Shettleworth et al., 2009; Anderson, 2013).

Memory serves as the foundation for learning by enabling the retention of information within the brain (Ghosh & Gilboa, 2014). Learning, in turn, involves acquiring new information that changes behavior and memory. It is central to the formation of associations necessary for adaptive behavior (Anderson, 2000).

Over the past several decades, group-average studies have played a foundational role in advancing our understanding of the neural architecture of memory. Using fMRI and lesion-based approaches, these studies have consistently highlighted the MTL—particularly the hippocampus—as a central hub for episodic memory encoding and retrieval (Squire et al., 2004; Nadel & Moscovitch, 1997). Additionally, prefrontal cortex (PFC) activity has been linked to strategic retrieval and working memory processes (Cabeza et al., 2002), while posterior parietal cortex engagement has been associated with attentional control during retrieval (Wagner et al., 2005). Neuroimaging studies contrasting patients with amnesia or Alzheimer's disease and HC have been especially influential in mapping deficits in memory-related networks, demonstrating widespread reductions in MTL activation and cortical connectivity in pathological populations (Schacter et al., 1996; Buckner et al., 2005). These group-level studies provided reproducible, population-wide inferences about the functional roles of specific brain regions in memory, helping to establish a canonical memory network.

In recent years, individual-level approaches have gained increased traction, demonstrating that memory processes are shaped by unique patterns of brain activity and connectivity that may be obscured by group-averaged analyses (Mueller et al., 2013; Finn et al., 2015). These studies emphasize that functional brain organization varies markedly across individuals and that such variability correlates with differences in memory performance, strategies, and brain-behavior coupling. These approaches leverage within-subject analyses, individualized functional mapping, and native-space connectivity to more effectively capture the unique neural configurations that support episodic memory (Braga et al., 2019; Gordon et al., 2020). Unlike traditional models that emphasize shared neural responses, this line

of research has shown that FC networks—especially those involving the MTL, PFC, and thalamus—exhibit stable, trait-like variability across individuals and are predictive of memory ability (Gordon et al., 2021; Gordon et al., 2017; Seitzman et al., 2019; Greene et al., 2020; Van Buuren et al., 2019; Gratton et al., 2020). For example, recent findings indicate that early-stage FC changes occur at the system level following a learning experience and are associated with individual variation in memory performance (Passiatore et al., 2021). Recent developments in precision fMRI techniques have revealed that these differences are not simply noise, but rather reflect distinct functional architectures—or "connectome fingerprints"—that are behaviorally meaningful and reproducible across sessions (Finn, 2015; Laumann et al., 2021; Lynch et al., 2021). Furthermore, studies employing individualized functional parcellation have found that person-specific connectivity patterns provide better explanatory power for memory performance than population-based atlases (Dworetsky et al., 2021; Kong et al., 2019).

These insights are particularly relevant for memory research, where individual differences in strategy, motivation, and attention can shape both the neural substrates of memory encoding and the consolidation processes. These findings challenge the assumption of homogenous brain organization and underscore the value of precision neuroscience approaches that respect interindividual differences (Dubois & Adolphs, 2016; Seghier & Price, 2018).

Collectively, these findings underscore the need to move beyond group-level inference to understand how individual brains support memory processes, providing a more nuanced and personalized understanding of cognitive function.

2.1.1. The role of MTL in memory

The MTL is not a single, uniform structure; rather, it is a complex network comprising the hippocampus, entorhinal cortex, perirhinal cortex, and parahippocampal cortex. Each of these

components contributes differently to various aspects of memory. The hippocampus is thought to support the binding of items and contextual information, making it crucial for relational and spatial memory (Eichenbaum et al., 2007; Davachi, 2006). The perirhinal cortex has been linked to familiarity-based recognition, while the parahippocampal cortex is more involved in processing contextual or spatial information (Ranganath & Ritchey, 2012).

Memory consolidation and storage over time are key focuses in the study of memory neuroscience. The traditional standard model of systems consolidation suggests that memories are initially stored in the MTL, especially the hippocampus, and slowly become stabilized in distributed neocortical regions over time (Frankland & Bontempi, 2005; Squire & Alvarez, 1995). This model suggests that the hippocampus plays a time-limited role in memory retrieval: as consolidation proceeds, the neocortex can support retrieval independently of the MTL. However, this account has been challenged by alternative models, such as the multiple trace theory (MTT) and transformation theory, which argue that the hippocampus continues to be involved in retrieving richly detailed, episodic memories even after long periods (Nadel & Moscovitch, 1997; Winocur & Moscovitch, 2011). According to the MTT, each reactivation of a memory creates a new trace in the hippocampus, strengthening the memory but not rendering it independent of hippocampal involvement. Transformation theory, in contrast, proposes that over time, episodic memories may be transformed into schematic or semantic representations that rely more on cortical networks. A large number of human neuroimaging studies have found that the hippocampus is involved in binding separate episodic elements into cohesive units (Henke et al., 1997; Eichenbaum, 2004; Tubridy & Davachi, 2011).

Recent neuroimaging studies have supported the idea that the neocortex is actively involved in memory encoding and retrieval even at early stages. For instance, work by Brodt et al. (2018), which utilized fMRI and behavioral paradigms, demonstrated that learning can induce rapid systems-level

consolidation in cortical regions, such as the posterior parietal cortex, thereby challenging the traditional view of slow consolidation dependent on hippocampal reactivation.

Similarly, evidence from simultaneous hippocampal and neocortical recordings in rodents has disclosed bidirectional communication during memory encoding and retrieval (Yadav et al., 2022), suggesting a more dynamic and distributed model of memory networks. Functionally, memory processing depends not only on local computations within the MTL, but also on interactions with broader cortical and subcortical networks. For example, the PFC has been implicated in strategic aspects of memory, such as encoding strategies, retrieval monitoring, and suppression of irrelevant information (Simons & Spiers, 2003; Blumenfeld & Ranganath, 2007). Overall, memory processing in neuroscience is increasingly understood as a dynamic and distributed phenomenon, involving reciprocal interactions between the MTL, cortical, and subcortical structures, such as the amygdala, basal ganglia, thalamus, and cerebellum (Tan et al., 2025).

2.1.2. Thalamic involvement in memory

The thalamus is increasingly recognized for actively regulating information transmission to cortical regions (Sherman, 2016; Nakajima and Halassa, 2017; Pergola et al., 2018; Nakajima et al., 2019), affecting behavioral performance (Bradfield et al., 2013a; 2013b; Jankowski et al., 2013; Mitchell and Chakraborty, 2013), far beyond its classical role as a passive relay station for sensory information.

2.1.2.1. Functional architecture of the thalamus

The thalamus is comprised of numerous nuclei, some of which have multiple cytoarchitectonic subdivisions. Connections and projections of these subdivisions are primarily established within the boundaries of particular cortical fields (Jones, 2007). Thalamic nuclei are not interconnected, with the exception of the reticular thalamic nucleus (RTN), which is reciprocally linked with all other nuclei. Thus, the thalamo-cortical pathways remain largely segregated. Consistently, neurologic patients with

thalamic lesions exhibit regional-specific loss of cognitive functions (reviewed by Carlesimo et al., 2011; Carrera and Bogousslavsky, 2006; Schmahmann, 2003; Van der Werf et al., 2003b). Furthermore, thalamo-cortical pathways are usually confined to a single hemisphere, except for those encompassing the midline nuclei and the RTN (Barbas et al., 2013; Ray and Price, 1993). In particular, prefrontal and temporal regions are important areas for the crosstalk between the thalamus and the cortex, which is critical for high cognitive functions. The anatomical basis of the thalamic-PFC network includes the mediodorsal nucleus (MD), which is the main nucleus projecting to the PFC (Barbas et al., 2013; Klein et al., 2010; Taber et al., 2004), along with the anterior nucleus (AN), the lateral dorsal nucleus (LD). These nuclei have intimate interconnections with cortical and subcortical areas (Figure II.1).

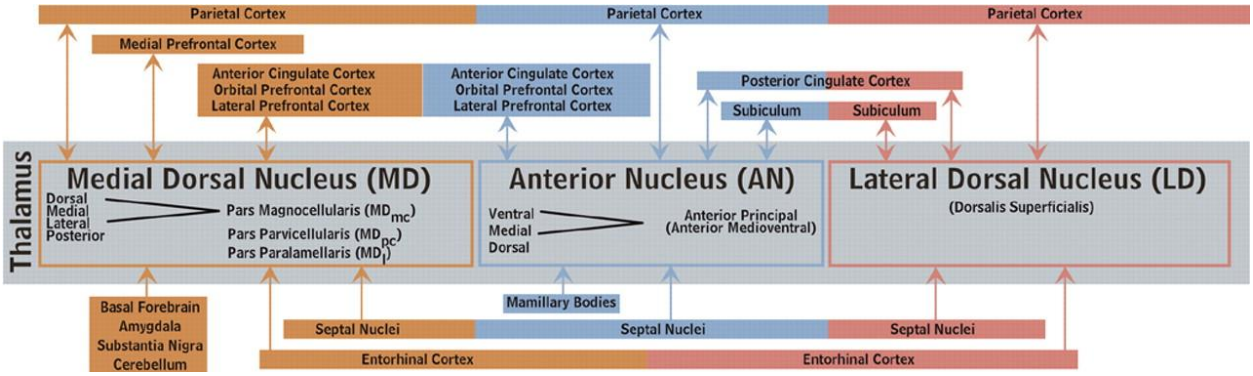

Figure II. 1 Functional architecture of the medial dorsal, anterior and lateral dorsal nuclei of the thalamus. The principal cortical and subcortical connections of the medial dorsal (orange), anterior (blue) and lateral dorsal (pink) nuclei of the thalamus. Reciprocal connections are indicated by color-coded circles. Afferents are indicated by color-coded triangles. The approximate location of fibers of passage are indicated by open symbols. Each thalamic nucleus is filled with the appropriate color on the left side of brain slices. Subdivisions and usual abbreviations for each nucleus are summarized in the circuit diagram. Some anatomic structures are labeled for orientation. Figure from Taber et al., 2004.

First, the MD nucleus has major reciprocal connections with the orbital, medial prefrontal, lateral prefrontal, and ACC regions (Yeterian & Pandya, 1988; Vogt, Pandya, & Rosene, 1987; Powell, 1973). Reciprocal connections with SMA and parietal cortices, as well as the frontal eye fields, have also been reported (Yeterian & Pandya, 1988; Vogt, Pandya & Rosene, 1987; Goldman-Rakic & Porrino, 1985). The amygdala, substantia nigra, and cerebellum also project to MD (Ilinsky et al., 1985; Russchen, Amaral & Price, 1987; Ray & Price, 1993; Aggleton & Mishkin, 1984). These connections

and the effects of injury to this region indicate a role in memory, mood, motivation, and the sleep/waking cycle. The MD comprises sub-components with different connectivity patterns (Barbas et al., 2013; Mitchell and Chakraborty, 2013; Pergola et al., 2012). The medial magnocellular portion receives afferents from the vmPFC and the MTL, and projects to the vmPFC and OFC (Aggleton, 2012). The lateral parvocellular portion, instead, is reciprocally connected to the dorsolateral PFC, which is also its major source of input (Byne et al., 2009; Mitchell and Chakraborty, 2013). The intralaminar nuclei project diffusely to the PFC, primarily to the striatum (Barbas et al., 1991; Preuss & Goldman-Rakic, 1987; Sadikot, Parent, & Francois, 1992). Similar to the magnocellular MD, the medial pulvinar (Pul) is connected to the PFC, temporal, and parietal cortices. The MD and the Pul, therefore, are nodes of a network of associative regions – "the seventh layer of the cortex" (LaBerge, 1997; Sherman and Guillery, 2011).

Second, the AN nucleus has major reciprocal connections with anterior and posterior cingulate cortices and the subiculum-presubiculum portion of the hippocampal complex (Yakovlev et al, 1960; Vogt, Pandya, and Rosene, 1987; Powell, 1973; Goldman-Rakic and Porrino, 1985; Bachevalier et al., 1997). Lesser connections with the orbital frontal, dorsolateral prefrontal, and parietal cortices have also been reported (Yeterian and Pandya, 1988; Goldman-Rakic and Porrino, 1985; Selemon & Goldman-Rakic, 1988; Cavada et al., 2000). AN also receives major input from the mamillary complex (Veazey et al., 1982; Bentivoglio et al., 1993). It may also receive a projection from the septal nuclei (Powell, 1973; Hreib et al., 1988). These connections and the effects of injury to this region indicate a role in memory, modulation of the sleep/waking cycle, and directed attention.

Third, the LD nucleus has major reciprocal connections with posterior cingulate and parietal cortices as well as the subiculum-presubiculum and entorhinal cortex portions of the hippocampal complex (Bentivoglio et al., 1993; Yeterian and Pandya, 1988; Vogt, Pandya, and Rosene, 1987; Selemon et al., 1988; Vogt, Rosene, and Pandya, 1979; Shibata and Yukie, 2003). It may also receive a

projection from the septal nuclei (Hreib, Rosene, and Moss, 1988; Powell, 1973). A role in integrating motivation and/or attention with sensory processes has been suggested (Yeterian & Pandya, 1988).

Beyond enhanced local cortical connectivity, thalamic circuits may also enhance connectivity across cortical areas (Zhou, Schafer, and Desimone, 2016; Saalman et al., 2012). The thalamus is globally connected with distributed cortical regions, and most thalamic subdivisions display network properties that can integrate multimodal information across diverse cortical functional networks (Hwang et al., 2017). Hence, at least three thalamic systems can be dissociated on the basis of their connectivity patterns: the medial thalamus connected to the MTL and the vmPFC (Mitchell & Dalrymple-Alford, 2005); the parvocellular MD, reciprocally connected to the dorsolateral PFC; and a lateral thalamic system (Lopez, Wolff, & Lecourtier, 2009; Mitchell & Dalrymple-Alford, 2005) connected to the striatum and diffusely to the PFC (Figure II.2).

Far from considering the thalamus as a passive relay station of information from sensory organs or subcortical structures to the cortex, it has become a critical hub region that could integrate diverse information being processed throughout the cerebral cortex, as well as maintain the modular structure of cortical functional networks (Hwang et al., 2017; Nakajima and Halassa, 2017).

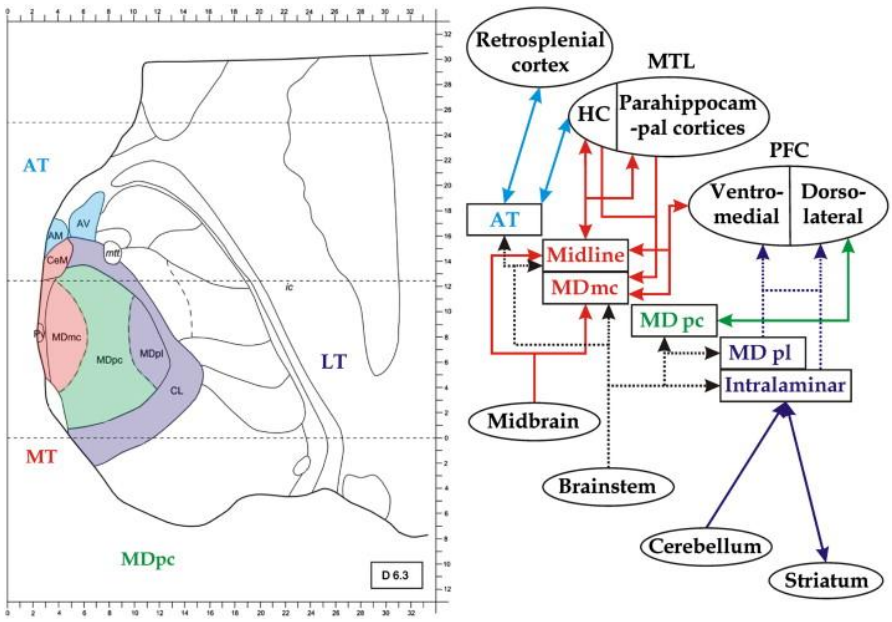

Figure II. 2 Thalamo-cortical connectivity. Left panel: spatial arrangements of the AT, MT, MDpc, and LT. The section is 6.3 mm superior to the intercommissural plane. Modified from Morel (2007). **Right panel:** Patterns of connectivity of the groups of nuclei described. Dashed arrows indicate diffuse projections. *Abbreviations*: AM, anteromedial nucleus; AV, anteroventral nucleus; CeM, centromedial nucleus; CL, centrolateral nucleus; LT, lateral thalamus; MDmc, magnocellular mediodorsal nucleus; MDpc, parvocellular mediodorsal nucleus; MDpl, paralamellar mediodorsal nucleus; MT, medial thalamus; MTL, medial temporal lobe; *mtt*, mammillothalamic tract; PFC, prefrontal cortex; Pv, paraventricular nucleus. Figure from Pergola et al., 2012.

2.1.2.2. Thalamic connections supporting memory

The functional architecture of thalamic involvement in cognition mirrors the patterns of anatomical and functional connectivity of its subdivisions with cortical and subcortical brain regions (Antonucci et al., 2021). The diverse cytoarchitecture of thalamic nuclei implies that different thalamic nuclei may specialize in distinct cognitive domains, while others serve multiple cognitive processes. Although previous investigations hinted at thalamic subdivisions' specificity for cognitive processes (Child and Benarroch, 2013; Funahashi, 2013; Jankowski et al., 2013; Lakatos et al., 2016; Marchand et al., 2014; Mitchell, 2015; Parnaudeau et al., 2018; Pergola et al., 2018; Saalmann, 2014), they have all focused on investigating one thalamic subdivision at a time, or one cognitive process at a time. Hwang et al. (2017) demonstrated how thalamic subdivisions are differentially involved in several brain functional

networks related to cognition, although without explicitly mapping the brain correlates of segregated domains.

Yet, it has been a challenge to localize function to thalamic structures because of two issues. First, highly localized thalamic lesions leading to specific clinical pathologies are uncommon. Second, imaging technologies have only recently acquired sufficient resolution to differentiate these structures from one another (Ji et al., 2019). For this reason, several neuroimaging studies have aggregated thalamic nuclei into thalamic subdivisions based on the adjacency between nuclei and the similarity of their connectivity patterns (Barbas et al., 2013; Pergola et al., 2012; Roy et al., 2022) to test hypotheses about the cognitive domain specificity of thalamocortical networks (Herrero et al., 2002; Antonucci et al., 2019, 2016; Wolff and Halassa, 2024). For example, one parceling considered the AT, MD, Pul, intralaminar nuclei, geniculate nuclei, ventral anterior region, and ventral lateral nucleus, assuming a relative functional cognitive independence between these subdivisions (Pergola et al., 2017; Di Carlo et al., 2020). In a recent meta-analysis, Antonucci et al. (2021) proposed a model that describes how thalamic subdivisions may differentially subserve neuropsychological domains. They showed that their parceling results only partially support the hypothesis of a thalamic subdivision cognitive specificity (Figures II.3 and II.4).

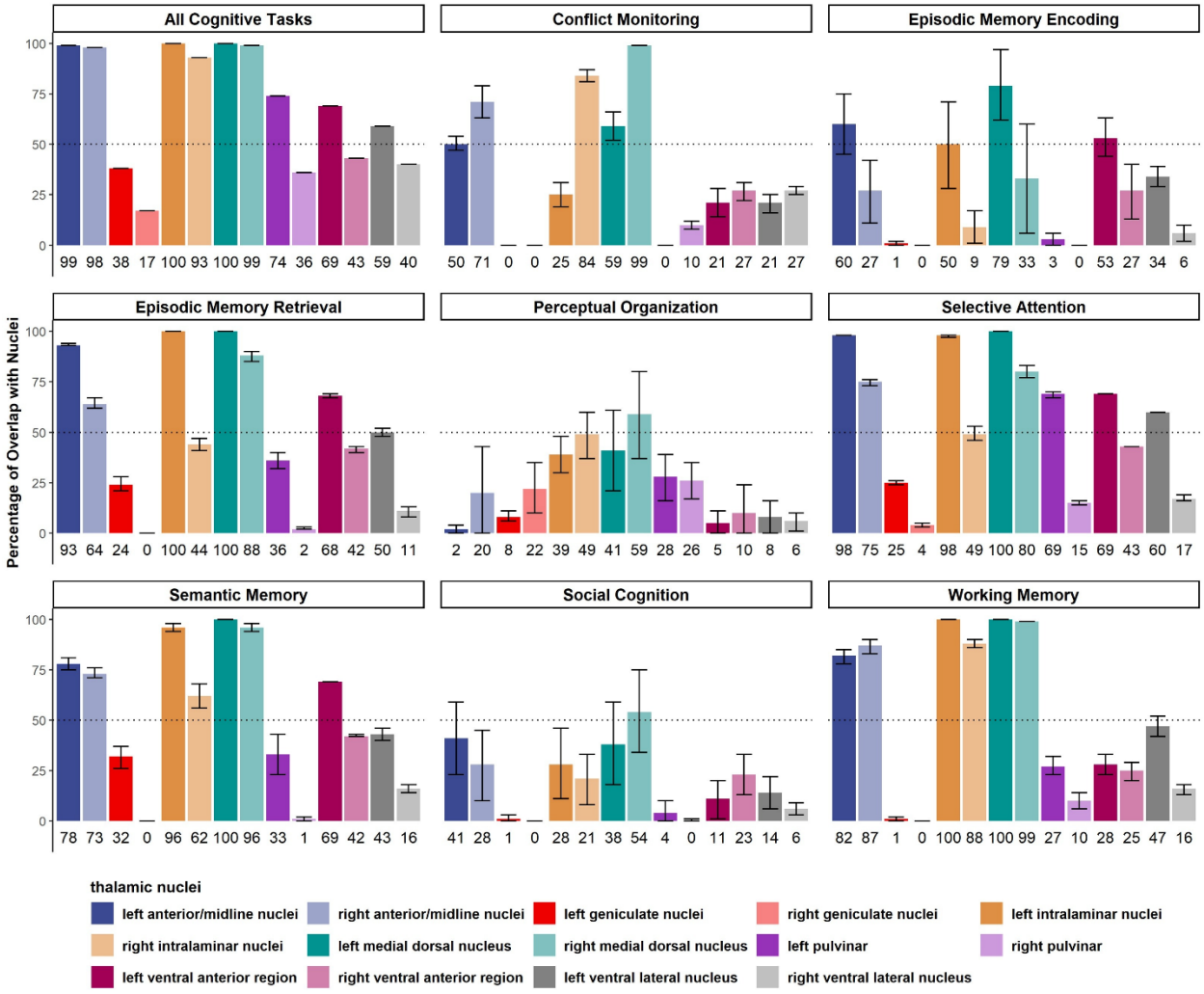

Figure II. 3 Contribution of each thalamic subdivision to each co-activation map. The percentage of overlap between each thalamic cluster and the seven specific thalamic subdivisions are weighted by the total size of subdivisions in the Thalamus Atlas (Krauth et al., 2010). Thalamic parceling is depicted only for seed-based maps based on at least four studies. Figure from Antonucci et al., 2021.

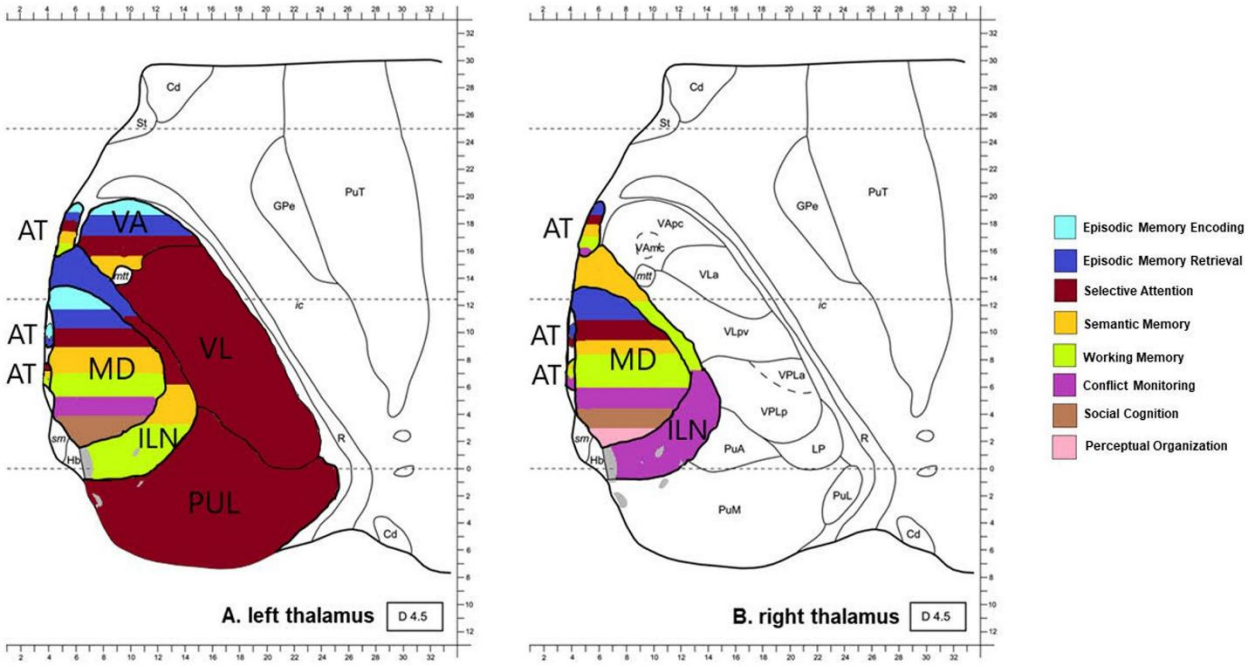

Figure II. 4 Spatial distribution of cognitive domain representation in thalamic subdivisions. Specific and flexible left (panel A) and right (panel B) thalamic contribution of nuclei with mean above 50 % of nuclei size. Abbreviations: AT = Anterior/Midline Nuclei; ILN = Intralaminar Nuclei; MD = Medio-Dorsal Nuclei; PUL = Pulvinar; VA = Ventral Anterior Nuclei. Figure from Antonucci et al., 2021.

Given that an individual cortical region receives inputs from multiple thalamic structures, Nakajima & Halassa (2017) suggest a model in which different thalamic inputs shift the connectivity weights within a cortical region, allowing it to shift between processing and sustaining locally generated categorical representation and participating in distributed categorical processing across multiple cortical nodes (Figure II.5). Testing this framework has undoubtedly expanded the understanding not only of thalamic function, but of cognition more broadly.

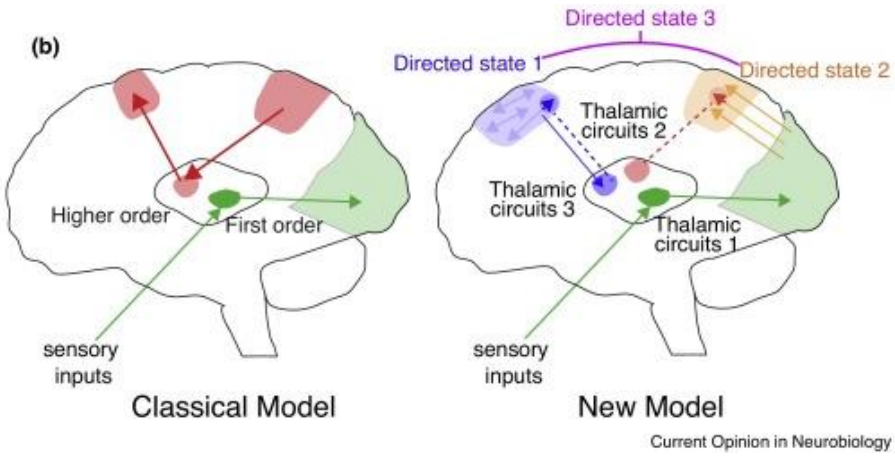

Figure II. 5 Thalamic control of functional cortical connectivity. The classical model reflects the idea of information relay from the LGN, which is shared by other sensory systems (green), and extended to account for thalamic nuclei that receive driving inputs from the cortex rather than subcortical or sensory inputs. The new model suggests that many of these circuits operate to enhance connectivity within a cortical region (*e.g.*, blue), or across different cortical areas (*e.g.*, orange). In these scenarios, thalamic inputs do not necessarily drive spiking (do not dictate cortical spike times), but rather enhance how spike times are generated in response to specific inputs (either local or from area A, B, C, *etc.*). A particular connectivity pattern reflects a directed arousal state, as defined in this review. Such states can encompass multiple cortical nodes and utilize multiple thalamic amplifiers (*e.g.*, directed state 3), depending on the complexity of the associated cognitive process. Figure from Nakajima & Halassa, 2017.

Growing interest in this fascinating structure is bolstered by consistent evidence from both animal and human studies, demonstrating that the connectivity pattern of the medial thalamus with the MTL and the PFC suggests a pivotal role of this brain region in memory. Particularly, specific thalamic nuclei are crucial for episodic memory function (Aggleton et al., 2011; Mitchell & Chakraborty, 2013), such as the AN (Aggleton & Brown, 1999; Mitchell & Dalrymple-Alford, 2005; see Aggleton et al., 2010 for a review), MD (Aggleton & Brown, 1999), and the intralaminar and the midline nuclei (Vertes, Hoover, Do Valle, Sherman, & Rodriguez, 2006; Vertes, 2006).

The AT has been associated with episodic memory processes (Jankowski et al., 2013; Stillova et al., 2015; Mitchell and Dalrymple-Alford, 2005), which in turn appears also involved in executive processing (Pergola et al., 2018; Van der Werf et al., 2003), decision making (Mitchell, 2015), and attention control (Antonucci et al., 2016). MRI studies have highlighted the role of the AN and MD nuclei of the thalamus in episodic memory processing, playing distinct roles during encoding and

retrieval (Pergola et al., 2012; Pergola et al., 2013a; Pergola and Suchan, 2013; Danet et al., 2015). Yet, the contributions of these nuclei to episodic memory appear quite different (Mitchell and Chakraborty, 2013). The AN, connected to the hippocampus and the retrosplenial cortex, is the primary thalamic hub for recognition mediated by recall and is involved in spatial memory (Aggleton et al., 2011). It is debated whether the role of the anterior nuclei prevails during retrieval or encoding (Aggleton and Brown, 1999; Pergola and Suchan, 2013; Geier, 2023). Damage to the AN has been associated with profound amnesia in both clinical and experimental settings, further underscoring its role in memory consolidation and recall (Danet et al., 2015; Segobin & Pitel, 2021). In contrast, the MD maintains strong bidirectional connections with the PFC and perirhinal areas and is thought to contribute primarily to the encoding and strategic organization of episodic memories (Blumenfeld et al., 2011; Pergola & Suchan, 2013). Functional neuroimaging studies have demonstrated that the activity of the MD nucleus is particularly relevant for tasks that require memory-guided decision-making and contextual integration (Pergola et al., 2012; Wagner et al., 2019). Since the AN and MD nuclei do not communicate directly with each other but are independently connected to cortical and subcortical regions (Acsády, 2022), they likely belong to distinct functional circuits (Wolff et al., 2015; Halassa and Sherman, 2019).

Altogether, these findings point to a need to reconceptualize the thalamus as playing a central role in the neural architecture of memory. Its function appears to be pivotal in supporting both the encoding of new information and the retrieval of stored content by modulating cortico-cortical communication in a state- and task-dependent manner. Given that FC patterns involving the thalamus vary significantly between individuals (Mueller et al., 2013; Gordon & Nelson, 2021), a deeper investigation of cortico-thalamic circuits may provide critical insights into the neural basis of interindividual differences in memory performance—a key objective of the current research.

2.1.3. A multi-session approach to study corticothalamic functional connectivity

The relevance of the thalamus in memory processing is further evidenced by studies using resting-state and task-based fMRI paradigms, which have demonstrated that thalamic FC with cortical regions—such as the PFC, parietal lobe, and medial temporal structures—predicts memory performance across individuals (Halassa & Sherman, 2019; Antonucci et al., 2021). Crucially, this connectivity is not static but dynamically reconfigures in response to task demands and cognitive state, suggesting a flexible role for the thalamus in coordinating distributed memory networks (Mitchell, 2015; Passiatore et al., 2021).

Although there are no direct structural connections between the MD thalamus and hippocampus (Vertes, 2006), it is thought to contribute to memory via an extrahippocampal circuit (Ketz et al., 2015). The thalamus, along with other brain regions, i.e., parahippocampal gyrus and the PFC, communicates with the hippocampus for the encoding of new memories (Pergola & Suchan, 2013), particularly, its connections underlie post-encoding memory processing. Thus, investigating FC between episodic memory regions shortly after encoding offers an opportunity to study the early stage of memory traces in the human brain.

As the nucleus reuniens appears to be involved in durable memory formation during encoding and consolidation in rodents, Wagner et al. (2019) hypothesized a similar role for the corresponding medioventral thalamus in humans. Overall, along with the medioventral thalamus, the AN, LD, and MD thalamus appear to be relevant for memory processing, particularly in encoding. Using a multi-session fMRI experiment, they investigated how these four thalamic regions contribute to different aspects of durable memory formation in humans. Human subjects underwent fMRI while intentionally encoding picture-location associations (Wagner et al., 2016).

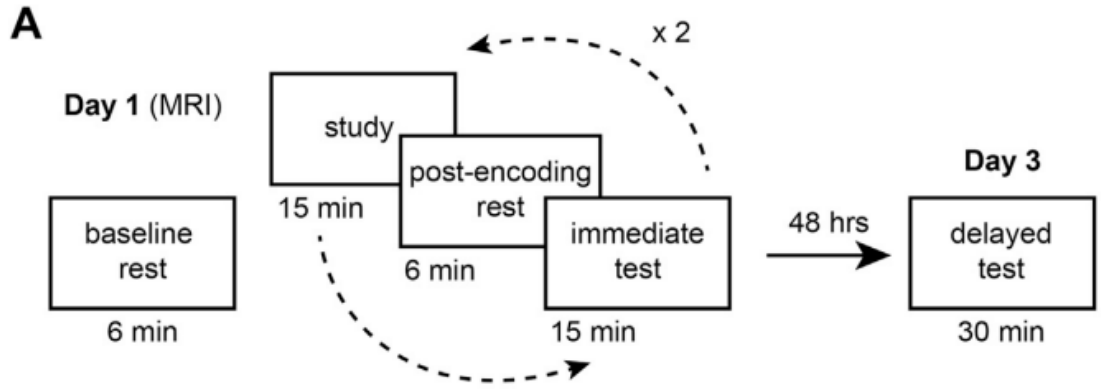

Figure II. 6 Study timeline. Subjects studied 192 unique picture-location associations and were tested for their immediate memory (immediate test, day 1) during two study-rest-test cycles (schematically indicated with dashed arrows), and during a delayed memory test 48 h later (delayed test, day 3). Subjects were scanned during an initial baseline resting-state period, as well as during resting-state periods after the study phase of each of the two cycles (post-encoding rest). Figure from Wagner et al., 2019.

Neural activity related to consolidation was assessed during resting-state periods following the encoding task, and compared to the baseline resting-state activity acquired before the task. To capture durable memories, study material was tested on the same day and 48 hours later. While "weak" memories could only be remembered at the immediate test, "durable" memories also persisted after 48 h. A depiction of the study timeline is shown in Figure II.6, adapted from Wagner et al. (2019).

Due to the small size of the medioventral thalamus, they combined it with the adjacent central medial nucleus to form an enlarged region and referred to it as the medioventral thalamus. Thalamic regions-of-interest (ROIs) are represented in Figure II.7. Wagner and colleagues (2019) found an increased coupling of the medioventral-, adjacent anterior-, lateral dorsal-, and mediodorsal thalamus with the hippocampus and surrounding MTL, as well as with anterior and posterior midline regions related to durable memory encoding. The medioventral and LD thalamus showed increased connectivity with posterior medial and parietal cortex from baseline to post-encoding rest, positively scaling with the proportion of durable memories formed across subjects. Additionally, the LD thalamus proved consolidation-related coupling with the inferior temporal, retrosplenial, and medial PFC. This multi-session fMRI paradigm suggests that thalamo-cortical crosstalk strengthens

mnemonic representations at initial encoding, and that cortical coupling of specific thalamic subregions supports consolidation thereafter.

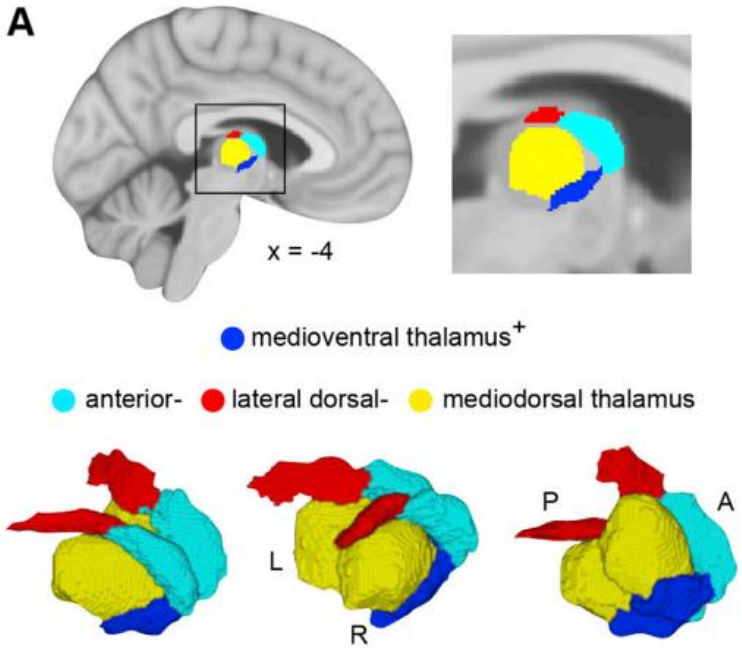

Figure II. 7 Thalamic regions-of-interest. The medioventral thalamus (blue), anterior thalamus (cyan), lateral dorsal thalamus (red), and mediodorsal thalamus (yellow). Slices for all figures are based on a mean structural scan. The magnified image (right part) shows the inset as marked on the sagittal brain slice. The lower part shows 3D-renderings of the four thalamic regions. Figure from Wagner et al., 2019.

Using a similar multi-session fMRI design that included a memory task interposed between two resting-state scans, a recent study of our group investigated how learning influences short-term changes in brain network configuration and how these changes relate to individual differences in episodic memory (Passiatore et al., 2021). They found that while the configuration within executive, default mode, and cerebellar circuits remained stable across sessions, the visual circuit showed post-learning reconfiguration. Notably, this change was modulated by memory performance: individuals with higher recall scores exhibited an FC pattern of MTL and PFC with the visual circuit, as well as thalamic FC within the executive control system, suggesting a performance-dependent adaptation of visual and memory-related circuits. Importantly, thalamic connectivity within the executive control network varied as a function of individual performance, showing a negative association with memory

ability during baseline and a positive one post-encoding. This suggests a dynamic, phase-specific involvement of the thalamus in supporting cortical integration and cognitive updating during memory processing—consistent with prior models of thalamic regulation of prefrontal activity (Antonucci et al., 2021; Halassa & Sherman, 2019; Pergola et al., 2018). These findings align with those of Wagner et al. (2019), who observed performance-related coupling between thalamic subdivisions, PFC, MTL, and parietal regions, supporting the view that thalamic-cortical crosstalk contributes to long-term memory stabilization. Despite these promising insights, the study was conducted on a relatively small sample and lacked external replication.

2.1.4. Hypothesis and aim of the study

Building on this knowledge, the current study aimed to expand the previous research questions by Passiatore et al. (2021) by investigating interindividual variability in corticothalamic contributions to memory. We complemented the previously published cohort by using a larger, independent sample and a native-space analysis approach that permits fine-grained functional segmentation of thalamic subdivisions.

As outlined in the General Introduction, the main hypothesis is that different individuals may recruit distinct neural resources that support their learning abilities. Specifically, we used a multi-scan fMRI experiment to test the hypothesis that interindividual differences in learning are contingent upon frontoparietal network configurations mediated by thalamic activity. Particularly, we aimed to analyze interindividual variability in thalamic recruitment before, during, and after memory encoding using an individualized spatial approach.

Firstly, we hypothesized that individual-level analytical approaches could yield greater predictive power for individual memory performances compared to group-level analysis. By comparing the two approaches, we aimed to further explore interindividual differences in the

recruitment of thalamic subdivisions related to memory, avoiding information loss caused by averaged group-level analysis. Secondly, we hypothesized that individuals employ distinct neural configurations before completing a memory task to support their individual performance in an associative memory task. We also proposed that thalamic activity during the task may influence cortical configurations during resting-state, thereby aiding learning. Specifically, we aimed to investigate the relationships between individual memory performance and cortico-thalamic recruitment at the individual level across the rs-fMRI scan, as well as activity during the encoding and retrieval phases of the memory fMRI task.

Finally, we hypothesized that diverse cortico-thalamic circuits recruited by different individuals may underlie variability in memory performance. We aimed to stratify individuals based on their activity during memory encoding and retrieval, highlighting different thalamocortical profiles that support memory performance.

This study emphasizes the use of an individualized approach, which prioritizes inter-individual variability rather than seeking commonalities across individuals by averaging participant data. Understanding the distinct roles of thalamic nuclei in episodic memory holds significance also for clinical populations, such as SCZ, in which memory processes and thalamic connectivity are impaired.

2.2. Materials and Methods

2.2.1. Participants

A total of 184 healthy adults were recruited for an fMRI experiment at the Ruhr University Bochum (RUB) and the University of Bari Aldo Moro (UNIBA).

RUB. We enrolled 36 healthy adults in the high-resolution multi-scan fMRI study, which was approved by the Faculty of Medicine Ethics Committee of Ruhr-Universität Bochum. In accordance with the Declaration of Helsinki, all participants provided written informed consent prior to their participation. All participants had normal or corrected-to-normal vision. Vision correction during MRI was achieved via contact lenses or MRI-compatible glasses. No individual had a history of drug or alcohol abuse within the last six months, head trauma with loss of consciousness, or any clinically relevant medical condition. For each participant, handedness was assessed through the Edinburgh Handedness Inventory (Oldfield, 1971). One ambidextrous participant was excluded to reduce heterogeneity for laterality (Ferrucci et al., 2013). Five participants did not complete the protocol due to technical failures during data collection, and one more was excluded because of data artifacts and excessive head motion during acquisition. Excessive motion was assessed using framewise displacement (FD), i.e., the displacement of one frame relative to the previous, with subjects excluded if mean FD>0.5 within a given scan (Power et al., 2012; Power et al., 2014). A total of 29 participants were included in the final sample for the resting-state analysis (mean age ± SD = 26 ± 3 years; range 19-33 years; gender ratio, m:f = 10:19; mean handedness ± SD = 0.89 ± 0.11). We further excluded two participants who did not complete the task fMRI, including 27 participants in the sample for the task activity analysis (mean age ± SD = 25 ± 3 years; range 19-32 years; gender ratio, m:f = 8:19; mean handedness ± SD = 0.9 ± 0.12).

UNIBA. We recruited an independent sample of 148 healthy adults for the multi-scan fMRI study, which was approved by the institutional Ethics Committee of the University of Bari Aldo Moro. All participants gave written informed consent before participating and underwent the same screening for inclusion as RUB participants. Five participants were excluded due to data artifacts, four were excluded for technical failure during data collection, twenty-two participants were excluded for excessive motion during the rs-fMRI acquisition, and nine more were excluded because of low performance on the associative memory task (see Behavioral Analysis). A total of 108 participants were included in the final sample for the resting-state analysis (mean age ±SD = 28±8 years; range 18-61 years; gender ratio, m:f = 56:52; mean handedness ± SD = 0.63 ± 0.41; mean intellective quotient ± SD = 106.6 ± 9.73 measured through the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981); mean education ± SD = 17.31 ± 2.3 measured through the Hollingshead Four Factor Index of Socioeconomic Status (Hollingshead and Frederick, 1964)). We further excluded four participants due to data artifacts and three more because of excessive motion in at least one of the task fMRI scans (mean FD > 0.5 mm). The final UNIBA sample for the task activity analysis included 101 participants (mean age ±SD = 27±8 years; range 18-59 years; gender ratio, m:f = 49:52; mean handedness ± SD = 0.64 ± 0.4; mean intellective quotient ± SD = 106.5 ± 9.71; mean education ± SD = 17.24 ± 2.25).

2.2.2. Experimental Design

The multi-scan fMRI procedure in the two samples consisted of (i) a baseline resting-state scan and (ii) the performance of an associative memory task (Figure II.8A-B), encompassing encoding and retrieval phases. In RUB, (iii) a second resting-state scan, i.e., post-encoding, was also collected (Tambini et al., 2010; Wagner et al., 2019; Passiatore et al., 2021).

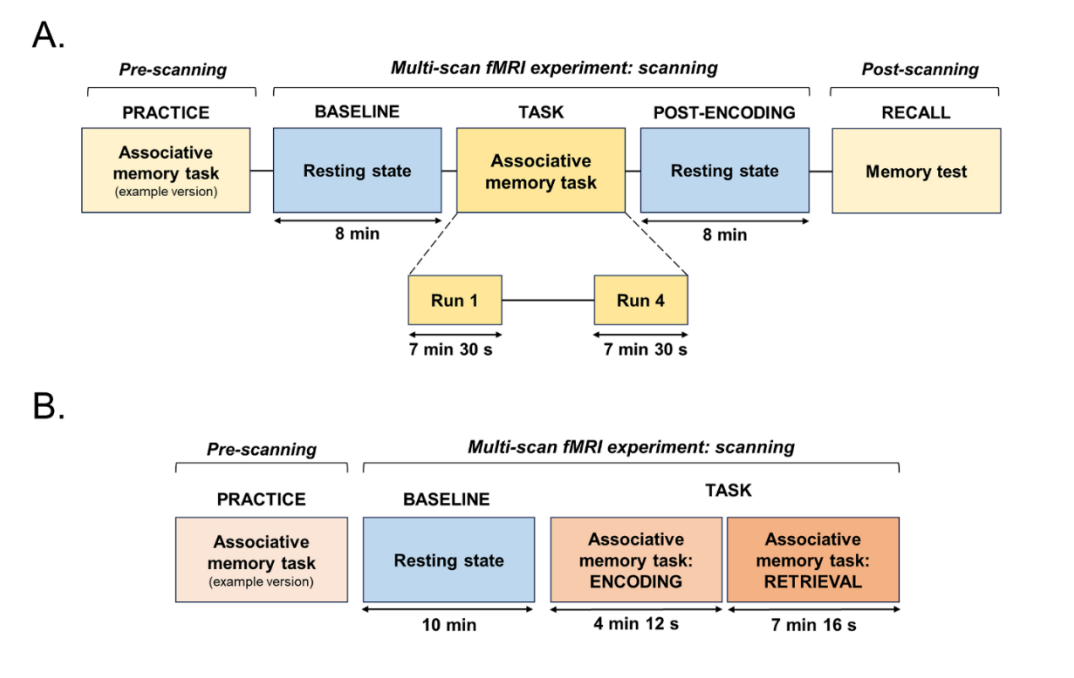

Figure II. 8 The multi-scan fMRI experiment. Diagrams illustrate the experimental procedure for the (A) RUB and (**B**) UNIBA samples.

RUB. During each 8-minute resting-state scan, participants were instructed to remain awake with their eyes open and fixate on the crosshair in the center of a black screen (Greicius et al., 2003; Damoiseaux et al., 2006). The event-related memory task was structured into 4 runs, each composed of a picture-pair association encoding phase followed by a delay and a retrieval phase (Figure II.9A). The whole scan for the associative memory task lasted approximately 30 minutes, with each run lasting 7 minutes and 30 seconds, including a total of 2 minutes and 20 seconds for the encoding, 1 minute and 40 seconds for the delay, and 3 minutes for the retrieval phase as described in Passiatore et al. (2021).

During the encoding phase, 20 picture-pairs per run (80 picture-pairs in total) were presented simultaneously on the screen for 2500 ms. Each picture was selected from the FRIDa dataset (Foroni et al., 2013), representing an object. Each pair consisted of one object on the left and one on the right, with their positions randomized across participants. Participants were instructed to study the association between the picture pairs and indicate which object in each pair was larger or smaller in

reality. Responses were made using two separate button pads ("Left" or "Right"). To ensure a balanced left-right orientation of the stimuli, participants identified the larger object in half of the trials and the smaller object in the other half of the trials. This approach facilitated equal attention to both objects in the pair while maintaining consistency in task demands. Interleaved with the encoding trials, we included control stimuli consisting of 6 pairs per run (24 pairs in total) of two-digit numbers located at the center of visually scrambled picture-pairs with a duration of 2500 ms. To match the motor demands of the "Larger/Smaller" trials, participants performed an "Even/Odd" judgment, responding with two separate button pads ("Left" and "Right" responses). At the beginning of each run, instructions randomly indicated whether the participant should identify the position of the "Even" or "Odd" number. These control trials served to reset the BOLD signal in the MTL to baseline while enhancing task-related activity (Stark and Squire, 2001).

During the delay, 10 novel picture-pairs per run (40 picture-pairs in total) were presented for 2500 ms while participants performed the same "larger/smaller" judgment task but were explicitly instructed not to memorize the associations. The picture-pairs presented during the delay were used to interfere with the previously studied pictures (Marini et al., 2017). Interleaved with the delay trials, we included control stimuli consisting of three pairs per run (12 pairs in total) of two-digit numbers, located at the center of visually scrambled picture pairs, with a duration of 2500 ms. As for the encoding phase, participants were instructed to identify the position of the "Even" or "Odd" number.

During retrieval, participants performed an "Old/New" object recognition task involving 30 single pictures per run (120 pictures in total), each presented for 2000 ms. Of these, 20 objects per run (80 in total) were studied during the encoding phase. The remaining 10 pictures per run (40 in total) were entirely novel. Participants were presented with two text boxes labeled "Old" and "New," corresponding to "Left" or "Right" responses. Interleaved with the retrieval trials, we included control stimuli consisting of 6 pairs per run (24-pairs in total) of two-digit numbers located at the center of

visually scrambled picture-pairs with a duration of 2500 ms. As for the encoding and the delay phase, participants were instructed to identify the position of the "Even" or "Odd" number.

Stimuli were presented via a back-projection system, and behavioral responses were recorded through an optic fiber response box to measure the hits (the count of correct responses) and reaction time (RT, measured in ms) for each trial. The experiment was performed using Presentation® software (Version 23.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Before the scanning session, all participants completed a practice run consisting of the three phases described above, using stimuli that were not included in the actual scanning session. After the scanning session, participants completed a forced-choice association test. They were first presented, one at a time, with all the pictures they had correctly identified as "Old" during the retrieval phase, i.e., cue items. For each cue item, participants rated their confidence in recalling its associated pair using a 4-point Likert scale, where 1 indicated certain recall and 4 indicated guesses. Next, two pictures were displayed on the screen: one was the associated picture from the encoding phase, and the other was a distractor picture shown during the delay. Participants were instructed to select the picture paired with the cue item studied in the encoding phase.

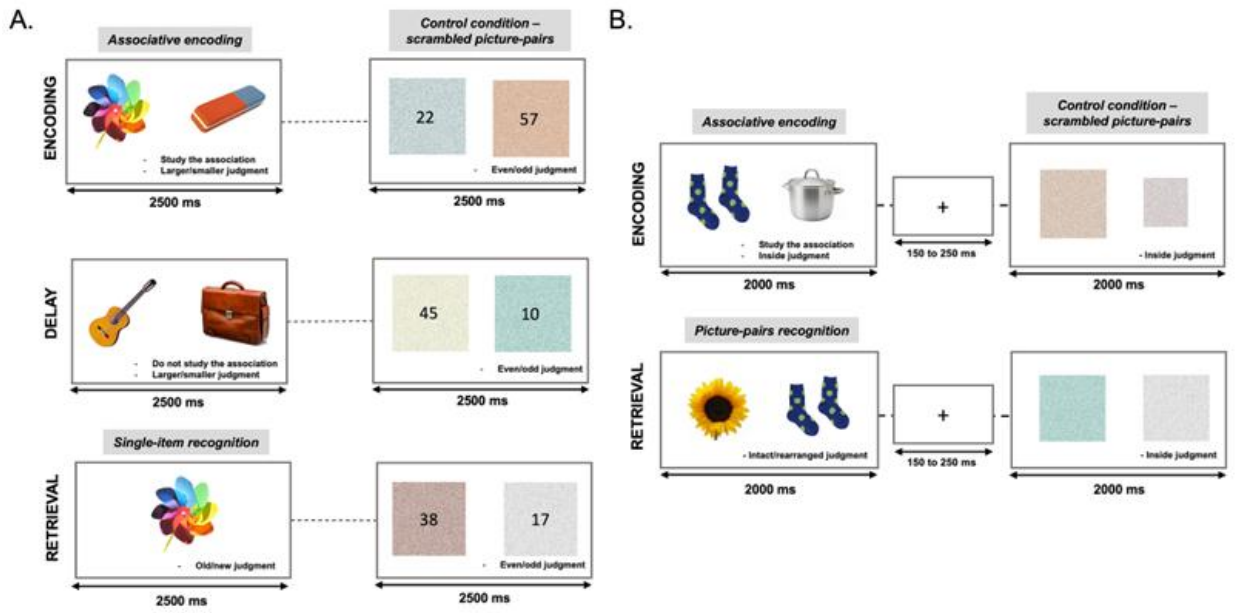

Figure II. 9 The associative memory task. Diagrams show the associative memory paradigms performed during fMRI in the (**A**) RUB and the (**B**) UNIBA samples.

UNIBA. During the 10-minute resting-state scan, participants were instructed to remain awake with their eyes open and to fixate on the crosshair in the middle of a white screen (Greicius et al., 2003; Damoiseaux et al., 2006). We used a revised version of the Relational and Item-Specific Encoding task (RISE; Ragland et al. (2012); Ragland et al. (2015); Passiatore et al. (2023)). The event-related task consisted of a picture-pair association encoding followed by a picture-pair associative retrieval (Figure II.9B). The whole scan for the associative memory task lasted approximately 10 minutes, with the encoding phase lasting 4 minutes and 12 seconds and the retrieval phase 7 minutes and 16 seconds as described in Passiatore et al. (2023).

During the encoding phase, 35 picture-pairs drawn from the FRIDa dataset (Foroni et al., 2013) were presented simultaneously on the screen for 2000 ms each, with one picture on the left and one on the right. Participants completed 35 relational trials (i.e., "Can one object fit inside the other?"), responding with the "Left" and "Right" buttons. They were explicitly instructed to study the association of the picture-pairs. Interleaved with the encoding trials, participants were presented with

10 visually scrambled picture pairs for 2000 ms. Half of the pairs were sized to a standard 530×530 pixels, while in the other half, one picture was upsized and the other downsized by 30% to maintain consistent screen brightness. Participants were instructed to indicate whether "one – scrambled - picture could fit inside the other," matching the "Left/Right" motor demands of the encoding task. The control trials, similar to those at the RUB, served to reset the BOLD signal in the MTL to baseline while enhancing task-related activity (Stark and Squire, 2001).

During the retrieval phase, participants were presented with 70 picture-pairs for 2000 ms each, including 35 "intact" picture-pairs, i.e., presented in the identical configuration of the encoding phase, and 35 "rearranged" picture-pairs. Participants were instructed to indicate whether picture pairs were "intact" or "rearranged" using the "Left" and "Right" buttons, respectively. Interleaved with the retrieval trials, 20 pairs of scrambled pictures were presented for 2000 ms as control stimuli. As for the encoding phase, participants were instructed to indicate whether "one – scrambled - picture could fit inside the other," matching the "Left/Right" motor demands of the retrieval task.

Stimuli were presented via a back-projection system, and behavioral responses were recorded through an optic fiber response box to measure accuracy (the count of correct responses) and RT (ms) for each trial. The experiment was performed using Presentation® software (Version 23.0).

Before the scanning session, all participants completed a practice version of the task that included 5 picture-pairs for the encoding phase, interleaved with 2 scrambled picture-pairs. In the retrieval phase, participants were presented with 5 "intact" pairs and 5 "rearranged" pairs. The stimuli used during training were distinct from those used in the scanning session. More details on stimuli selection are described by Passiatore et al. (2023).

2.2.3. MRI data processing and statistical analysis

2.2.3.1. MRI data acquisition and processing

Characteristics of the MRI acquisitions are shown in Table II.1. We followed the same preprocessing pipeline for both samples using the Statistical Parametric Mapping (SPM) version 12 (http://www.fil.ion.ucl.ac.uk/spm), the Computational Anatomy Toolbox (CAT12, http://dbm.neuro.uni-jena.de/cat/), and the Advanced Normalization Toolbox – ANTs (Avants, 2009). We visually inspected all MRI scans to retain only data unaffected by technical artifacts.

Quality-based inclusion criteria were the absence of blurring, ringing, or wrapping on structural (sMRI) and fMRI raw scans (Wood and Mark Henkelman, 1985; Lu et al., 2019), the absence of excessive cropping (>75%; Fornito et al., 2011; Geerligs and Henson, 2016), and the absence of excessive noise, poor image contrast, or poor boundaries in segmented sMRI scans (Song et al., 2006).

Scanner	Sequence	TR/TE (ms)	Volumes	Voxel size (mm)	FOV (mm)/slices
RUB					
Philips Achieva 3T	sMRI: T1-weighted	9/4.1	1	0.83 × 0.83 × 0.9	288/244
	Rest fMRI: EPI	2,500/35	185	1.44 × 1.44 × 3.2	160/39
	Memory task fMRI: EPI	2,500/26	170 × 4	1.5 × 1.5 × 2	160/29
UNIBA					
Philips Ingenia 3T	sMRI: T1-weighted	8.3/3.8	1	1 × 1 × 1	256/180
	Rest fMRI: EPI	3,000/28	200	3 × 3 × 4	240/36
	Rest fMRI: EPI	2,000/30	240	3 × 3 × 3.6	240/38
	RISE fMRI: EPI	2,000/30	117 + 212	3 × 3 × 3.6	240/38

Table II. 1 Description of the structural and functional MRI scans. Acquisition parameters are reported for RUB and UNIBA samples. Abbreviations: TR = Time Repetition; TE = Time Echo; FOV = Field of view; 3T = 3 Tesla; EPI = Echo planar imaging.

Individual sMRI and fMRI scans were reoriented to the anterior commissure-posterior commissure (AC-PC) line with the origin in the AC. The segmentation of T1-weighted scans served to estimate grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) (Ashburner and Friston, 2005). Bias correction removed intensity non-uniformities to improve coregistration with the fMRI scans. Thalamic structural segmentation on the unbiased individual T1-weighted scans was performed through the Thalamus Optimized Multi-Atlas Segmentation (THOMAS; Su et al. 2019;

Vidal et al., 2024) based on multiple scan registration steps and multi-atlas label fusion. We obtained twelve thalamic segmentations in the individual native space for each participant (Figure II.10, Step 1). The same procedure was performed on the Montreal Neurological Institute (MNI) 152 Template, obtaining twelve thalamic segmentations in the MNI space for group-level parcellations.

Because of the coarse resolution and high correlation of the fMRI signal across neighbor voxels, we anatomically grouped thalamic nuclei into four major subdivisions: anterior, including anterior ventral, ventral anterior nuclei, and the mammillothalamic tract; posterior, including pulvinar, lateral and medial geniculate nuclei; medial, including mediodorsal and centromedian nuclei; ventral, including ventral lateral posterior, ventral posterior lateral and ventral lateral nuclei (Antonucci et al., 2019; Passiatore et al., 2021). The habenula was excluded due to its small size and proximity to the third ventricle, which may affect the accuracy of the related fMRI signal detection.

All echo planar imaging (EPI) fMRI scans were corrected for the differences in acquisition time between slices within a single volume, realigned to the mean scan, coregistered with the unbiased T1-weighted scan, resampled to isotropic voxel sizes of 2 mm^3 for RUB and 3 mm^3 for UNIBA (reflecting their respective original acquisition resolutions) and smoothed using a 6-mm full width at half-maximum isotropic 3D Gaussian kernel. EPI fMRI scans were examined for excessive motion exceeding the initial acquisition resolution (RUB: >2 mm in maximum translation and 1.5 degrees in maximum rotation; UNIBA: >3 mm in maximum translation and 1.5 degrees in maximum rotation; mean FD>0.5 mm for both samples).

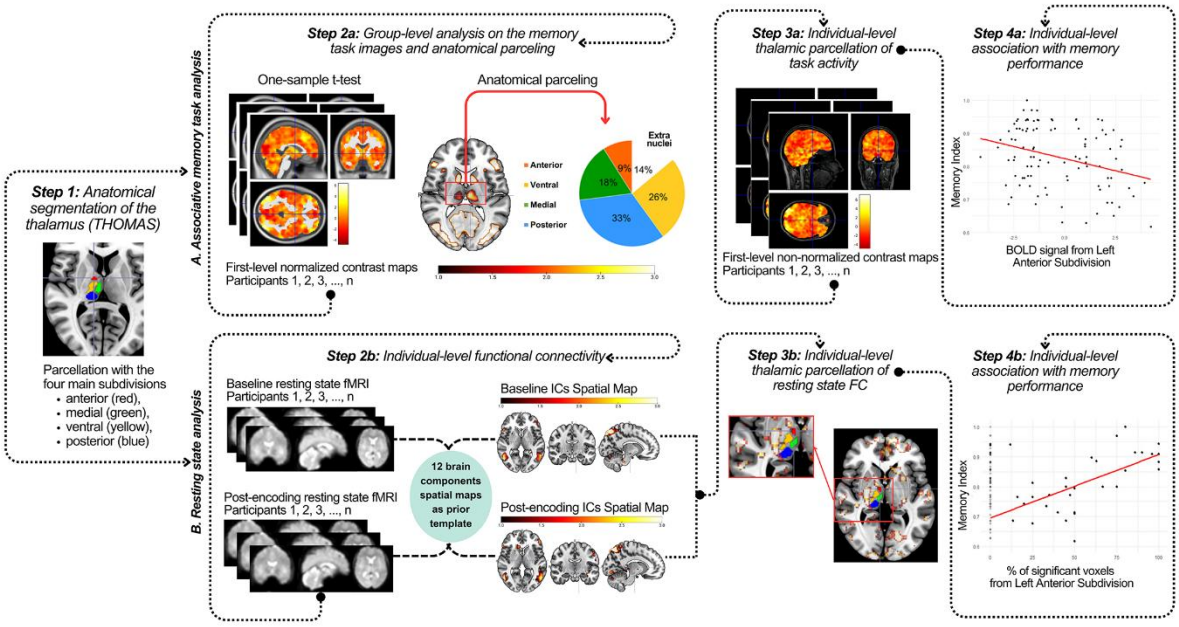

Figure II. 10 Outline of the analysis. Step 1 represents the anatomical segmentation of the thalamus on the individual T1-weighted scans through THOMAS. Panel A summarizes the associative memory task analysis pipeline, including the analysis at the group level on the normalized task activity contrast maps and the anatomical parceling of the thalamic clusters (Step 2A); the analysis at the individual level task activity within thalamic parcellations' by overlapping the individual thalamic subdivisions segmented in the individual native space on the non-normalized task activity maps at α<0.05 (Step 3A); the association between the associative memory index and the BOLD signal extracted from each thalamic subdivision within the individual activity contrasts across task fMRI sessions through multiple regression models (Step 4A). Panel B summarizes the resting state analysis pipeline, including the analysis of group-level networks' spatial maps extracted through ICA on the baseline and post-encoding rs-fMRI sessions (Step 3A); the analysis of individual-level thalamic parcellations within resting state networks' spatial maps by overlapping the individual thalamic subdivisions segmented in the individual native space on the non-normalized resting state networks' spatial maps at α<0.05 (Step 3B); the association between the associative memory index and the percentage of significant voxels in each thalamic subdivision extracted from the resting state networks' spatial maps across rs-fMRI sessions through two-part regression models (Step 4B). Abbreviations: IC = Independent Component; FC = functional connectivity; BOLD = Blood Oxygenation Level Dependent.

At the first level, a general linear model (GLM) was specified to estimate task-related BOLD responses. Each participant's smoothed task-related scans were modeled into a design matrix consisting of one row for each scan and one column for each explanatory variable. Explanatory variables included a stimuli onset vector representing the task events (i.e., Hits, False Alarms - FA, Misses, No Response - NR, Correct Rejections - CR, and control trials), convolved with a canonical hemodynamic response function implemented by SPM12. For the memory task collected at the RUB,

the onset vector incorporated the classification of Hits responses derived from the post scanning memory test, specifically categorized as Hits with a subsequent recall of paired picture (H+), Hits with incorrect subsequent recall, with confidence ratings between 1 and 3 (H-), and Hits followed by a correct or incorrect recall of paired pictures with confidence ratings equal to 4 (H0), as detailed in Table II.2. Additionally, twenty-four motion parameters (Friston, 2003), WM, and CSF mean signals were included as nuisance regressors to account for residual motion artifacts and non-GM signals. The design matrix was high-pass filtered with a cutoff of 128 s to remove low frequency drifts. Global scaling was not applied to the data.

To identify event-related activations, contrasts of interest were computed at the first level (Table 3). For the task collected at the RUB, we modeled the following contrast vectors of +1 for the first condition and -1 for the second condition specified:

- trials involving the successful encoding of pictures, represented by the union of H+ and H- responses (hereafter referred to as H+ ∪ H-), excluding trials associated with ratings equal to 4 (H0) at the post-scanning test, were contrasted with control trials ("Even/Odd" judgment on scrambled picture-pairs) during the encoding phase;
- the successful retrieval of "Old" single items (H+ ∪ H-) was contrasted with control trials during the retrieval phase;
- the successful retrieval of "Old" single items (H+ ∪ H-) was contrasted with the correct identification of "New" single items (CR) during the retrieval phase.

For the task collected at UNIBA, we modeled the following contrast vectors of +1 for the first condition and -1 for the second condition specified:

- trials involving the successful encoding of picture-pairs (Hits; including only trials successfully retrieved during the retrieval phase) were contrasted with control trials (Scrambled picture-pairs) during the encoding phase;

• the successful retrieval of "intact" picture-pairs (Hits) was contrasted with control trials during the retrieval phase;

• the successful retrieval of "intact" picture-pairs (Hits) was contrasted with the correct identification of "rearranged" picture-pairs (CR). For the group-level analysis, individual fMRI contrasts were warped into the standard MNI space using affine and non-linear transformations via the Symmetric Normalization function in ANTs (Avants, 2009; Avants et al., 2011).

2.2.3.2. Behavioral analysis

The response classification in the two experimental paradigms is summarized in Table II.2.

Categories	Sample	Description
Hits	RUB; UNIBA	RUB: responses successfully classifying an old item as old during retrieval of pictures studied during the encoding UNIBA: responses successfully classifying an intact picture-pair as intact during retrieval of picture-pairs studied during the encoding
H^+	RUB	"Hits with subsequent recall of paired picture"
H^-	RUB	"Hits with incorrect subsequent recall, with confidence ratings between 1 and 3"
H^0	RUB	"Hits followed by correct or incorrect recall of paired pictures with confidence ratings equal to 4"
CR	RUB; UNIBA	RUB: responses successfully classifying a new item as new during retrieval of pictures studied during the encoding UNIBA: responses successfully classifying a rearranged picture-pair as rearranged during retrieval of picture-pairs studied during the encoding
FA	RUB; UNIBA	RUB: responses incorrectly classifying a new item as old UNIBA: responses incorrectly classifying a rearranged picture-pair as intact during the retrieval
Misses	RUB; UNIBA	RUB: responses incorrectly classifying an old item as new UNIBA: responses incorrectly classifying an intact picture-pair as rearranged during the retrieval
NR	RUB; UNIBA	The participants failed to respond within the given time
Numbers	RUB	Even/odd number judgment trials with scrambled picture-pairs
Scrambled	UNIBA	Inside judgment events with scrambled picture-pairs with similar/different dimensions

Table II. 2 Behavioral performance classification during the fMRI associative memory task phases in the RUB and the UNIBA samples. The Hits classification during the subsequent memory test in RUB is shown in italics. Note: The description of the Hits and CRs count refers to a single-item recognition in RUB and to picture-pair recognition in UNIBA. Abbreviations: H = hits; CR = correct rejections; FA = false alarms; NR = no responses.

RUB. The memory performance rating was determined by responses during the fMRI retrieval phase and the subsequent forced-choice memory test. We included individuals with a Hits rate >0.65 during the retrieval (Ragland et al., 2012; Passiatore et al., 2021). Since the subsequent memory trial classification included a 'guess' option with a confidence rating of 4 (H0), correct responses could result from accurate recall or correct guessing. To focus on recall trials, H0 trials were excluded when calculating the individual memory index. The memory index was calculated as the ratio between the sum of the number of correct (H+) and partial (H-) recall – with a confidence rating between 1 and 3

– in the post-scanning test (see Table II.2 for a description) and the total number of the valid trials during the encoding [(H+ ∪ H-)/(trials – no responses)] (Pergola et al., 2013a; Passiatore et al., 2021). Additionally, we calculated an index for the correct recognition of a new image as the ratio between the CR and the number of valid trials. We also extracted RTs parsed by response type.

UNIBA. We classified performance based on responses during the fMRI retrieval phase. We included individuals with a Hit rate >0.65 during the retrieval (Ragland et al., 2012; Passiatore et al., 2023). The Hit rate corresponded to our memory index, denoting the correct recognition of 'intact' picture-pairs as the ratio of the Hit count on the total number of valid trials during the encoding [Hits/(trials – no responses)]. We also computed a correct recognition index as the ratio of the CR to the number of valid trials, denoting the accurate recognition of 'rearranged' picture-pairs. We extracted RTs parsed by response type.

In both samples, we used the Shapiro-Wilk test to assess deviations from the normal distribution of the individual memory performance and the correct recognition indices. We performed a Wilcoxon signed-rank test on paired samples to compare the individual memory performance and the correct recognition indices to study any differences in the individuals' ability to retrieve 'old' pictures or 'intact' picture-pairs (Hits) rather than 'new' pictures or 'rearranged' ones (CR), depending on the experimental paradigm in the two samples. In both samples, we also tested the effect of response types (Hits vs. CR) on performance speed (RTs) using the Wilcoxon paired signed-rank test according to previous evidence suggesting faster RTs in retrieving 'new' pictures rather than the 'old' ones (Pergola et al., 2013a; Passiatore et al., 2021).

We used the Wilcoxon rank-sum test to compare the memory indices calculated in the RUB and UNIBA samples to evaluate whether the two different memory task paradigms measured comparable memory performance. To determine effect size (r equivalent), we used the rstatix R package (Fritz et al., 2012; Tomczak and Tomczak, 2014) by dividing the Wilcoxon Z-transformed

values by the square root of the number of observations (Rosenthal et al., 1994; Rosenthal and Rubin, 2003). We also evaluated the associations between the memory index and age and sex in both samples through separate linear models (Dependent variable: memory index; independent: age, sex); also, we assessed the association between the memory index with the intellective quotient and the educational level available only in the UNIBA sample (Dependent variable: memory index; independent: intellective quotient, educational level). We set statistical significance at $\alpha < 0.05$.

2.2.3.3. Group-level task-related activity

We evaluated task-related activity at the group-level by entering the individual normalized contrast maps for each memory phase defined in Table II.3 into separate one-sample t-tests using SPM12 (Figure II.10, panel A, step 2a). All statistics were non-parametrically corrected for multiple comparisons through the Threshold-Free Cluster Enhancement approach (TFCE, Smith and Nichols (2009)). Significance was set at $\alpha < 0.05$, using the false discovery rate (FDR) for multiple comparison correction (Benjamini, 1995).

Phase	Contrast	Description
RUB		
Encoding	$H^+ \cup H^- >$ scrambled picture-pairs	Successful encoding followed by correct and incorrect subsequent recall—with confidence ratings between 1 and 3—of the associated picture versus scrambled picture-pair trials with even/odd numbers
Retrieval	$H^+ \cup H^- >$ scrambled picture-pairs	Successful retrieval followed by correct and incorrect subsequent recall—with confidence ratings between 1 and 3—of the associated picture versus scrambled picture-pair trials with even/odd numbers
	$H^+ \cup H^- >$ CR	Successful retrieval of old pictures versus successful detection of new pictures
UNIBA		
Encoding	Hits > scrambled picture-pairs	Successful encoding of picture-pairs versus scrambled picture-pair trials
Retrieval	Hits > scrambled picture-pairs	Successful retrieval of intact picture-pairs versus scrambled picture-pair trials
	Hits > CR	Successful retrieval of intact picture-pairs versus successful detection of rearranged picture-pairs

Table II. 3 First-level activity contrasts for the fMRI associative memory tasks. Contrast descriptions are reported for the RUB and the UNIBA samples. Abbreviations: H+ = Hits with subsequent recall of paired picture; H- = Hits with incorrect subsequent recall, with confidence ratings between 1 and 3; CR= Correct Rejections.

We extracted the individual BOLD signal from the significant thalamic cluster through Marsbar (Brett et al., 2002). We related thalamic BOLD signal and memory performance through separate multiple regression grouped by memory phases and hemisphere (Dependent variable:

memory index; independent: averaged BOLD signal across the thalamic cluster; nuisance covariates: age, sex, handedness) in the R environment (https://www.r-project.org). Results were corrected for multiple comparisons pFDR<0.05.

To localize the significant thalamic activity into specific thalamic subdivisions, we performed a descriptive anatomical parceling of the group-level activity maps (Antonucci et al., 2019). We used Marsbar to compute the percentage of spatial overlap between the significant thalamic clusters and the thalamic subdivisions in the MNI space obtained with THOMAS (Figure II.10, panel A, step 2a).

2.2.3.4. Individual-level task-related activity and the anatomical parceling of the thalamus

To further investigate interindividual differences in the recruitment of thalamic subdivisions related to memory, we conducted the same analysis in the individual native space to enhance interindividual differences and avoid information loss caused by averaged group-level analysis. We overlapped the individual thalamic subdivisions segmented in the native space and the standardized contrast maps in the individual native space (Figure II.10, panel A, step 3a) to minimize registration bias and maximize sensitivity to detect individual regional effects that can be impacted by registration errors (Friston, 2003; Avants, 2009). Then, we extracted the averaged BOLD signal from the significant voxels ($\alpha <$ 0.05) within a given subdivision.

After scaling the distributions of individual BOLD signals extracted from each thalamic subdivision grouped by memory phases and hemisphere, we explored their relationships with the memory index through separate multiple regression models (Dependent variable: memory index; independent: BOLD signal in the thalamic subdivision; nuisance covariates: age, sex, handedness.

See Figure II.10, panel A, step 4a), setting significance at pFDR<0.05. Analyses were conducted separately for the RUB and UNIBA datasets.

2.2.3.5. Resting-state functional connectivity analysis

To explore the degree of variability in corticothalamic FC during the resting-state, we estimated individual-level FC maps through independent component analysis (ICA) in the baseline resting-state scans in both samples.

For the data acquired at the RUB, we performed the same analysis on the post-encoding resting-state scan. We used the multi-objective optimization ICA algorithm (MOO-ICAR) (Du et al., 2020) implemented in the Group ICA Toolbox (https://trendscenter.org/software/gift/) (Calhoun et al., 2001; Calhoun and Adali, 2012). The MOO-ICAR algorithm extracts individual-level FC spatial maps based on a spatial reference (Figure II.10, panel B, step 2b). We employed the spatial reference from Iraji et al. (2019) to extract reliable FC patterns in high-order functional networks, including attentive, auditory, cerebellar, default mode (DMN), language, bilateral frontoparietal (FPN), sensorimotor, salience, primary and secondary visual networks topologically determined based on their spatial and temporal properties from previous studies (Beckmann et al., 2005; Damoiseaux et al., 2006; Damoiseaux et al., 2008; Smith et al., 2009; Allen et al., 2011; Iraji et al., 2016). We inversely warped the spatial reference into each individual native space to match the preprocessed non-normalized resting-state scans used as the input. Additionally, twenty-four motion parameters (Friston, 2003) were regressed out from the individual FC network spatial maps to mitigate the impact of motion on connectivity estimates. For the group-level analysis, the obtained individual FC networks spatial maps were warped into the standard MNI space using affine and non-linear transformations via the Symmetric Normalization function in ANTs.

We evaluated the thalamic recruitment within each cortical network at the group level by entering each network's normalized individual FC maps into separate one-sample t-tests using SPM12. Significant clusters were defined at $\alpha < 0.05$ and corrected for multiple comparisons using the family-wise error rate (FWE). We employed the mask of the whole thalamus in the MNI space segmented

through THOMAS to localize the effects. Then, we extracted the IC loadings from the thalamic significant cluster through Marsbar. We associated thalamic IC loadings and memory performance through separate multiple regression grouped by network and hemisphere (Dependent variable: memory index; independent: averaged IC loadings across the thalamic cluster; nuisance covariates: age, sex, handedness), setting significance at pFDR<0.05.

2.2.3.6. Association between individual-level functional connectivity of thalamic parcellations during resting-state and memory performance

To further investigate interindividual differences in terms of spatial recruitment of cortico-thalamic FC related to memory, we calculated the degree of spatial overlap between the thalamic subdivisions and the significant voxels included in each non normalized spatial networks' map at the individual level (Figure II.10, panel B, step 3b), defining voxel significance at α <0.05. Networks' spatial maps include both positive and negative FC values. Positive values denoted correlated voxels to the individual time course, while negative values denoted anticorrelated voxels to the individual time course (Calhoun et al., 2001; Calhoun and Adali, 2012). We separately extracted the number of voxels significantly correlated and anticorrelated within each network's spatial map, along with the averaged FC from each thalamic subdivision.

To investigate the relationship between corticothalamic recruitment and memory performance across individuals, we analyzed the associations of memory performance with both (i) the proportion of significant voxels in each thalamic subdivision over the number of voxels in the entire thalamus grouped by hemisphere (Figure II.10, panel B, step 4b), and (ii) the averaged FC extracted bilaterally from each thalamic subdivision.

For the analysis of the proportion of significant voxels in each thalamic subdivision, we quantified the number of individuals showing non-zero values in thalamic subdivision recruitment,

since the spatial localization of thalamic signals may vary at the individual level, possibly not covering all the thalamic subdivisions. Thus, we restricted the analyses to thalamic subdivisions, including at least N=10 non-zero observations (Table S.1). We used two-part regression models for zero-inflated data using the *twopartm* R package (Belotti et al., 2015). A preliminary natural logarithmic transformation was applied to the proportion of recruited voxels in each thalamic subdivision to mitigate the right-skewness of the distributions. The obtained values were used as the predictor in the second part of the model, while a binary variable indicating whether thalamic recruitment was equal to zero (0) or not (1) was used as an independent variable in the first part of the model. In the first part of the model, a binomial regression was fitted with a logit link function to measure the probability of a non-zero outcome. The second part of the model is a GLM (Dependent variable: memory index; independent: proportion of significant voxels in the thalamic subdivision; nuisance covariates: age, sex, handedness). To facilitate the interpretation of results, we also calculated the averaged marginal effects (AME; Belotti et al. (2015)), i.e., the estimated impact of the proportions of thalamic voxels recruited on the memory index calculated by considering the combined effect of the probability of the thalamic recruitment occurring (first part) and the magnitude of the thalamic recruitment association with the memory index (second part), and the confidence intervals (CI) at 0.95.

Similarly, for the analysis of FC in each thalamic subdivision, GLMs were used to test the effect of memory performance on the averaged FC, since FC did not include zero values (Dependent variable: memory index; independent: averaged FC in the thalamic subdivision; nuisance covariates: age, sex, handedness). All analyses were conducted separately for the RUB and UNIBA datasets.

In the RUB sample, we also computed separate two-part regressions and GLMs for the post encoding resting-state spatial maps to assess changes in cortico-thalamic recruitment after the performance of the associative memory task. In addition, we analyzed the association of memory performance with the variation in terms of thalamic recruitment from the baseline to the post

encoding resting-state. We computed a change score representing the difference in the proportion of significant voxels between the post- and the baseline resting-state scans for each thalamic subdivision, as well as the difference of the FC between post-encoding and baseline resting state.

Positive values indicate an increase in the proportion of significant voxels or FC from the baseline to the post-encoding resting state, while negative values indicate a decrease in the proportion of significant voxels or FC from the baseline to the post-encoding resting state. Zero values indicate no change in the proportion of significant voxels across resting-state scans. We tested the relationship between memory performance and the change score for each thalamic subdivision within each network and grouped by hemisphere through separate GLM (Dependent variable: memory index; independent: change score; nuisance covariates: age, sex, handedness), setting significance at pFDR<0.05.

2.2.3.7. Path analysis on the individual-level thalamic recruitment across resting-state and memory phases

To investigate the relationships between the individual associative memory performance and the individual-level cortico-thalamic recruitment across baseline resting-state FC, the activity during the encoding and the retrieval phases, serial path analyses within a SEM framework were implemented using the lavaan R package (Rosseel, 2012). We included the thalamic recruitment as the percentage of significant voxels extracted during the baseline resting-state IC spatial maps as the predictor (X); the individual task-related activity in the encoding (M1) and retrieval (M2) during the associative memory task as mediators; the index of the associative memory performance as the outcome (Y). The path analysis considered the thalamic subdivisions significantly associated with associative memory performance in the previous analysis. Age, sex, and handedness were used as covariates, and a dichotomic variable was used to control for non-zero observations. Significance was set as $\alpha < 0.05$

(bootstrap: 5,000 runs). We used the *effectsize* R package (Ben-Shachar et al., 2020) to calculate the magnitude of the effects reported as R2 (Liu et al., 2023).

2.2.3.8. K-means clustering on the individual-level thalamic recruitment during the associative memory task

To test our hypothesis that cortico-thalamic circuits recruited by different individuals may underlie variability in memory performance, we performed a K-means clustering using the *kmeans* R function (MacQueen, 1967; Hartigan and Wong, 1979) on the individual task-related activity during i) the encoding and ii) the retrieval phases during the associative memory task, previously included in the indirect paths of the path analysis. Also, the cluster analysis considered the thalamic subdivisions significantly associated with memory performance in the previous analysis. We performed clustering only on the UNIBA dataset, which included enough observations to capture variability across individuals. We determined the optimal number of clusters for the left and right hemispheres through the NbClust R package (Charrad et al., 2014). We employed the majority rule to select the optimal number of clusters. The cluster stability was calculated through the *bootcluster* R package considering S >0.8 (bootstrap: 5,000 runs; Yu et al. (2019)). We separately compared the memory index, the proportion of significant voxels within the baseline resting-state spatial maps, the individual BOLD estimates during the encoding and retrieval phases during task performance, age, education, and intellective quotient across clusters through the Wilcoxon rank-sum test, and the sex was compared through the Chi-Square test. We set the test significance at pFDR<0.05.

2.3. Results

2.3.1. Behavioral results

This section presents the results of the behavioral analysis. Table II.4 shows behavioral performance on the RUB task, as already published by Passiatore et al. (2021), along with behavioral performance collected at UNIBA.

Categories	*N* mean ± SD	Total *N* of trials	% Mean ± SD	Mean RT ± SD (ms)	RT min:max (ms)
RUB					
Hits	68 ± 6	80	86.02 ± 7.6	2,513 ± 406	1,735:3,131
H^+	40 ± 13	80	50.31 ± 16	1,099 ± 63	889:1,148
H^-	22 ± 11	80	27.72 ± 14.7	1,064 ± 62	846:1,116
H^0	6 ± 7	80	7.99 ± 9.97	349 ± 385	0:956
CR	36 ± 2	48	76.56 ± 5.54	1,178 ± 59	925:1,245
FA	2 ± 2	48	5.65 ± 4.93	12 ± 65	0:346
Misses	10 ± 6	80	12.09 ± 7.03	1,416 ± 100	972:1,484
NR	1 ± 2	128	0.94 ± 1.2	-	-
UNIBA					
Hits	28 ± 3	35	80.38 ± 9.31	1,119 ± 126	764:1,465
CR	31 ± 3	35	88.99 ± 8.34	1,130 ± 125	789:1,500
FA	3 ± 3	35	8.95 ± 7.49	1,326 ± 232	812:1,856
Misses	6 ± 3	35	16.58 ± 9.3	1,296 ± 212	845:1,957
NR	2 ± 2	70	6.47 ± 7.9	-	-

Table II. 4 Behavioral performance during the associative memory task in the RUB and UNIBA samples. The Hits classification during the subsequent memory test in RUB is shown in italics. Abbreviations: CR = Correct Rejections; FA = False Alarms; H+ = Hits with subsequent recall of paired picture; H- = Hits with incorrect subsequent recall, with confidence ratings between 1 and 3; H0 = Hits followed by correct or incorrect recall of paired pictures with confidence ratings equal to 4; NR = No responses; RT = Reaction Time; SD = Standard Deviation.

RUB. The memory index ranged between 0.4 and 0.9 (Mean ± SD = 0.78 ± 0.13), significantly deviating from the normal distribution (Shapiro-Wilk test, p=7e-03). The correct rejection index – computed on CR responses of "New" single objects – ranged between 0.48 and 0.83 (Mean ± SD = 0.76 ± 0.08), also deviating from the normal distribution (Shapiro-Wilk test, p=4e-04). Thus, non-parametric statistical tests were preferred for comparing behavioral indices.

The Wilcoxon paired rank-sum tests comparing the Hits and CR responses showed no significant differences (Z=-1.42, r=-0.2, p=0.16), while when comparing RTs related to Hits and CR, participants

were faster in recognizing "Old" single objects compared to the "New" ones (Z= 5.78, r=-0.87, p=7.4e-09). The memory index was not significantly associated with either age (t(27)=-0.6; p=0.6) or sex (t(27)=-2.1; p=0.07), excluding related nuisance effects on the performance levels.

UNIBA. The memory index ranged between 0.6 and 1 (Mean ± SD = 0.82 ± 0.1), significantly deviating from the normal distribution (Shapiro-Wilk test, p=3e-03). The correct rejection index – computed on CR responses of the 'rearranged' picture pairs – ranged between 0.65 and 1 (Mean ± SD = 0.91 ± 0.07), also deviating from the normal distribution (Shapiro-Wilk test, p = 8.91e-07).

The Wilcoxon paired rank-sum tests comparing the Hits and CR responses showed that participants better recognized "Intact" picture-pairs than the "Rearranged" ones (Z=-6.05, r=-0.59, p=1.49e-09), while, when comparing RTs related to Hits and CR, we found no significant effects of response type on RTs (Z=-1.01, r=-0.09, p=0.32). The memory index was not associated with age (t(106)=-0.4; p=0.7), sex (t(106)=-0.18; p=0.9), intellective quotient (t(106)=-0.8; 0.4), and educational level (t(106)=1.046; p=0.3), excluding related nuisance effects on the performance levels.

The Wilcoxon rank sum test on the memory indices across samples did not show a significant difference (Z=-1.18, r=-0.1, p=0.24). Therefore, despite the differences in the experimental paradigms, the two samples did not show significant differences in memory performance.

2.3.2. fMRI results

2.3.2.1. Group-level activity shows thalamic involvement during task performance

This section presents the results of the group-level task-activity analyses conducted across both samples (RUB and UNIBA). Task-activity analyses at the group level showed consistent activation patterns across samples (pTFCE-FDR <0.05), extensively reported in Table II.5.

Phase: contrast	Region	Side	MNI coordinates			Cluster extent (voxels)	Z value
			x	y	z		
RUB							
Encoding: $H^+ \cup H^-$ > scrambled picture-pairs	Middle temporal gyrus	R	38	−78	20	13,693	6.56
	Middle occipital gyrus	L	−38	−80	12	1,643	6.34
	Caudate	L	−18	14	12	243	2.97
	Insula	L	−38	−4	14	436	2.28
	BA30	R	18	−68	8	64	2.81
	Middle frontal gyrus, Frontal inferior orbital gyrus	R	32	36	−18	549	4.34
	Middle temporal gyrus	L	−36	−78	16	1,694	6.10
	Thalamus	R	6	−16	6	352	4.05
	Thalamus	L	−8	−10	0	258	3.43
	Precuneus	L	−6	−52	18	132	3
	Hippocampus	L	−18	−6	−22	211	3.37
	Hippocampus	R	22	−8	−22	256	3.04
	Amygdala	R	26	−2	−22	39	3.62
	Fusiform gyrus	L	−30	−2	−38	1,260	4.84
	BA20	L	−30	−38	−20	445	6.03
	Fusiform gyrus	R	32	−38	−18	1,248	6.03
	Inferior frontal gyrus, BA47	L	−32	34	−16	768	5.57
	BA17	R	16	−76	8	16	2.17
	Middle occipital gyrus, BA19	R	42	−80	10	32	5.79
Retrieval: $H^+ \cup H^-$ > scrambled picture-pairs	Middle temporal gyrus	R	42	−84	16	3,371	5.32
	Fusiform gyrus, BA19	R	36	−86	12	964	5.13
	Fusiform gyrus	L	−34	−50	−12	505	5.05
	Parahippocampal gyrus	R	24	−38	−16	614	4.69
	Thalamus	R	18	−22	12	255	3.32
	Thalamus	L	−18	−16	10	74	2.9
	BA28	L	−18	−16	−14	36	4.14
	Amygdala	L	−22	0	−14	104	4.15
	Parahippocampal gyrus	L	−34	−18	−28	184	4.19
	Hippocampus	L	−28	−20	−16	252	4.16
	Cuneus	L	−14	−60	20	334	4.85
Retrieval: $H^+ \cup H^-$ > CR	Angular gyrus	L	−54	−62	32	288	3.84
	Caudate head	L	−8	6	0	795	4.31
	Middle temporal gyrus, BA22	R	58	−50	12	158	3.51
UNIBA							
Encoding: Hits > scrambled picture-pairs	Visual associative cortex – Lingual gyrus	R	33	−76	−13	726	9.14
	Visual associative cortex – Lingual gyrus	L	−27	−76	−10	374	9.13
	Parahippocampal gyrus, BA19	R	27	−52	−7	433	8.98
	Fusiform gyrus	R	30	−67	−13	329	9
	Fusiform gyrus	L	−30	−64	−13	342	8.73
	Inferior frontal gyrus, BA47	R	42	26	−7	106	3.99
	Inferior frontal gyrus	R	54	32	11	52	4.06
	Inferior frontal gyrus, BA47	L	−48	32	11	345	7.73
	Thalamus	L	−18	−31	−1	35	6.11
	Thalamus	R	21	−28	−1	39	6.26
	Parahippocampal gyrus	R	24	−19	−19	130	5.62
	Amygdala	L	−18	−7	−16	46	6.34
	Amygdala	R	21	−1	−19	46	5.19
	Hippocampus	L	−33	−16	−16	160	3.61
	Frontal lobe, inferior frontal gyrus	R	54	32	11	52	4.06
Retrieval: Hits > scrambled picture-pairs	Fusiform gyrus, BA19	L	−27	−52	−10	438	9.07
	Fusifom gyrus, BA37	L	−30	−43	−16	339	9.02
	Parahippocampal gyrus	L	−24	−61	−10	363	8.96
	Inferior frontal gyrus	R	42	26	20	149	7.59
	BA6	R	39	−13	47	152	5.14
	Precentral gyrus	R	27	−22	59	136	3.74
	Postcentral gyrus, BA4	R	30	−34	62	45	3.97
	Postcentral gyrus, BA3	R	39	−28	56	41	2.38
	Superior frontal gyrus, BA8	L	−3	14	53	240	4.70
	Medial frontal gyrus	L	−6	26	47	58	4.12
	Cingulate gyrus, BA32	R	9	20	44	51	3.53
	BA6	R	6	14	50	34	3.31
	Thalamus	L	−3	−19	−1	150	3.33

Table II. 5 Whole-brain activation during the associative memory task in the RUB and the UNIBA samples. Peak activations were detected through one-sample t-tests on the following contrasts: (i) H+∪ H- > Scrambled picture-pairs during encoding and retrieval (RUB), (ii) Hits > Scrambled picture-pairs during encoding and retrieval (UNIBA), (iii) H+∪ H- > CR during retrieval (RUB; UNIBA). All peaks are shown at pTFCE-FDR<0.05; cluster extent = 10 voxels. Abbreviations: MNI = Montreal Neurological Institute; BA = Brodmann Area; PFC = prefrontal cortex; L = left; R= right.

RUB. In the encoding contrast, i.e., RUB: H+ ∪ H- > Scrambled picture-pairs, we found activation in the bilateral fusiform gyrus, the bilateral temporal gyrus, the right middle occipital gyrus, the right amygdala, the bilateral hippocampus, and the thalamus. The anatomical parceling of the active voxels falling into the thalamus showed that the effects were bilaterally located in the posterior subdivisions (Figure II.11A). We observed significant activations in the left fusiform gyrus, the BA19, the bilateral parahippocampal gyrus, the left hippocampus, the left amygdala, and the bilateral thalamus within the retrieval contrast, i.e., H+ ∪ H- > Scrambled picture-pairs. The anatomical parceling showed that the thalamic clusters were bilaterally located in the posterior and the ventral subdivisions (Figure II.11A). During activation contrasts related to the successful recognition of "Old" single objects compared with the successful detection of "New" single objects, i.e., H+ ∪ H- > CR, we found activation in the left cuneus, left angular gyrus, the left caudate head, the middle temporal gyrus, and the thalamus. The thalamic cluster mainly overlapped with the right anterior subdivision (Figure II.11A).

When testing the association between the averaged BOLD signal extracted from the thalamic clusters detected at the group level and the memory index, the BOLD signal extracted from the right thalamic cluster within the retrieval contrast, i.e., H+ ∪ H- > CR, was associated with higher memory performance (t(23)=2.9; r=0.51; p=9e-03; pFDR=0.04); instead, the averaged BOLD signal from the left thalamic cluster within the encoding contrast, i.e., H+ ∪ H- > Scrambled picture-pairs, was not significantly associated with the memory index (t(23)=-0.89; r=-0.18 puncorr=0.38).

UNIBA. In the encoding contrast, i.e., Hits > Scrambled picture-pairs, we found activation in the bilateral fusiform gyrus, the bilateral visual associative cortex, the right parahippocampal gyrus, the bilateral amygdala, the left hippocampus, and the bilateral thalamus (Figure II.11B). The anatomical parceling of the active voxels falling into the thalamus within the encoding contrast showed that the effects were bilaterally located in the posterior subdivision (Figure II.11B). We observed

significant activations in the left fusiform gyrus, the bilateral parahippocampal gyrus, the left hippocampus, the right caudate head, and the thalamus within the retrieval contrast, i.e., Hits > Scrambled picture pairs. The anatomical parceling showed that the thalamic clusters were bilaterally located in the posterior and the ventral subdivisions (Figure II.11B). During activation contrasts related to the successful recognition of "Intact" picture-pairs compared with the successful detection of "Rearranged" picture-pairs, i.e., Hits > CR, we found activation in the left middle frontal gyrus, the left ACC, the bilateral precuneus, the left caudate head, and the bilateral thalamus. The thalamic cluster mainly overlapped with the left ventral, the posterior, and the medial subdivisions (Figure II.11B).

Differently from the results in the RUB sample, when studying the associations between the thalamic signal at the group level and the memory performance, the BOLD signal extracted from the left thalamic cluster within the encoding contrast, i.e., UNIBA: Hits > Scrambled picture-pairs, was significantly associated with higher memory performance ($t(97)=2.86$; $r=0.28$; $p=5e\text{-}03$; $pFDR=0.03$); instead, the averaged BOLD signal extracted from the right thalamic cluster within the retrieval contrast, i.e., Hits > CR, was not associated with the memory index ($t(97)=0.97$; $r=0.1$ $puncorr=0.33$).

In summary, although the retrieval demands may differ in that single pictures were presented in the task collected at RUB and picture-pairs were presented in the task collected at UNIBA, the activation patterns, including those involving the thalamus, largely exhibited consistency across samples. Despite the whole-brain similarity in activity patterns between the two samples during the encoding, some disparities emerged when comparing thalamic parcellations in the retrieval phase contrasts. These differences may be attributed to variations in retrieval demands, sample size, and differences in image acquisition resolution between the samples (Turner et al., 2018; Bossier et al., 2020; Alkemade et al., 2022; Sambuco, 2022; Taylor et al., 2023). Moreover, while some significant associations between memory performance and the thalamic signal localized at the group level were

detected, associations were not consistent across memory phases and samples. This inconsistency may reflect individual variability in the spatial localization of thalamic activation.

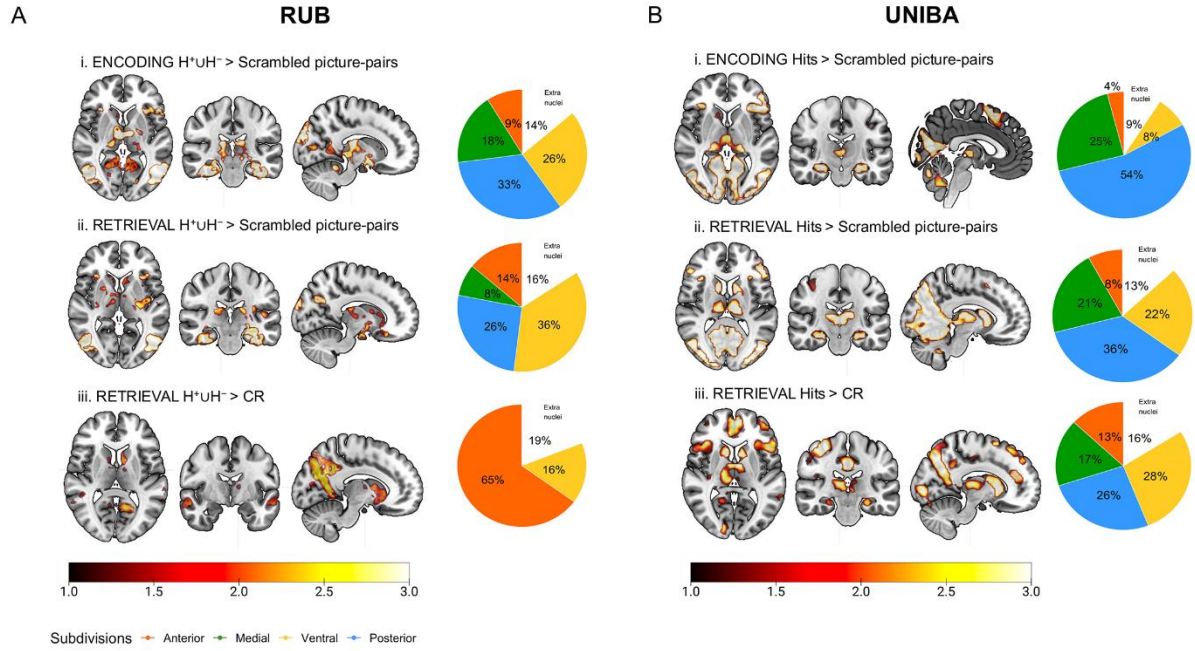

Figure II. 11 Group-level activity during the associative memory task. On the left, sections depict whole brain activation resulting from the one-sample t-tests on individual contrasts during the RUB sample's associative memory task fMRI sessions ($p_{FDR\text{-}TFCE}$<0.05, k=10). On the right, pie charts show the anatomical parceling of the thalamic clusters in the following contrasts: (i) H+; H- > Control during encoding, (ii) H+; H- > Control during retrieval, (iii) H+; H- > CR during retrieval. **B.** On the left, sections depict whole brain activation resulting from the one sample t-tests on the individual contrasts during the associative memory task fMRI sessions in the UNIBA sample ($p_{FDR\text{-}TFCE}$<0.05, k=10). On the right, pie charts show the anatomical parceling of the thalamic clusters in the following contrasts: (i) Hits > Control during encoding, (ii) Hits > Control during retrieval, (iii) Hits > CR during retrieval. Abbreviations: H+ = Complete recall; H− = Partial Recall; Hits = Responses successfully classifying an unchanged picture-pair; CR = Responses successfully classifying a new item as new in RUB and a changed picture-pair as changed in UNIBA.

2.3.2.2. Individual-level medial and anterior thalamic activity during the associative memory task is associated with memory performance

Task-activity analysis at the individual level consistently showed a significant involvement of the anterior and the medial subdivisions of the thalamus across samples, as shown in Figure II.12.

RUB. During the encoding, i.e., the H+ ∪ H- > Scrambled picture-pairs contrast, the decreased activity of the bilateral anterior subdivision was significantly associated with a higher memory index (Left: t(23)=-2.98; r=-0.52; p=7e-03; pFDR=0.03. Right: t(23)=-3.53; r=-0.60; p=1e-03; pFDR=7e-03; Figure II.12A-B). During the retrieval, i.e., the H+ ∪ H- > Scrambled picture-pairs contrast, increased activity of the bilateral medial subdivision was significantly associated with a higher memory index (Left: t(23)=3.54; r=0.6; p=1e-04; pFDR=7e-03. Right: t(23)=3.80; p=9e-04; r=0.63; pFDR=4e-02; Figure II.12A-C). No significant associations were observed in the ventral and posterior subdivisions, nor in the H+ ∪ H- > CR contrast during retrieval (pFDR>0.05).

UNIBA. Similarly to what we observed in RUB, during the encoding of the associative memory task, i.e., Hits > Scrambled picture-pairs, the decreased activity of the bilateral anterior subdivision was significantly associated with a higher memory index (Left: t(97)=-2.47; r=-0.24; p=7e-03; pFDR=0.03. Right: t(97)=-2.71; r=-0.28; p=3e-03; pFDR=0.01; Figure II.12A-B). During the retrieval, i.e., the Hits > Scrambled picture-pairs contrast, increased activity of the bilateral medial subdivision was significantly associated with a higher memory index (Left: t(97)=4.29; r=0.40; p=3e-05; pFDR=8e-04. Right: t(97)=3.44; r=0.33; p=4e-04; pFDR=2e-03; Figure II.12A-C). No significant associations were identified in the ventral and posterior subdivisions (pFDR>0.05), as shown in Figure II.12A, nor in the Hits > CR contrast during retrieval (pFDR > 0.05).

In summary, examining the BOLD signal extracted from thalamic subdivisions in individual native space, thereby enhancing individual signal localization, we consistently detected activation patterns associated with the associative memory performance across samples during the encoding and retrieval phases, in contrast to the results observed at the group level.

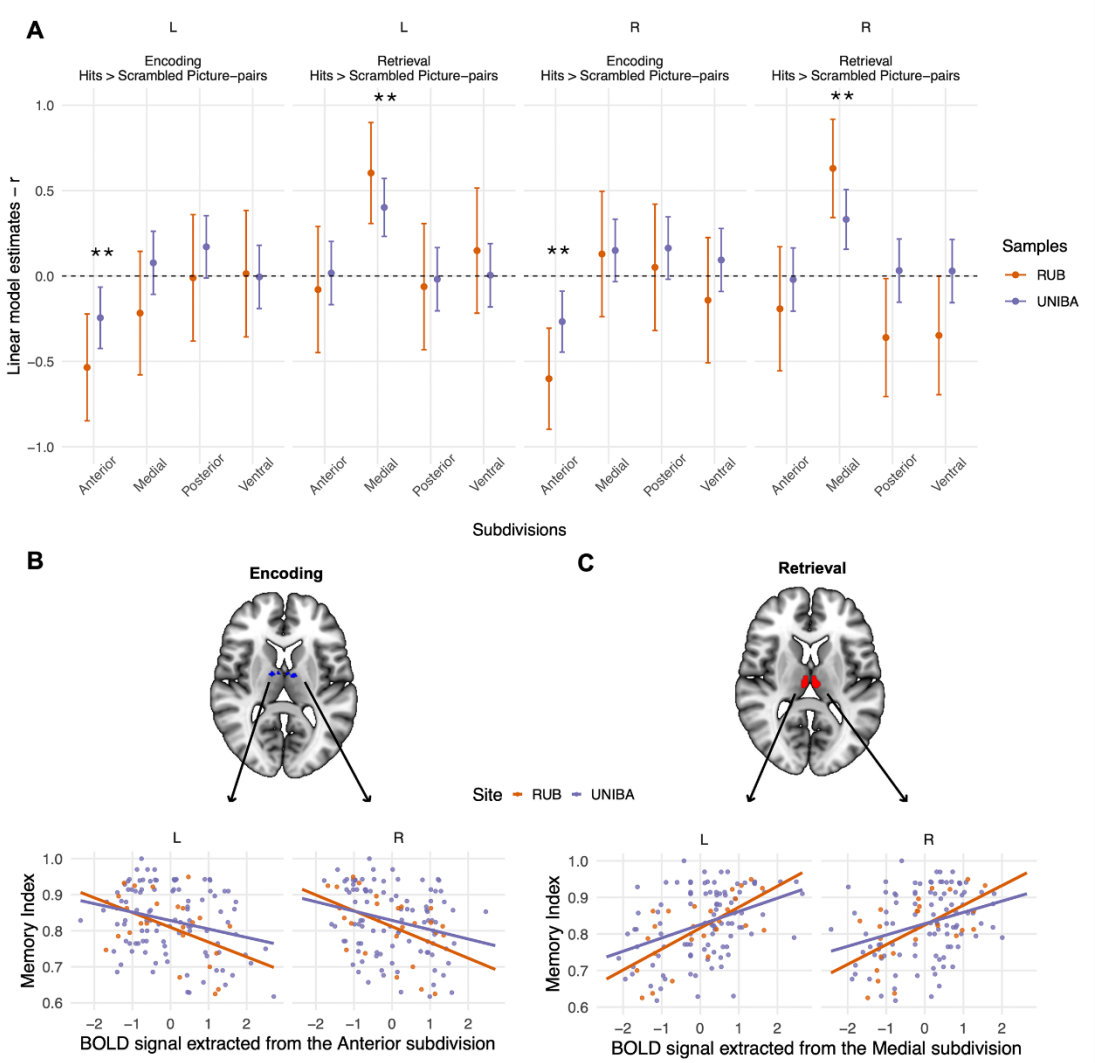

Figure II. 12 Individual-level association between thalamic parcellations during task performance and the associative memory index. A. Barplots show the regression's r estimates from the association between the BOLD signal extracted from thalamic subdivisions during the encoding and the retrieval sessions and the associative memory index. Significant results at $p_{FDR}<0.05$ are marked with asterisks. B. Section shows the anterior thalamic parcellation in blue, including anterior ventral, ventral anterior nuclei, and mammillothalamic tract segmented with THOMAS. The scatterplots show the negative association between the BOLD signal extracted from the anterior subdivision and the associative memory index during encoding across samples (RUB, Left: t(27)=-2.98; r=-0.54; p=7e-03; p_{FDR}=0.03. Right: t(27)=-3.53; r=-0.60; p=1e-03; p_{FDR}=7e-03. UNIBA, Left: t(101)=-2.47; r=-0.24; p=7e-03; p_{FDR}=0.03. Right: t(101)=-2.71; r=-0.28; p=3e-03; p_{FDR}=0.01). C. Section shows the medial thalamic parcellation in red, including the mediodorsal and the centro-median nuclei segmented with THOMAS. The scatterplots show the positive association between the BOLD signal extracted from the medial subdivision and the associative memory index during retrieval across samples (Left: t(27)=3.54; r=0.6; p=1e-04; p_{FDR}=7e-03. Right: t(27)=3.80; p=9e-04; r=0.63; p_{FDR}=4e-02. UNIBA, Left: t(101)=4.29; r=0.40; p=3e-05; p_{FDR}=8e-04. Right: t(101)=3.44; r=0.33; p=4e-04; p_{FDR}=2e-03). Abbreviations: R=right; L=left.

2.3.2.3. Individual-level medial thalamic parcellations during the baseline resting-state within the frontoparietal network are associated with memory performance

Group-level one-sample t tests on normalized functional networks' spatial maps detected significant thalamic signals within the DMN, the left FPN (LFPN), and the right FPN (RFPN). Statistics are reported in Table II.S1. However, the BOLD signal extracted from the significant clusters showed no significant association with the memory index (pFDR>0.05).

Instead, when evaluating the thalamic contribution within the DMN (Figure II.13A), LFPN (Figure II.13B), and RFPN (Figure II.13C) in the individual native space, the regression models showed significant associations between the medial and posterior subdivisions during the baseline resting-state with the memory index. GLM results are depicted in Figure II.13D and detailed below. Table II.S2 shows the quantification of non-zero observations for each thalamic subdivision across cortical networks.

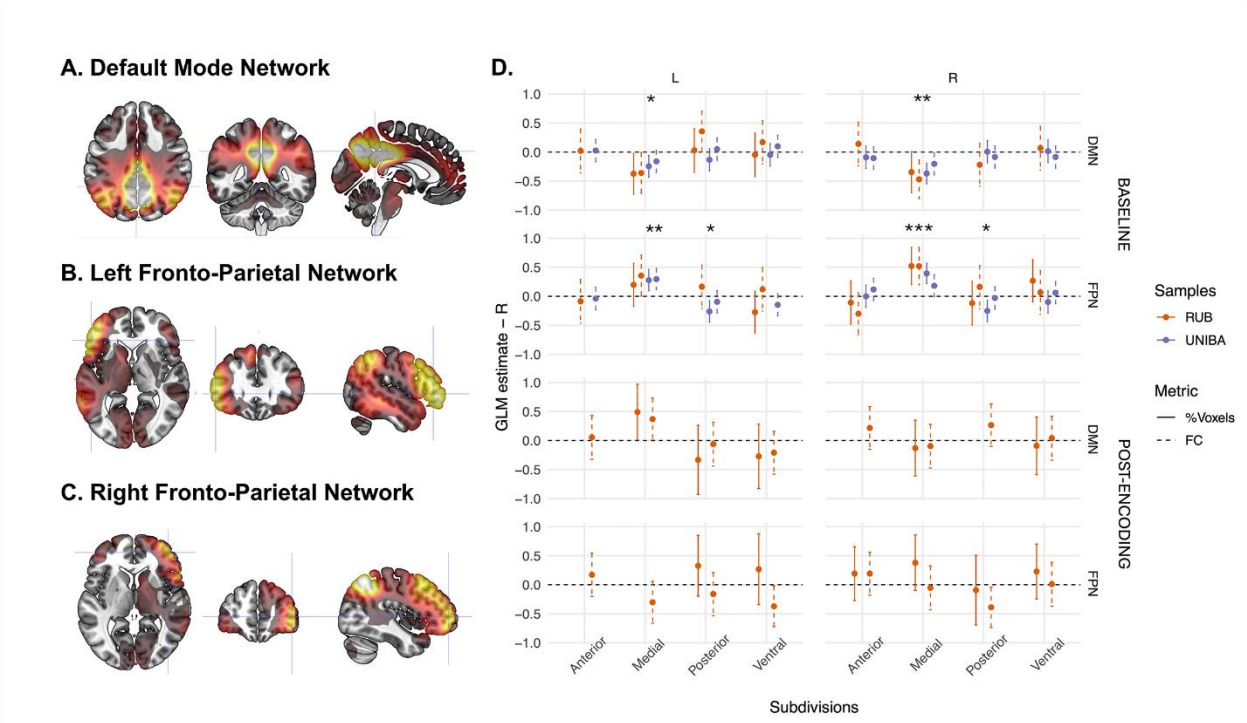

Figure II. 13 Individual-level association between thalamic parcellations during baseline resting state and the associative memory index. Sections depict the group-level networks' spatial maps at Z>3 of the (A) LFPN, (B) RFPN, and (C) DMN. D. Barplots show the generalized linear regression's r estimates resulting from the two-part models associating the percentage of significant voxels extracted from the four thalamic subdivisions across resting state sessions and the associative memory index. Significant results at $p_{FDR}<0.05$ are marked with asterisks. Abbreviations: R=right; L=left; DMN = default mode network; RFPN = right frontoparietal network; LFPN = left frontoparietal network.

RUB. During baseline resting-state, the memory index was positively associated with a greater proportion of significant voxels (t(24)=2.99; r=0.52; p=0.01; pFDR=0.03; AME=0.08; 95% CI=[0.03, 0.14]) and greater FC (t(24)=2.97; r=0.51; p=7e-03; pFDR=0.05) in the right medial subdivision within the RFPN (Figure II.13, upper right panel D), while the left medial subdivision in the LFPN did not reach statistical significance testing the proportion of significant voxels (t(24)=0.98; r=0.19; p=0.33; AME=0.05; 95% CI=[-0.05, 0.14]) and the FC (t(24)=1.86; r=0.35; p=0.07; Figure II.13, upper left panel D). In the opposite direction, a lower memory index was marginally associated with a higher proportion of active voxels of the bilateral medial subdivision within the DMN (Left: t(24)=-1.98; r=-0.44; p=0.05; AME=-0.07; 95% CI=[-0.14, -0.001]. Right: t(24)=-1.80; r=-0.34; p=0.07; AME=-0.06; 95% CI=[-0.12, -0.005]), as well as a marginal association was found with greater FC of

the left medial subdivision within the DMN (Left: t(24)= 1.89; r=-0.36; p=0.07), while a nominally significant association was found with greater FC of the right one (Right: t(24)=-2.61; r=-0.47; p=0.02; pFDR=0.1), as shown in Figure II.13, upper panel D.

When analyzing the post-encoding resting state, the proportion of significant voxels extracted from the right medial subdivision within the RFPN did not show a significant association with the memory index (t(24)=1.36; r=0.25; p=0.17; AME=0.06; 95% CI=[-0.03, 0.15]), as well as the FC from the medial subdivision (t(24)=-0.26; r=-0.05; p=0.8), as shown in Figure II.13, lower right panel D. Similarly, the proportion of active voxels within the bilateral medial subdivision of the DMN (Left: t(24)=2.18; r=0.41; p=0.03; pFDR=0.14; AME=0.11; 95% CI=[0.001, 0.21]. Right: t(24)= 0.47; r=-0.1; p=0.64; AME=-0.02; 95% CI=[-0.10, 0.06]), as well as the FC from the bilateral medial subdivision (Left: t(24)=2.1; r=0.4; p=0.05. Right: t(29)=-0.48; r=-0.09; p=0.64), only showed statistical trends in the left hemisphere in the opposite direction of the baseline patterns (Figure II.13, lower panel D). However, the GLM conducted on the change score representing the voxels' recruitment subtraction between the post- and pre-encoding resting-state scans highlighted a significant association between the change score of the left medial subdivision within the DMN and the memory index (t(24)=2.64; r=0.47; p=0.01; pFDR=0.05) in the opposite direction of the relationship observed during the baseline resting-state. All other pFDR>0.05. Similar association patterns were observed between the FC change score of the left medial subdivision within the DMN and the memory index (t(24)=2.65; r=0.47; p=0.01; pFDR=0.11) and bilaterally in the medial subdivision within the FPNs (Left: t(24)=-2.68, r=-0.48; p=0.01, pFDR=0.11; Right: t(24)=-1.99; r= 0.38; p=0.06, pFDR=0.33).

UNIBA. During baseline resting-state, a higher memory index was associated with a greater proportion of significant voxels in the left and right medial subdivision within the LFPN and RFPN, respectively (Left: t(104)=2.94; r=0.46; p=3.2e-03, pFDR=5.72e-03; AME=0.03; 95% CI=[0.02, 0.14].

Right: t(104)=4.36; r=0.59; p=1.28e-05; pFDR=5.13e-05; AME=0.1; 95% CI=[0.03, 0.15]), and the FC from the left medial subdivision within the LFPN (t(104)=3.17; r=0.29; p=2e-03; pFDR=0.02), while the FC from the right subdivision within RFPN showed only marginal significance (t(104)=1.88; r=0.17; p=0.06), as shown in Figure II.13, upper panel D. In the opposite direction, we found that a lower memory index was associated with a higher proportion of active voxels of the bilateral medial subdivision within the DMN (Left: t(104)=-2.58; r=-0.24; p=0.01; pFDR=0.03; AME=-0.05; 95% CI=[-0.08, -0.01]. Right: t(104)=-4.06; r=-0.4; p=4.92e-05; pFDR=3.44e-04; AME=-0.07; 95% CI=[-0.11, -0.04]) and the FC from the right medial subdivision (t(104)=-2.12; r=-0.21; p=0.04), while the FC from the left medial subdivision was not significant (t(104)=-1.67; r=-0.17; p=0.1). Moreover, we found that a higher proportion of significant voxels from the left posterior subdivision in the LFPN and the right posterior subdivision in the RFPN was significantly associated with a lower memory index (Left: t(104)=-2.76; r=-0.26; p=5.72e-03; pFDR=5.72e-03; AME=-0.07; 95% CI=[-0.12, -0.02]. Right: t(104)=-2.66; r=-0.25; p=7.91e-03; pFDR=0.01; AME=-0.07; 95% CI=[-0.11, -0.04]). However, the association patterns were not significant when analyzing FC from the posterior subdivision (Left: t(104)=-0.98; r=0.09; p=0.33. Right: t(104)=-0.29; r=0.03; p=0.77), as shown in Figure II.13, upper panel D.

Consistent findings across the two samples suggest potentially opposite configurations of the medial subdivision within the DMN compared to the bilateral FPN in supporting the subsequent associative encoding. The switch between the two resting-state scans associated with the memory index in the RUB sample highlights a change in spatial recruitment and FC of the medial subdivision within the DMN between the two resting-state scans, with the memory task interposed.

2.3.2.4. Individual-level thalamic recruitment across memory phases indirectly affects memory performance

Our analysis of cortico-thalamic recruitment supporting associative memory baseline during resting-state proved the medial subdivision's prominent role across samples. Consequently, we included only the proportion of voxels in the medial subdivisions during baseline resting-state as a predictor in our SEM (X). We included the right anterior BOLD estimates during encoding (M1) and the right medial BOLD estimates during retrieval (M2) as serial mediators, as they were consistently associated with the memory index across samples. We separately considered the left and right hemispheres. As we reported no significant associations in the post-encoding resting-state in RUB, we did not include the post-encoding scan in the path analyses.

RUB. The total effect of the SEM on the right hemisphere was significant (std. β [bootstrapped 95% CI] = 0.37 [0.02; 0.7]; R2 = 0.12; p=0.04; Figure II.14, right panel A). We found a significant indirect effect of the right anterior thalamic BOLD signal during encoding (M1: std. β [bootstrapped 95% CI] = -0.33 [-0.56; -0.09]; R2 = 0.2; p=0.006; Figure II.14, right panel A) negatively influencing the association between the proportion of voxels in the medial subdivision within the RFPN during the baseline resting-state and the associative memory performance. Also, we found a significant indirect effect through the right medial thalamic BOLD signal during retrieval (M2: std. β [bootstrapped 95% CI] = 0.48 [0.12; 0.8]; R2 =0.2; p=0.008; Figure II.14, right panel A). In the left hemisphere, we found a significant indirect effect through the right medial thalamic BOLD signal during retrieval (M2: std. β [bootstrapped 95% CI]=0.41 [0.04; 0.8]; R2 =0.14; p=0.02), while no other indirect effects (M1: std. β [bootstrapped 95% CI] =-0.17 [-0.41; 0.05]; R2 =0.07; p=0.1; M1-2: std. β [bootstrapped 95% CI]=-0.03 [-0.12; 0.04]; R2 =0.02; p=0.9; Figure II.14, left panel A) or the total model (std. β [bootstrapped 95% CI] = 0.10 [-0.21; 0.42]; R2 =0.01; p=0.5) were significant.

UNIBA. The total effect of the path analysis on the right hemisphere was significant (std. β [bootstrapped 95% CI] =0.16 [0.07; 0.24]; R2 = 0.12; p=0.001; Figure II.14, right panel B), in line with the analysis performed in the RUB sample. Also, we found a significant indirect effect of the right anterior thalamic BOLD signal during encoding (M1: std. β [bootstrapped 95% CI] = -0.03 [-0.06; 0.002]; R2 =0.04; p=0.04; Figure II.14, right panel B) negatively influencing the association between the proportion of voxels in the medial subdivision within the RFPN during the baseline resting-state and the associative memory performance. The indirect effect through the right medial thalamic BOLD signal during retrieval showed significance in the same direction found in the RUB results (M2: std. β [bootstrapped 95% CI] = 0.03 [0.002; 0.60]; R2 =0.04; p=0.03; Figure II.14, right panel B). The same model in the same network was significant also in the left hemisphere (std. β [bootstrapped 95% CI]=0.16 [0.07; 0.26]; R2 =0.1; p=0.001; Figure II.14, left panel B) considering the indirect effects of left anterior BOLD signal during encoding (M1: std. β [bootstrapped 95% CI] = -0.039 [-0.08; -2e-04]; R2 =0.04; p=0.05; Figure II.14, left panel B) and the left medial BOLD signal during retrieval (M2: std. β [bootstrapped 95% CI]=0.046 [0.0002; 0.09]; R2 =0.04; p=0.05; Figure II.14, left panel B).

Overall, results show that the BOLD signal from the anterior subdivision during encoding and the BOLD signal from the medial subdivision during retrieval have independent indirect effects on the association between the recruitment of the thalamic subdivision during baseline resting-state and individual memory performance.

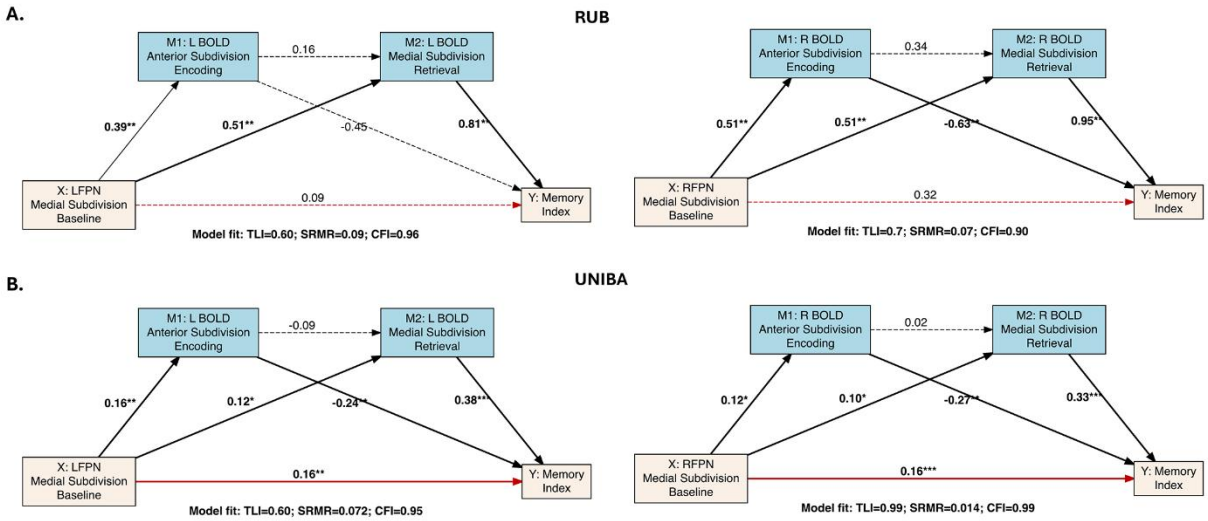

Figure II. 14 Indirect effects of BOLD task activity between baseline resting state and the memory performance. Diagrams represent the SEM on (**A**) the RUB and (**B**) UNIBA samples testing the relationship between the proportion of positive active voxels extracted from the right medial thalamic subdivision during baseline resting state and the memory index as mediated by the BOLD signal extracted from the anterior thalamic subdivision during the encoding and the BOLD signal extracted from the medial thalamic subdivision during retrieval. Standardized β coefficients for each path are displayed. Significant direct paths are depicted with bold red lines, while significant indirect paths are reported with bold black lines. Continuous lines represent significant paths, while dashed lines represent a nonsignificant effect. Significant was set at $\alpha < 0.05$ and is marked with asterisks. The goodness of fit is reported for each model as the standardized root mean square residual (SRMR), the Tucker–Lewis index (TLI), and the comparative fit index (CFI). As the rule of thumb guidelines is that CFI ≥ 0.95, TLI ≥ 0.95, and RMSEA ≤ 0.05 represent an excellent fitting model, the right hemisphere model in the UNIBA sample shows the best model fit. R, right; L, left; RFPN, right frontoparietal network; LFPN, left frontoparietal network.

2.3.2.5. K-means clustering on the individual-level thalamic recruitment during the associative memory task differentiates individuals' pathways supporting memory performance

To clarify whether individuals differentially recruit distinct cortico-thalamic pathways, we performed a k means clustering of observations in the UNIBA datasets. Clustering performances are depicted in Figure II.15A-B, and an optimal number of clusters k=3 has been identified using both the left and the right hemisphere data (Left: Cluster 1, N=27; Cluster 2, N=44, Cluster 3, N=30. Right: Cluster 1, N=34; Cluster 2, N=39, Cluster 3, N=28. Figure II.15C). The overall clustering stability reached S=0.93 for the right hemisphere and S=0.89 for the right hemisphere.

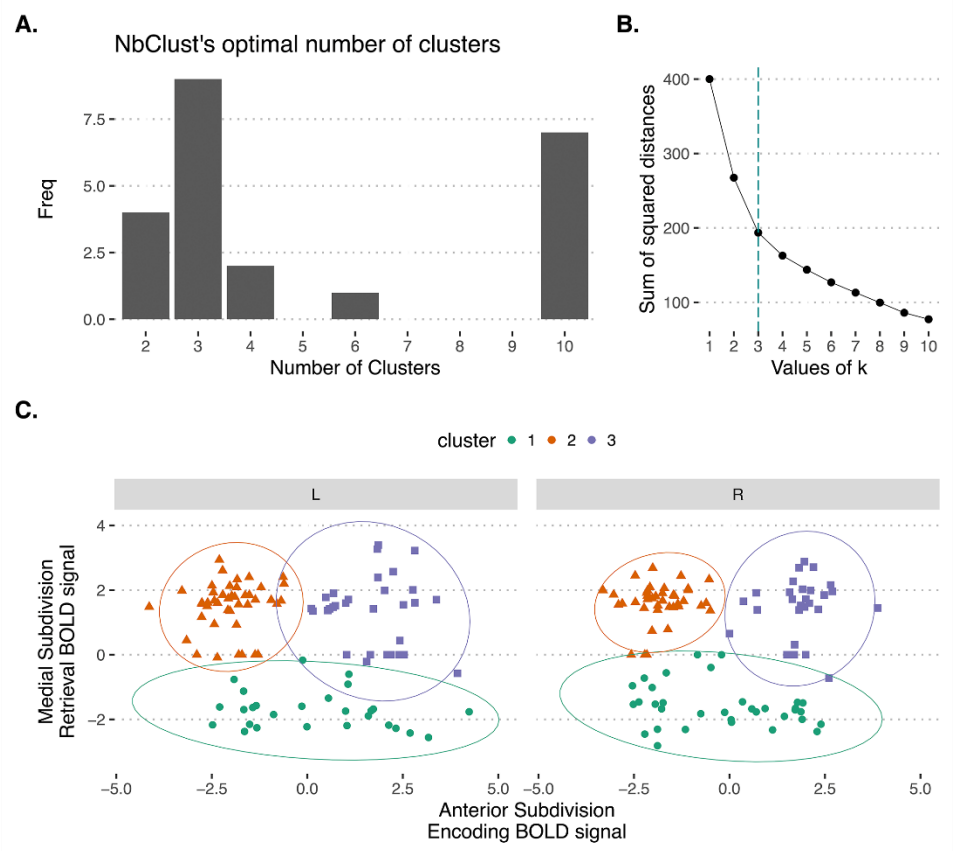

Figure II. 15 k-mean clustering performance in the UNIBA sample. A. Barplot shows the optimal number of clusters determined through the NbClust method (k = 3). **B.** Sum of square distances for determining the best number of clusters through the elbow method (k = 3). **C.** Clusters obtained for the left and right hemispheres are represented on the distribution of the BOLD signal extracted from the anterior subdivision during the encoding (x-axis) and the medial subdivision during the retrieval (y-axis). Cluster 1 is depicted in green, Cluster 2 in orange, and Cluster 3 in violet. R, right; L, left; BOLD, blood oxygenation level-dependent.

In the right hemisphere, Cluster 1 was characterized by i) the lowest performance (Cluster 1 vs. Cluster 2: Z=-3.10, r=0.36, p=2e-03, pFDR=5e-03; Cluster 1 vs. Cluster 3: Z=-1.33, r=0.17, p=0.18, pFDR=0.2. Figure II.16A), ii) the lowest proportion of significant active voxels in medial subdivision during baseline resting-state (Cluster 1 vs. Cluster 2: Z=-2.75, r=0.6, p=6e-03, pFDR=0.01; Cluster 1 vs. Cluster 3: Z=-3.09, r=0.66, p=2e-03, pFDR=5e-03. Figure II.16B), iii) a higher BOLD signal from the anterior subdivision during the encoding compared to Cluster 2 (Z=4.02, r=0.47, p= 5.74e-05, pFDR= 5.74e-05), but lower compared to Cluster 3 (Z=-5.03, r=0.53, p= 5.07e-07, pFDR=1.01e-06), and iv) lowest BOLD signal from the medial subdivision during the retrieval

(Cluster 1 vs. Cluster 2: Z=-7.29, r=0.85, p=3e-13, pFDR = 9e-13; Cluster 1 vs. Cluster 3: Z=-6.59, r=0.83, p= 4.51e-11, pFDR = 9.02e-11. Figure II.16C). This pattern suggests that lower recruitment of the medial subdivision before the associative encoding may be detrimental to memory performance regardless of the activation level of the anterior and medial subdivisions during the task performance.

Cluster 2 represents the best performers (Cluster 2 vs. 1: Z=3.10, r=0.36, p=2e-03, pFDR=5e-03; Cluster 2 vs. 3: Z=-1.18, r=0.22, p=0.07, pFDR=0.13) showing i) an intermediate level of medial thalamic recruitment during baseline resting-state, yet not significantly different from Cluster 3 (Z=-1.94, r=0.42, p=0.06), ii) a pattern of lower activation of the anterior subdivision during encoding and iii) greater activation of the medial subdivision during the retrieval. This pattern suggests that greater recruitment of the medial subdivision, compared to Cluster 1, prior to associative encoding, along with deactivation of the anterior subdivision during encoding and increased activation of the medial subdivision during associative retrieval, characterizes the best performers.

Cluster 3 represents individuals with a nominally intermediate performance, even though no significant differences were reported with the performance of individuals in Cluster 2 (Z=-1.82, r=0.22, p=0.07, pFDR=0.13) or Cluster 1 (Z=-1.33, r=0.17, p=0.18, pFDR=0.2). In contrast, individuals in Cluster 3 showed the highest recruitment of the medial subdivision during baseline resting-state, yet not different from Cluster 2 (Z=1.94, r=0.42, p=0.06, pFDR=0.06), and a pattern of enhanced activation of both the anterior subdivision during encoding and the medial subdivision during the retrieval. We found a similar pattern in the left hemisphere, as shown in Figure II.16. These patterns suggest that the highest recruitment of the medial subdivision during the baseline supports an optimal medial subdivision activation during the task performance. This applies even without the suppression of the anterior subdivision during the encoding to promote memory performance.

The Sørensen–Dice index on cluster membership across hemispheres showed similarity at 0.63 for Cluster 1, 0.72 for Cluster 2, and 0.51 for Cluster 3, suggesting a substantial yet partial

consistency across hemispheres. Clusters showed no differences (pFDR>0.05) in age, sex, education, and intellective quotient, suggesting no nuisance effects of these variables on the clustering results.

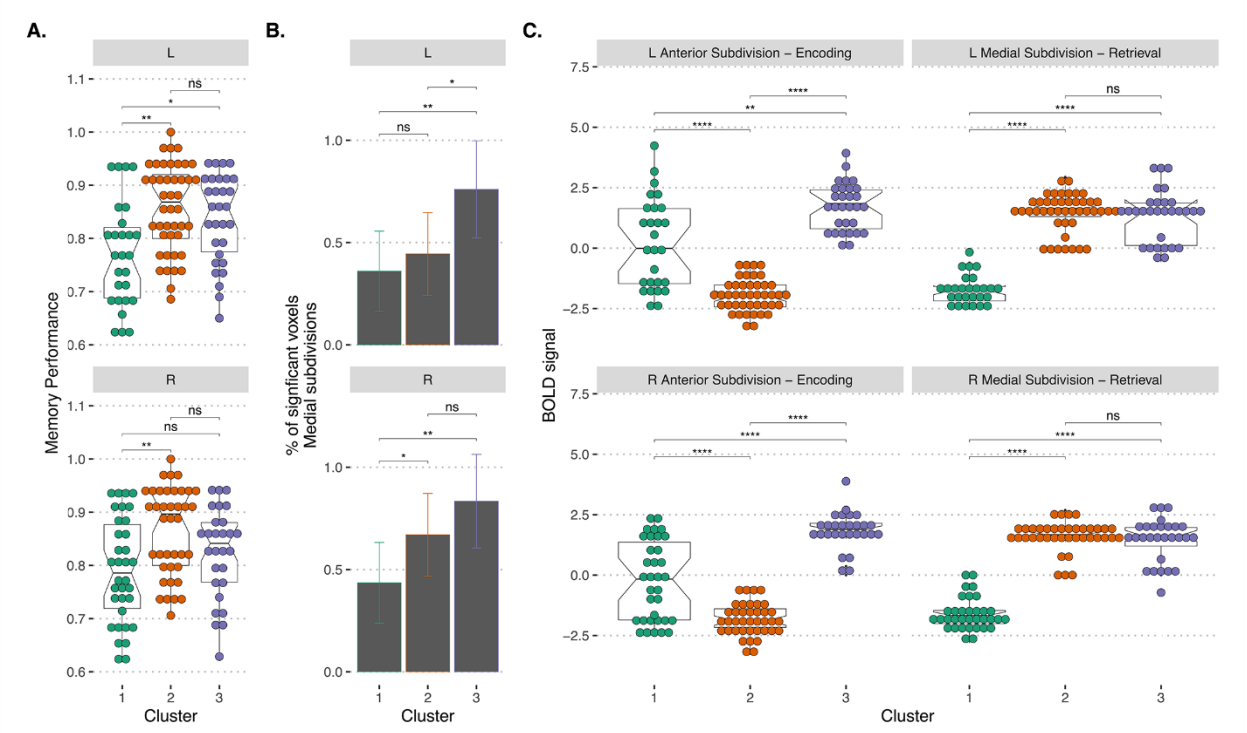

Figure II. 16 Clusters of individuals in the UNIBA dataset use different thalamocortical pathways. A. Boxplots show the differences in associative memory performance across the clusters of individuals obtained through the k-mean clustering. For the left hemisphere, Cluster 1 shows the lowest performance indices compared to Cluster 2 (Z=-3.53, r=0.42, p=4.12e-04, p_{FDR}=0.001) and Cluster 3 (Z=-2.61, r=0.35, p=0.009, p_{FDR}=0.017). Clusters 2 and 3 show no differences (Z=1.01, r=0.10, p=0.3). The same pattern is preserved in the right hemisphere (Cluster 1 vs. 2: Z=-3.10, r=0.36, p=0.002, pFDR=0.005; Cluster 1 vs. 3: Z=-1.33, r=0.17, p=0.18; Cluster 2 vs. 3: Z=-1.18, r=0.22, p=0.07). **B.** Barplots show the differences in the percentage of significant voxels within the RFPN and the LFPN during baseline resting-state across the clusters of individuals identified through the k-mean clustering. For the left hemisphere, Cluster 1 shows lower medial subdivision recruitment compared to Cluster 2 (Z=-1, r=0.16, p=0.5) and Cluster 3 (Z=-3.31, r=0.69, p=9.3e-04, p_{FDR}=0.003). Also, Cluster 3 shows higher medial subdivision recruitment compared to Cluster 2 (Z=2.29, r=0.55, p=0.22, p_{FDR}=0.045). The same pattern is preserved in the right hemisphere (Cluster 1 vs. 2: Z=-2.75, r=0.66, p=0.006, pFDR=0.013; Cluster 1 vs. 3: Z=-3.09, r=0.65, p=0.002, pFDR=0.005), but Cluster 2 and 3 comparison shows only a statistical trend (Z=-1.94, r=0.42, p=0.05, p_{FDR}=0.06). **C.** Boxplots show the differences in BOLD signal extracted from the anterior subdivision during the encoding and the medial subdivision during the retrieval across the clusters of individuals obtained through the k-mean clustering. For the left hemisphere, Cluster 2 shows the lowest BOLD signal from the anterior subdivision compared to Cluster 1 (Z=-5.43, r=0.64, p=5.76e-08, p_{FDR}=1.15e-07) and Cluster 3 (Z=-9.43, r=0.89, p=4.26e-21, p_{FDR}=1.28e-20). Cluster 1 shows a lower BOLD signal from the anterior subdivision than Cluster 3 (Z=-3.29, r=0.43, p=1e-03, p_{FDR}=1e-03). Also, Cluster 1 shows the lowest BOLD signal from the medial subdivision compared to Cluster 2 (Z=-7.03, r=0.83, p=2.04e-12, p_{FDR}=6.12e-12) and 3 (Z=-6.44, r=0.85, p=1.22e-10, p_{FDR}=2.44e-10). Clusters 2 and 3 show no significant differences in the BOLD signal extracted from the medial subdivision (Z=1.01, r=0.1, p=0.4). The same pattern is preserved in the right hemisphere concerning the BOLD signal extracted

from the anterior subdivision during the encoding (Cluster 1 vs 2: Z=4.02, r=0.47, p=5.74e-05, p_{FDR}=5.74e-05; Cluster 1 vs 3: Z=-5.02, r=0.64, p=5.07e-07, p_{FDR}=1.01e-06; Cluster 2 vs. 3: Z=-8.95, r=0.8, p=3.41e-19, p_{FDR}=1.02e-18) and from the medial subdivision during the retrieval session (Cluster 1 vs. 2: Z=-7.29, r=0.85, p=3e-13, p_{FDR}=9e-13; Cluster 1 vs. 3: Z=-6.59, r=0.84, p=4.51e-11, p_{FDR}=9.02e-11; Cluster 2 vs. 3: Z=0.7, r=0.08, p=0.5). Abbreviations: R=right; L=left; BOLD = Blood Oxygenation Level Dependent.

2.4. Discussion

Building on prior evidence linking individual memory performance to medial thalamic connectivity with frontoparietal cortices prior to learning, our investigation delves into the role of cortico-thalamic recruitment in supporting memory across two independent samples. Briefly, we found that the neural activity in the anterior and medial thalamic subdivisions is associated with memory encoding and retrieval during task performance. The engagement of these thalamic subdivisions with cortical networks differs not only across scans but also among individuals, correlating with memory performance. The major strength of this study lies in the implementation of two independent samples, overcoming the main limitation—namely, the lack of a replication dataset—present in the previous study by Passiatore et al. (2021).

2.4.1. Interindividual variability in thalamic activity during memory task

Our group-level analysis of task activity showed consistent activations of the MTL and the frontoparietal cortices across samples, representing well-established components of the episodic memory network (Cansino et al., 2002; Muzzio et al., 2009; Aminoff et al., 2013; Opitz, 2014). Significant task-dependent thalamic activity extended across all thalamic subdivisions, with the predominant engagement of the posterior and medial subdivisions during associative encoding. Greater involvement of the anterior subdivision was descriptively observed in the retrieval contrasts compared to encoding, suggesting possible overlapping processes across samples (Mayes et al., 2007). However, our experimental framework does not allow us to determine that the brain performs identical operations under two related yet different memory paradigms. Indeed, dissimilarities in anterior thalamic spatial recruitment were reported across samples, likely reflecting the different cognitive demands during retrieval. Past research indicates the critical role of anterior subdivisions in

processes associated with recognition and recall (Pergola et al., 2013a; Sweeney Reed et al., 2014; Carlesimo et al., 2015), while the medial subdivision is implicated in executive functioning and attentional processing supporting memory encoding and retrieval (Pergola et al., 2012).

It is worth noting that the observed activities may encompass perceptual effects not necessarily related to memory processes (Saalmann and Kastner, 2011; Antonucci et al., 2021), implicating the posterior and ventral subdivisions, which are comparatively less involved in memory (Van Der Werf et al., 2003; Hwang et al., 2017). Two key considerations arise regarding the role of these small regions compared to well-studied areas like the cortex, where signals are distributed over larger areas (Hermes et al., 2012; Raimondo et al., 2021): first, the fMRI spatial resolution may lead to mixed signals when detecting effects in small regions (Pergola et al., 2013b; Phillips et al., 2021), as shown by the correlations across subdivision FC; second, anatomical studies have shown considerable variability in thalamic structures across individuals (Johansen-Berg et al., 2005; Fama and Sullivan, 2015; Georgescu et al., 2020). Therefore, spatially averaged group analyses may overlook individual variations in subcortical signals. We thus prioritized individual signals from segmented thalamic subdivisions in native space, thereby reducing registration errors and providing a more precise assessment of the contributions of smaller regions to behavior (Seibert et al., 2012; Meyer et al., 2017). Our individual-level findings consistently show deactivation of the anterior thalamic subdivision during associative memory encoding, followed by activation of the medial subdivision during retrieval in both samples, despite differences in retrieval task demands. This pattern replicates previous evidence linking the deactivation of the anterior subdivision to subsequent memory performance (Wagner et al., 2019; Geier et al., 2020). Deactivating the anterior thalamus may facilitate the encoding of new memories without concurrent retrieval interference. This hypothesis may also explain palinopsia in patients with acute anterior thalamic ischemia, in which mnestic traces overlap (Ghika-Schmid and Bogousslavsky, 2000; Cipolotti et al., 2008; Carlesimo et al., 2011). A possible mechanism involves the hippocampal

reactivation of prior traces through the anterior thalamus, whose suppression is thus essential for memory encoding (Aggleton and O'Mara, 2022). Pergola and Suchan (2013) proposed that the anterior nuclei may be required for instantiating the cortical synaptic plasticity patterns necessary for transferring memory traces from the hippocampus to the neocortex. Therefore, optimal encoding may involve the activation of medial subdivision alongside anterior suppression during tasks.

The discrepancy between our group-level and individual-level results underscores the value of our methodological approach. It enables us to relate individual-level neurobiology underlying memory processes in such small regions while addressing behavioral variability.

2.4.2. Task comparability and the nature of associative retrieval

One possible concern in interpreting our findings is the comparison of results from two different memory paradigms. While associative encoding is common in both tasks, their retrieval phases differ in structure and cognitive demands. During fMRI retrieval, both RUB and UNIBA tasks involve recognition, which entails single-item recognition (without associative retrieval demand) and picture-pair recognition (with cue-based associative retrieval demand), respectively. However, the RUB task includes a post-scanning forced-choice associative test, allowing us to retrospectively assess associative memory performance. To address this asymmetry, our approach was to 'enrich' the interpretation of retrieval-related neural activity by considering the broader context of associative encoding and post-task recall, thereby capturing associative contributions indirectly.

Nonetheless, the assumption that the UNIBA retrieval task exclusively taps into associative memory is not without debate. As Mayes et al. (2007) and others have argued, recognition of composite stimuli—such as picture pairs—can be supported by familiarity when the items are perceived as a single, unitized representation rather than as two distinct elements requiring relational binding (Yonelinas, 1999; Quamme et al., 2007; Elfman et al., 2008). This unitization process allows

familiarity, typically considered a non-associative mechanism, to support performance in tasks that appear associative on the surface. Therefore, both tasks may involve a mixture of associative and non-associative processes, depending on how participants encode and retrieve the stimuli. This similarity in underlying processes may explain the consistent brain activation patterns observed across samples during retrieval, particularly in the medial subdivision, which is implicated in both executive control and familiarity-based recognition (Mitchell & Chakraborty, 2013; Pergola et al., 2013a). Moreover, our behavioral results support the idea that both tasks may involve overlapping cognitive mechanisms. In the RUB sample, participants were faster in recognizing previously seen items than rejecting new ones, suggesting a strong familiarity component. In contrast, the UNIBA sample showed higher accuracy for intact pairs than for rearranged ones, consistent with associative retrieval demands, but without significant differences in reaction times. These patterns suggest that, although the tasks differ in format, they may converge in the underlying cognitive operations and neural substrates they engage, particularly in the thalamocortical circuits that support memory retrieval.

Ultimately, although we acknowledge that we cannot claim full equivalence in latent cognitive processes across tasks, the convergence of behavioral and neural findings across samples strengthens the validity of our conclusions. Future studies using within-subject designs and parametric manipulations of associative strength could help disentangle the contributions of familiarity and recollection more precisely.

2.4.3. Functional activation and deactivation of the anterior and medial thalamic subdivisions

The observed deactivation of the anterior thalamic subdivision during memory encoding may reflect a gating mechanism that facilitates the suppression of interfering mnemonic traces. This interpretation aligns with prior evidence suggesting that the anterior thalamus, through its connections with the

hippocampus and retrosplenial cortex, plays a pivotal role in recall-based recognition (Aggleton et al., 2011; Pergola et al., 2013a). Its transient deactivation during encoding could prevent the reactivation of previously stored information, reducing proactive interference and enhancing the encoding of novel associations (Aggleton & O'Mara, 2022; Geier et al., 2020). In neuroimaging, the terms "activation" and "deactivation" typically refer to positive and negative BOLD responses, respectively. These signal changes arise from neurovascular coupling—the relationship between neural activity and hemodynamic parameters such as cerebral blood flow (CBF), cerebral blood volume (CBV), and the cerebral metabolic rate of oxygen consumption ($CMRO_2$). While the positive BOLD response (PBR) is generally interpreted as a marker of increased neural activity, the physiological basis of the negative BOLD response (NBR) remains a topic of ongoing debate. The PBR is characterized by a well-defined temporal profile, including an initial dip (0–3 s), a main peak (3–11 s), and a post-stimulus undershoot (11–17 s). In contrast, the NBR is less consistently shaped but is typically defined by a sustained signal decrease below baseline during the 3–11 s window. Varying and often conflicting reports on its vascular and metabolic contributors have hindered the identification of its physiological basis. As a result, discussions of the NBR have considered a variety of possible mechanisms, which can be broadly categorized into vascular and neuronal origins (Devor et al., 2008; Nagaoka et al., 2006; Schridde et al., 2008; Shih et al., 2009). The vascular hypothesis, referred to as the *hemodynamic steal hypothesis* or the *blood-steal hypothesis,* is based on the blood-steal mechanism. According to this effect, NBR in one area arises because increased CBF in adjacent activated regions leads to reduced perfusion in nearby areas (Harel et al., 2002; Hu and Huang, 2015; Kannurpatti and Biswal, 2004; Ma et al., 2017; Puckett et al., 2014), and venous backpressure effects (Bandettini, 2012; Goense et al., 2012; Huber et al., 2014; Shmuel et al., 2006). Neuronal explanations include both decreased excitatory activity and active inhibition (Boorman et al., 2010; Huber et al., 2014; Logothetis, 2002; Mullinger et al., 2014; Nakata et al., 2019; Shmuel et al., 2006, 2002; Wilson et al., 2019). In the *neural inhibition hypothesis* (Shmuel et

al., 2002; Buzsaki et al., 2007; Devor et al., 2008), NBR are caused by local neural inhibition, for example, deactivation of the DMN in conjunction with activation in distributed task-positive networks (Raichle et al., 2001; Gusnard et al., 2001) or in the ipsilateral somatosensory cortex (Schafer et al., 2012; Klingner et al., 2010). This hypothesis is supported by observations of decreased local field potential activity in primates (Shmuel et al., 2006) and rats (Boorman et al., 2010, 2015; Lee et al., 2019). Also, in humans, correlations between NBR and 8–13 Hz alpha-frequency EEG power have been observed (Mullinger et al. 2014; Wilson et al. 2019) as well as decreases in oxygen metabolism in human motor and visual cortex NBR regions (Mullinger et al. 2014; Pasley et al. 2007; Shmuel et al. 2002; Stefanovic et al. 2004). Moreover, NBRs have been linked to decreased $CMRO_2$ in visual and motor cortices (Stefanovic et al., 2004; Pasley et al., 2007), and to GABAergic inhibition (Koush et al., 2021). These findings suggest that NBRs may reflect active cortical suppression rather than a passive hemodynamic artifact. However, whilst such studies suggest that the NBR reflects, at least in part, a measure of cortical deactivation (Sten et al. 2017), caution should be taken in assuming the NBR represents the simple neurophysiological inverse of the PBR (Devi et al. 2022; Huber et al. 2014; Mayhew et al. 2014; Mullinger et al. 2014; de la Rosa et al. 2021).

In our study, in the context of the anterior thalamus, we found a consistent association between negative BOLD during encoding and better memory performance across samples, supporting the hypothesis of a functional suppression mechanism, possibly aimed at minimizing interference from previously encoded traces (Wagner et al., 2019; Pergola et al., 2018). Conversely, the medial thalamic subdivision, particularly the MD nucleus, showed increased activation during retrieval, supporting its involvement in executive control processes such as attentional allocation and goal-directed memory search (Mitchell & Chakraborty, 2013; Antonucci et al., 2020). Finally, while most NBR research has focused on cortical regions, subcortical structures, such as the thalamus, present additional challenges due to their smaller size and distinct vascular architecture (Mishra et al., 2011).

Thus, these findings provide additional evidence on the meaning of the NBRs, particularly in favor of the neural inhibition hypothesis, suggesting a functional dissociation between thalamic subdivisions, with the anterior nucleus acting as a gatekeeper during encoding and the medial nucleus facilitating controlled retrieval processes.

2.4.4. The role of the medial thalamic subdivision within cortico-thalamic interplay across learning stages

Recent studies have shown that multi-scan fMRI effectively captures individual differences in neural recruitment during memory and learning tasks (Tambini et al., 2010; Brodt et al., 2018; Wagner et al., 2019; Passiatore et al., 2021). Our previous research found changes in thalamic FC with the FPN as a function of learning, comparing baseline to post-encoding resting-state (Passiatore et al., 2021), suggesting a consistent association between thalamic FC and performance across participants.

Upon closer examination of the same dataset reported by Passiatore et al. (2021) and the larger UNIBA sample, we observed a distinct outcome. Our results highlight the prominent role of the medial subdivisions in cortico-thalamic recruitment within the bilateral FPNs during baseline resting state, predicting memory performance. In contrast, the best performers exhibited reduced recruitment of the medial subdivision in the bilateral DMN during baseline. The involvement of medial subdivisions prior to task performance may facilitate preparatory operations across cortical circuits instrumental to encoding (Sweeney-Reed et al., 2015; Perakyla et al., 2017), possibly influencing arousal and attentiveness during task performance (Saalmann et al., 2012). Indeed, the medial thalamic subdivision likely interacts with cortical activity during high-load cognitive processes, which subsequently contribute to supporting memory (Pergola et al., 2018). Hence, in the context of memory, the function of medial subdivisions may extend beyond information retrieval to facilitate the processing of information that will be later transmitted to the corticothalamic circuit for subsequent recall (Antonucci et al., 2020).

The results obtained from the structural equation modeling framework analysis bolster the hypothesis regarding the broader role of the medial thalamic subdivision in distributed processing for executive control (Antonucci et al., 2020). The independent indirect pathways involving the anterior and mediodorsal nuclei account for different proportions of variance in performance, indicating that mediodorsal thalamus recruitment during retrieval is not reliant on anterior nuclei activity. This functional dissociation aligns with lesion studies showing differential cognitive deficits in patients with focal tuberothalamic versus paramedian thalamic damage (Pergola & Suchan, 2013), where anterior lesions impair recall and spatial memory, while mediodorsal lesions affect executive functioning and strategic retrieval.

Interestingly, our data show an absence of direct correlations between anterior/encoding and medial/retrieval pathways, despite both contributing indirectly to memory performance. This pattern proposes an individual stratification, in which the individuals recruiting anterior/encoding pathways are different from those showing medial/retrieval mediation. This perspective is supported by clustering analyses across individuals based on thalamic-frontoparietal configurations during task performance. Specifically, the optimal memory profile involves medial thalamic recruitment during baseline resting state, anterior thalamic deactivation during encoding, and medial thalamic activation during retrieval. However, even in the absence of anterior/encoding and medial/retrieval recruitment, recruiting the medial subdivision during the baseline resting state predicts better memory performance, underscoring its preparatory role in cognitive readiness and attentional set (Sweeney-Reed et al., 2015; Hwang et al., 2022).

2.4.5. Variability in Lateralization of Thalamic Contributions to Memory

Clustering analysis also indicated different clustering membership across hemispheres, suggesting variability in lateralization related to memory processes. This means that participants did not consistently recruit the same thalamic subdivisions across hemispheres. This variability in recruitment patterns across individuals and hemispheres suggests a nuanced lateralization of thalamic contributions to memory, possibly indicating that the lateralization of thalamic function is not fixed but rather reflects dynamic, context-dependent configurations of cortico-thalamic circuits.

Hemispheric asymmetries in thalamic connectivity have been observed in both healthy populations and clinical groups, with implications for the specialization of verbal versus spatial memory (Zhou et al., 2020; Bocchetta et al., 2021). While classical models of hemispheric specialization suggest a left-hemisphere dominance for verbal memory and a right-hemisphere bias for visuospatial memory (Kelley et al., 1998; Golby et al., 2001), recent neuroimaging studies have shown that this dichotomy is not absolute and may vary significantly across individuals depending on task demands, cognitive strategies, and intrinsic network organization (Seghier & Price, 2018; Kong et al., 2019).

This variability may be influenced by individual differences in structural connectivity between the thalamus and cortical association areas. For example, diffusion tensor imaging (DTI) studies have shown asymmetries in thalamocortical tracts, particularly those connecting the mediodorsal thalamus to the PFC, which may underlie lateralized patterns of executive control and memory retrieval (Johansen-Berg et al., 2005; O'Muircheartaigh et al., 2011). Moreover, functional asymmetries in thalamic activation have been observed in clinical populations, such as in patients with temporal lobe epilepsy or SCZ, where disrupted lateralization patterns are associated with cognitive impairments and altered network dynamics (Zhou et al., 2020; Hwang et al., 2022).

Importantly, the lateralization of thalamic contributions to memory may also interact with age and neurodevelopmental stages. Despite finding no effect of age on memory performance in both samples, data from individuals aged 18 to 61 years were included in the clustering, indicating a wide range of variability. In younger individuals, hemispheric specialization may be less pronounced, allowing for more bilateral recruitment of thalamic and cortical regions during memory tasks. In contrast, aging and neurodegenerative processes may exacerbate hemispheric asymmetries or lead to compensatory recruitment of contralateral regions (Cabeza et al., 2002). These findings underscore the need for individualized approaches in both research and clinical assessment, as group-level analyses may obscure meaningful lateralized patterns that are critical for understanding memory function and dysfunction. Further investigations in larger age groups are warranted to study neurodevelopmental or aging effects associated with memory.

2.4.6. Abnormal memory functioning in clinical conditions

The present findings, which highlight interindividual variability in thalamic recruitment patterns, suggest that distinct neural strategies may underlie similar behavioral outcomes. This is relevant for a clinical translation, as it is well-known that the thalamus functioning is altered in many disorders, including Alzheimer's disease (Braak & Braak, 1991), SCZ (Popken et al., 2000), Parkinson's disease (Henderson et al., 2000), autism spectrum disorder (Woodward et al., 2017), as well as frontotemporal dementia (Bocchetta et al., 2021) and major depressive disorder (Zhu et al., 2021), in which thalamic atrophy and hypoconnectivity are often associated with impairments in episodic memory, attention, and executive function.

Specifically, damage to the MD is known to produce symptoms similar to those observed after damage to the PFC, including diminished behavioral flexibility and working memory impairment (Daum & Ackermann, 1994; Van der Werf et al., 2000, 2003). These deficits appear to be more closely

linked to damage of the lateral MD nucleus, which establishes strong and reciprocal connectivity with the dorsolateral PFC, while damage to the medial portion produces deficits in long-term memory (Halassa & Kastner, 2017; Van der Werf, 2000). These findings support the important role of the medial subdivision in high cognitive processes, as it can fine-tune cortical activity to face ongoing behavioral challenges (Schmitt et al., 2017; Rikhye et al., 2018; Halassa and Kastner, 2017). This explains why MD dysfunctions can produce several mental conditions.

In this study, the medial thalamus—through its engagement with FPN networks—emerges as a key node in preparatory and executive processes that support memory encoding and retrieval. Its functional integrity may serve as a predictive marker for cognitive decline in aging and neurodegenerative conditions (Fama & Sullivan, 2015; Hwang et al., 2017), as well as in psychiatric disorders such as SCZ, where episodic memory deficits are a core cognitive symptom (Ragland et al., 2015; Antonucci et al., 2020). Multiple measures have shown that MD-PFC structural (Girarldo-Chica et al., 2018) and functional connectivity (Woodward & Heckers, 2016; Pergola et al., 2015) are diminished in SCZ. These abnormalities are believed to disrupt thalamo-prefrontal circuits, which are critical for working memory, episodic encoding, and retrieval. Additional evidence supporting this link comes from recent deep-brain stimulation results from treatment-resistant SCZ patients, which indicate that targeting inhibitory inputs to the MD significantly reduces their hallucinations and delusions (Nucifora et al., 2019). Converging evidence from high-resolution fMRI and diffusion imaging supports the notion that thalamic dysconnectivity represents a robust biomarker of cognitive impairment in SCZ (Anticevic et al., 2014; Hwang et al., 2022) and may even precede the clinical onset of psychosis in high-risk individuals (Sarpal et al., 2020). Furthermore, the researchers found that thalamic connectivity in areas of the FPN correlated with cognitive function, including verbal learning and memory (Woodward & Heckers, 2016). Using PET to explore cognitive abnormalities in SCZ, Andreasen et al. (1996) demonstrated the existence of a prefrontal-thalamic-cerebellar network that

was dysfunctional in patients with SCZ compared with healthy participants in recalling complex narrative material (Andreasen et al., 1996).

The use of spatially informed metrics—such as the proportion of significant voxels within thalamic subdivisions—may offer more sensitive biomarkers than traditional FC measures, particularly in capturing subtle individual differences relevant for clinical prognosis and treatment response (Seghier & Price, 2018; Keller et al., 2023). These metrics may also be useful in tracking longitudinal changes in thalamic function in response to pharmacological or behavioral interventions. Ultimately, understanding the phase-specific contributions of thalamic subdivisions to memory processes, as demonstrated in this study, may inform the development of targeted diagnostic tools and therapeutic strategies across a spectrum of brain disorders.

2.4.7. Limitations

In addition to the limitations already mentioned, our study lacks a control group performing an fMRI task unrelated to memory, which would test the specificity of our findings. We cannot rule out that differences between the two resting-state scans in the RUB experiment are attributable to test-retest reliability or the time elapsed, rather than to the memory task itself. Given the differences in retrieval task demands across the two datasets, our results should not be regarded as an exact replication. Nonetheless, the similarity in brain patterns observed during retrieval suggests that both single-items and picture-pairs involve thalamo-cortical contributions tapping into associative memory as captured by fMRI (Vatansever et al., 2021; Ambrus, 2024), despite the varying cognitive demands of the tasks. Moreover, although our results were consistent across two fMRI datasets acquired at different resolutions, recent advancements in enhancing signal quality in subcortical regions, particularly the thalamic nuclei (Alkemade et al., 2022), warrant cautious interpretation of these findings. Additionally, we recognize the limitation of the sample size in the clustering analyses. Therefore, we encourage you

to view the clustering results from the UNIBA dataset as an important step toward replication with larger groups. Finally, despite finding no age effects on memory performance in our sample, further investigations across larger age groups are warranted to examine neurodevelopmental or aging effects on memory.

2.4.8. Conclusions

In this study, we show that different individuals may recruit distinct cortico-thalamic resources, which, in turn, are associated with learning success. A prior brain configuration involving the recruitment of the medial subdivision within the FPN may be advantageous for subsequent recall of encoded memories, with the deactivation of the anterior subdivision during encoding and activation of the medial subdivision during retrieval. The strength of this study is its individualized approach, which emphasizes individual differences rather than group-level average analysis that identifies similarities among participants, thereby highlighting the uniqueness of each person's learning processes and related brain organization. Understanding the distinct roles of thalamic nuclei in episodic memory holds significance for clinical conditions affected by learning and memory impairments. Enhancing our understanding of thalamic functions and their link to mental disorders is critically important from a basic neuroscience viewpoint, and it also has the potential to guide clinical applications, including deep brain stimulation and neurofeedback.

Chapter 3

Study 2: fMRI Reward Response Classifies Human Participants into Sign- and Goal-Trackers and Is Associated with Response to Antipsychotics

This chapter is based on the manuscript titled **"Reproducible Human Reward Imaging Phenotypes Exhibit Differential Sensitivity to Dopamine D2 Receptor Antagonism"**, which has been submitted for publication.

Introduction

In environments, individuals are constantly exposed to a multitude of stimuli and cues that can influence their behavior. A rewarding stimulus is an object, event, activity, or situation that has the potential to motivate an individual to approach and engage with it (Schultz, 2015). Rewards possess appetitive and motivational properties, which can increase the likelihood of the behaviors that immediately precede them; this is known as "positive reinforcement". In contrast, "aversive" stimuli suppress the occurrence of preceding behaviors and are referred to as "negative reinforcers" (Ernst & Spear, 2009). From an evolutionary perspective, rewards are crucial for survival, as individuals must simultaneously avoid threats and seek out positive stimuli to ensure their well-being.

Animal studies have identified several psychological components of reward: "liking" (core reactions to hedonic impact), "wanting" (the motivation process of incentive salience), and "learning" (Pavlovian or instrumental associations and cognitive representations) (Berridge & Robinson, 2003). Evolutionarily, "wanting" probably developed first to motivate the search for food and mates, while

"liking" evolved later to improve this pursuit. For optimal functioning, "liking" and "wanting" should cohere, encouraging individuals to pursue actions that lead to positive outcomes. However, they are distinct processes, with discriminable neural mechanisms (Berridge & Robinson, 2016; Berridge and Dayan, 2021). Specifically, the mesolimbic dopaminergic systems mediate "wanting", but not "liking," for the same reward (Berridge et al., 1989; Robinson & Berridge, 1993).

Extensive studies in rodents discovered behavioral variability across mice in reward learning, distinguishing ST, who attribute incentive salience to reward-predictive cues, from goal-trackers GT, who focus on outcomes (Flagel & Robinson, 2009; Flagel et al., 2010; Gillis & Morrison, 2019, for a review see Felix & Flagel, 2024). Human studies replicating this model are limited, often underpowered, and focused on learning paradigms, with sparse evidence on the underlying neural processes (Colaizzi et al., 2023; Schad et al., 2024; Joyner et al., 2017).

In animal studies, researchers often examine reward-cue responses in a state of deprivation using primary rewards, such as food and water (Lovic et al., 2011; Tomie et al., 2008; Meyer et al., 2012). However, when studying humans, this state cannot be replicated for several reasons, including ethical concerns. Therefore, secondary rewards (i.e., monetary) are used as substitutes for the reward system in humans, as various studies have shown that primary and secondary reinforcements activate the brain in a similar manner (Sescousse et al., 2013). In these experimental paradigms, participants are given straightforward instructions to complete tasks, eliminating the need for prior learning (Knutson et al., 2000; Sescousse et al., 2013; Sambuco et al., 2020). The bulk of neuroimaging on rewards has been carried out with group-level derived BOLD contrasts using the MID (Knutson et al., 2000, 2001; Chen et al., 2022; Oldham et al., 2018; Wilson et al., 2018), which allows temporally separating the reward into two phases: the anticipation phase, when cues trigger reward prediction, and the outcome phase, in which reward feedback appears in case of correct performance (Haber & Knutson, 2010; Sescousse et al., 2013). These studies have established a robust and reproducible

model of the brain's reward system, anchored in the mesolimbic dopaminergic pathway (Schultz, 2015; Berridge & Kringelbach, 2015). Specifically, the reward system involves multiple brain structures that are responsible for reward-related processing, including dorsal and VStr—that includes the NAc—amygdala, insula, thalamus, various brainstem nuclei, and ACC and OFC, vmPFC and lateral frontal cortices (Knutson et al., 2001; O'Doherty et al., 2004; Haber & Knutson, 2010; For a detailed paragraph on the reward system in humans, explored using the MID, refer to paragraph 1.4.2 of Chapter 1). As such, in preclinical models of ST/GT, intra-accumbens injection of a DA antagonist substance significantly altered sign-tracking behavior, but not goal-tracking (Darvas et al., 2014; Saunders & Robinson, 2012), showing that ST behavior depends on striatal DA.

Importantly, this variability is relevant to psychiatric disorders such as SCZ, ADHD, impulse control disorders, obsessive–compulsive disorder, posttraumatic stress disorder, and addiction (Felix and Flagel, 2024). In particular, the cue-driven, DA-sensitive sign-tracking behavior provides a mechanistic parallel to the hypothesis of aberrant salience in psychosis, where dysregulated DA signaling leads to inappropriate attribution of significance to otherwise neutral cues. Therefore, this model holds significant promise for advancing our understanding of the psychological and neurobiological mechanisms that contribute to vulnerability to psychopathology.

Building on the validation of an individualized methodological approach with fMRI in the context of memory (Study 1), this Study 2 aims to apply an unsupervised machine learning technique, i.e., clustering, to translate the ST/GT framework into humans via fMRI to obtain a neuroimaging biomarker of incentive salience, which may be relevant to treatment response variability and clinical status. These studies lay the basis for fMRI-based clinical predictions and more personalized treatments.

3.1.1. The "Learning" component of the reward: the Sign-tracker and Goal-tracker animal model

Our behavior is shaped by stimuli or cues in the environment. Cues in the environment (e.g., sights, sounds, smells) play a significant role in associative learning processes because they predict positive or negative outcomes, which are essential for controlling behavior. The 'learning' component of reward is a process that enables cues that predict favorable outcomes to elicit motivated behavior. However, we know that individuals vary in their responses to environmental cues, and this variability is largely driven by differences in the propensity to attribute incentive value to a reward cue (Flagel et al., 2009). Stimuli associated with reward—either natural rewards (i.e., food, sex, water) or drugs of abuse—become imbued with incentive salience, gaining the power to control behavior (Berridge & Robinson 2003).

Individual variation in reward learning can be observed using one of the simplest reward paradigms, Pavlovian conditioning, which refers to the behavioral and physiological changes that occur when a predictive relationship is established between a neutral stimulus and a consequent biologically significant event (Pavlov, 1927, 2010). In animal research, researchers used the "autoshaping" procedure, in which the brief presentation of a localizable discrete cue acting as a conditioned stimulus (CS) is predictive of the delivery of a food reward (unconditioned stimulus, US) immediately after its termination (Brown & Jenkins, 1968). In this framework, two distinct neurobiological and behavioral profiles emerged based on their propensity to attribute incentive salience to reward cues (anticipation) and to directly focus on reward outcome: sign-trackers (ST) and goal-trackers (GT). Specifically, although no action is required to obtain the food reward, upon lever-cue presentation, ST approach and interact with the cue itself: nibbling it, sniffing it, and pressing on it (Flagel et al., 2009; Kuhn et al., 2018). They only approach the food delivery site upon retraction of

the lever-cue, at which time the food reward is delivered (Figure III.1). In contrast, GT, during the lever-cue presentation, acknowledges and sometimes orients toward the cue but then moves directly toward the location where the reward is to be delivered (Figure III.1). They primarily encode the predictive value of such cues, using them as signals that a reward is forthcoming, whereas ST attribute not only predictive but also incentive value, rendering the cue itself a "motivational magnet" capable of driving behavior (Anselme & Robinson, 2020; Berridge, 2007; Robinson & Flagel, 2009). Finally, another subset of the rats, termed intermediate responders (INT), exhibits both sign- and goal-tracking conditioned responses, with no clear preference for either and often vacillating between the two (Flagel et al., 2008, 2009; Meyer et al., 2012; Saunders & Robinson, 2012; Flagel & Robinson, 2017; Fitzpatrick & Morrow, 2020). In these studies, a large population of rats is approximately divided in one third of ST, one third INT responders, and one third GT, based on a score indicating the propensity of an individual rat to approach the CS (Figure III.1).

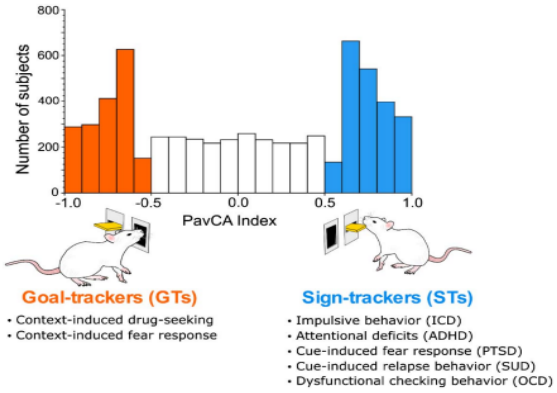

Figure III. 1 Distribution of the Sign-Tracker/Goal-Tracker Phenotypes in a Large Rat Population and Associated Behaviors of Relevance to Psychiatric Disorders. This histogram represents a frequency distribution of the propensity to attribute incentive salience to a reward cue in 6,182 Sprague-Dawley rats. The Pavlovian conditioned approach index is a composite score used to assess the propensity of an individual rat to approach the lever conditioned stimulus (sign-tracking) versus the food cup (goal-tracking) across five conditioning sessions. An index of −0.5 to −1.0 indicates that the rat is a GT. An index of 0.5 to 1.0 indicates that the rat is an ST. An index between −0.5 and 0.5 indicates that the rat is an intermediate responder. Figure adapted from Felix and Flagel, 2024.

Sign-tracking and goal-tracking behavior reflect different attributions of incentive salience

This variation in the conditioned behavior is not a matter of differences in the strength of Pavlovian learning, as both ST and GT learn the CS-food association. It rather reflects differences in the propensity to attribute incentive salience or motivational value to discrete localizable CS (Flagel et al., 2009; Meyer et al., 2012; Robinson & Flagel, 2009; Saunders & Robinson, 2012; Lovic et al., 2011), leading to differentiated stable phenotypes (Figure III.2). Incentive salience is the motivational mechanism where the CS, or more generally, an environmental cue, becomes attractive, triggering a "wanting" for the associated reward.

This attribution of value has been behaviorally quantified with the number of lever presses during cue presentation in the Pavlovian conditioned approach (PCA). In the PCA, the behavior of both ST and GT changed over time due to experience, but the nature of the conditioned response that emerged in the two groups was significantly different. Specifically, ST, which attribute high incentive salience to the CS, quickly learned to approach and manipulate the cue (lever) signaling impending reward delivery, resulting in significantly higher frequency of lever presses directed at the cue, as well as decreased latency to approach the lever. On the other hand, GT learned to approach the location where the reward would be delivered, showing minimal lever pressing and instead increasing their entries into the food magazine during the CS period (Flagel et al., 2009; Meyer et al., 2012). Further, INT responders, who exhibit both sign- and goal-tracking behaviors, show a mixed pattern of lever pressing and magazine entries, often fluctuating across trials (Meyer et al., 2012; Flagel et al., 2008, 2009). Lovic and colleagues (2011), and more recently, Gillis & Morrison (2019), replicated these results, emphasizing the attribution of incentive salience to reward anticipation cues over time in ST but not GT. This variability suggests that lever pressing is a sensitive and scalable behavioral measure for parsing individual differences in reward learning.

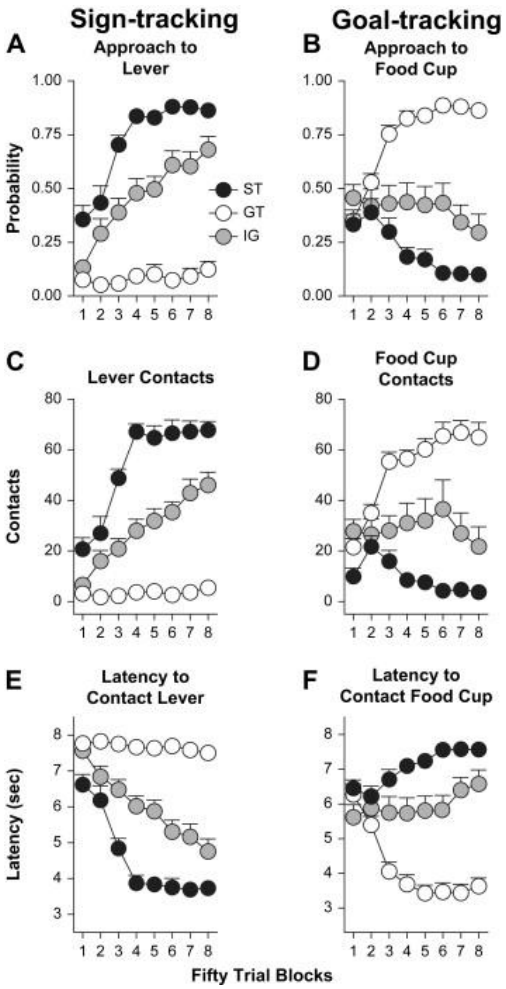

Figure III. 2 Sign-tracking and goal-tracking behavior. The black circles represent the one-third of animals that made the most lever presses, or sign-trackers (ST; $n = 14$); open circles represent those in the bottom third, or goal-trackers (GT; $n = 14$); and gray circles represent those animals comprising the middle third, or the intermediate group (IG; $n = 14$). Data are expressed as mean ± S.E.M. and illustrated in 50-trial (2-session) blocks. Note that both sign-trackers and goal-trackers developed a learned (conditional) response (CR) at about the same rate, as indicated by the change in behavior over the course of training. However, for ST the CR was directed towards the CS-lever; whereas for the GT it was directed towards the food cup. Figure adapted from Flagel et al., 2009.

3.1.2. Translating the ST/GT animal model in humans

Over the past decade, translational efforts to identify individual differences in the propensity to sign-track among humans have yielded results that closely mirror findings from animal studies (Garofalo & di Pellegrino, 2015; Colaizzi et al., 2020, 2023; Schad et al., 2020), supporting the validity and

relevance of the ST/GT framework across species, although with heterogenous methodologies (Heck et al., 2025).

The literature on human ST and GT encompasses a wide range of experimental approaches, including both computerized paradigms and more ecologically valid, real-life tasks. A core challenge in translating this model to humans lies in the difficulty of directly measuring physical approach behaviors—such as latency to contact a cue or lever pressing—which are commonly used in animal studies. A widely used paradigm in human studies is the combination of PCA with "Pavlovian-to-instrumental transfer" (PIT), which refers to the process by which a Pavlovian reward-paired cue acquires incentive motivational properties that drive choices (Holmes et al., 2010). Beyond measurement methods, studies differ widely in their categorization of participants into ST and GT groups. Categorization is typically based on a "response bias score" toward the CS, derived using tertile or median splits; however, some studies have adopted alternative techniques, such as latent profile analysis (Cope et al., 2023; Versace et al., 2016). Some frameworks include a third intermediate group for individuals displaying mixed behavioral patterns (Cope et al., 2023), as seen in animal models (Flagel et al., 2008, 2009; Saunders and Robinson, 2012), and sometimes exclude INT from further analysis (Colom, 2023; Schad et al., 2020; but Dinu et al., 2024, used both methods; Fitzpatrick & Morrow, 2020).

Findings in animal studies showed heightened striatal activation in ST relative to GT in response to reward-associated cues, a pattern interpreted as reflecting greater attribution of incentive salience to the cue itself (Flagel et al., 2007, 2010; Darvas et al., 2014; Singer et al., 2016; Gillis & Morrison, 2019). In contrast, GT focus on the reward outcome, demonstrating a more cognitively controlled strategy that facilitates goal-directed behavior (Pitchers et al., 2018). Building on this knowledge, Garofalo & di Pellegrino (2015) investigated, for the first time, the extent to which such individual differences might affect the influence of reward-paired cues in humans. In the first task, participants

learned an instrumental response that led to a reward; then, in the second task, a visual Pavlovian cue was associated with the same reward. Individuals were characterized as ST or GT based on their learned eye-gazes (i.e., conditioned response) directed toward either a visual cue (CS) or reward (US) location upon cue (CS) presentation. Finally, in a third task, the PIT paradigm with a monetary reward was tested by measuring the preference for the reward-paired instrumental response when the task-irrelevant reward-paired cue was presented, in the absence of the reward itself. In ST, but not in GT individuals, reward-related cues biased behavior, resulting in an increased likelihood to perform the instrumental response independently paired with the same reward when presented with the task-irrelevant reward-paired cue, even if the reward itself was no longer available (i.e., stronger PIT effect).

More recently, Schad et al. (2020) combined eye-tracking and fMRI in humans to investigate the direction of the eye gaze and the pupillary response of healthy adult participants engaging in a Pavlovian conditioning task during fMRI. Similar to Garofalo and di Pellegrino (2015), individuals classified as ST showed a preference for eye-tracking the reward-associated cue (CS), whereas those categorized as GT primarily focused on the location of the reward (US), progressively shifting their gaze away from the cue (CS) and toward the reward (US) as the trials continued (Figure III.3, left). The sign-tracking response correlated with brain activation in the NAc, VTA, dorsal striatum, amygdala, and vmPFC (Figure III.3, right). In further alignment with the ST/GT rodent model (Robinson & Flagel, 2009) and in agreement with Garofalo and di Pellegrino (2015), the results of a PIT task indicated that both rewarding and aversive conditioned stimuli had acquired incentive salience in ST but not GT (Schad et al., 2020).

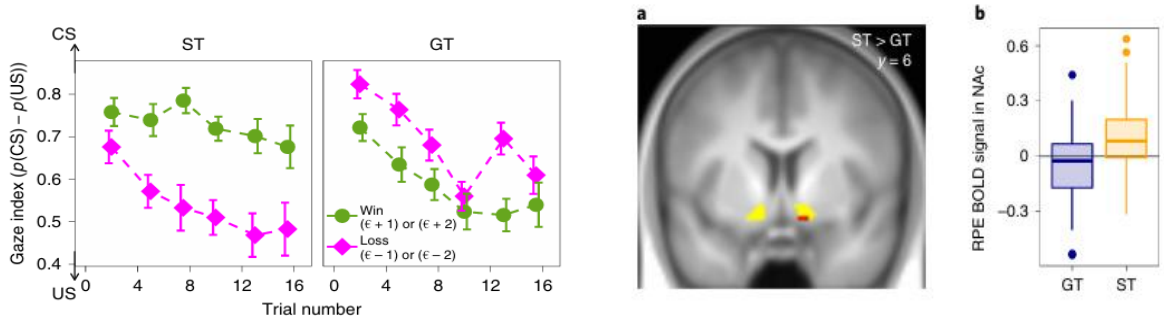

Figure III. 3 (On the left) Assessment of sign- and goal-trackers via eye-tracking. Evolution of gaze index for sign-trackers (left) and goal-trackers (right). (On the right) NAc BOLD response in sign-trackers versus goal-trackers. a. Pavlovian conditioning: a temporal-difference reward prediction error explains the right NAc BOLD response in sign-trackers better than in goal trackers (red, unmasked; a priori VOI marked in yellow; F(1, 75) = 10.88; SVC PFWE = 0.026; MNI coordinates [12, 6, –14]; η2 p ¼ 0:122 I to 0.242; n = 78 participants). y = 6 indicates the illustrated plane in MNI coordinates. b. RPE signal at the peak response difference in NAc. Box-and whisker plots show the median (centre line), upper and lower quartiles (box limits), 1.5× the interquartile range (whiskers) and outliers (points). Figure from Schad et al., 2020.

In summary, the ST/GT framework has clear translational value in understanding individual differences in motivated behavior. However, existing human studies are limited in number, often rely on small samples, and typically employ complex learning paradigms (Garofalo and di Pellegrino, 2015; Cope et al., 2023; Schad et al., 2020; Schettino et al., 2024; Colaizzi et al., 2023). These constraints hinder the application of this approach in clinical research, particularly among individuals with cognitive or learning impairments (Raio et al., 2023; Antonucci et al., 2020; Khandaker et al., 2011; Koenen et al., 2009; Barulli et al., 2013).

3.1.3. Striatal Dopamine Regulation of Incentive Salience in ST and GT

In the ST/GT model, ST and GT differ in the extent to which they rely on the reward circuit, so that, in response to a specific food- or drug-related cue, ST show greater neuronal activation throughout this circuit (Flagel et al., 2011; Yager et al., 2015). Additionally, the ST/GT model has greatly improved our understanding of the role of DA in reward learning (Flagel, Clark et al., 2011; Iglesias et al., 2023).

Recent findings suggest that sign-tracking behavior is mediated by prefrontal and ventral striatal DA and a relatively unresponsive cholinergic system (Felix & Flagel, 2024), relying primarily on "bottom-up" subcortical hypothalamic-thalamic-striatal pathways, enhancing the propensity to attribute incentive salience to reward cues with a more automatic response style toward the cue (Sarter & Phillps, 2018; Tomie et al., 2008; Enkel et al., 2019). In contrast, goal-directed behavior, in the face of cues, depends on cholinergic neurotransmission and is not dependent on DA, and relies on cortical cognitive mechanisms, reflecting a more "top-down" form of learning, that may inhibit that propensity (Flagel & Robinson, 2017; Haight, Fuller, Fraser, & Flagel, 2017; Sarter & Phillips, 2018; Enkel et al., 2019).

Several studies, including pharmacological ones, explored the role of DA for ST and GT behavior in rats, confirming that sign-tracking is DA-dependent behavior (Flagel, Clark et al., 2011; Saunders & Robinson, 2012; Chow et al., 2016; Iglesias et al., 2023). For example, when rats were given an injection of DA antagonist flupenthixol in the core of the NAc during Pavlovian training, both sign- and goal-tracking behaviors were initially suppressed. However, under drug-free conditions, only goal-tracking behavior recovered (Flagel, Clark et al., 2011). Additional studies have shown that blocking DA receptors locally in the NAc (Figure III.4; Saunders & Robinson, 2012) or lesioning the NAc (Chow et al., 2016) selectively impairs ST, but not GT. More recently, optogenetic inhibition of DA neurons in the VTA during cue presentation prevented the development of sign-tracking behavior (Iglesias et al., 2023). According to this framework, DA mediates the transformation of reward-predictive cues into salient, "wanted" stimuli (Berridge, 1996, 2007; Berridge & Robinson, 1998), and is not essential for "liking" (pleasure) or for learning associations *per se* (Berridge et al., 2009). Furthermore, ST also exhibit faster DA reuptake due to a higher density of DA transporters, which may increase their sensitivity to psychostimulants (Singer et al., 2016). Additionally, D1 receptors, which are associated with excitatory signaling, are necessary for the acquisition of sign-tracking

behavior (Clark et al., 2013). In contrast, D2 receptors, but not D3, appear to regulate both sign- and goal-tracking expression, although the precise mechanisms remain unclear (Fraser et al., 2016), as altered D2 receptor expression in the PFC and striatum has been linked to addiction vulnerability in both animal models and humans (Briand et al., 2008; Flagel et al., 2016; Asensio et al., 2010; Volkow et al., 2006). These findings underscore the importance of dopaminergic mechanisms in the VStr, including NAc, in driving incentive salience and relapse risk, particularly in individuals with a sign-tracking phenotype.

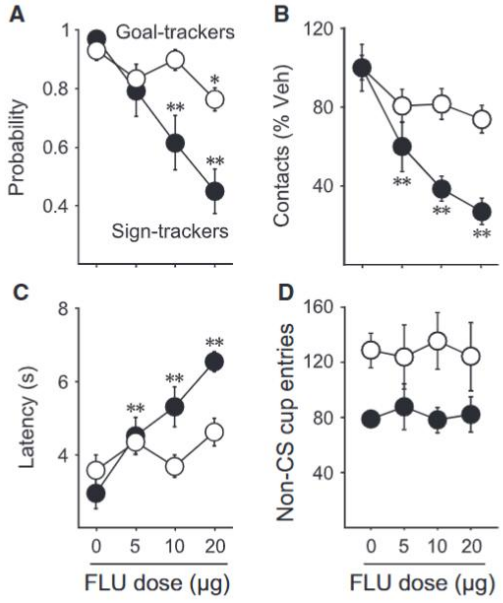

Figure III. 4 Effects of flupenthixol on STs (n = 16) and GTs (n = 13). Data presented correspond to trials 13–25 of experimental sessions. (A) Probability of making a ST or GT CR. (B) Number of contacts in STs and GTs CRs (statistical analysis was done on the raw data, but in B the data are expressed as a % of vehicle, in order to directly illustrate the relative size of the effect of flupenthixol that would otherwise be obscured by group baseline differences in total responding). (C) Latency to make a ST or GT CR. (D) Number of non-CS food cup entries. *P < 0.05, **P < 0.01 (relative to vehicle). Error bars indicate SEM. Figure from Saunders & Robinson, 2012.

3.1.4. A clinical application: alterations in reward processes in schizophrenia

Decades of research have emphasized the important role of presynaptic mesostriatal DA dysfunction, especially the increased capacity for DA synthesis and release, in the pathogenesis of psychosis. In

psychiatric disorders characterized by psychosis, such as SCZ (Gold et al., 2008; Strauss & Gold, 2012; Howes et al., 2012, 2020; Radua et al., 2015), this dysfunction in signaling the incentive salience of reward (Berridge & Robinson, 1998) is directly connected to impairments in the "wanting" of a reward (i.e., anticipation; Radua et al., 2015; Shackman et al., 2025).

In particular, SCZ is a neurodevelopmental brain disorder (Weinberger, 1987; Murray et al., 1992; Weinberger, 2017) with a complex clinical manifestation, including positive symptoms (delusions, hallucinations, disorganized thinking and behavior), negative symptoms (affective flattening, social withdrawal, apathy, loss of initiative), and cognitive symptoms (widespread cognitive deficits in domains like attention, working memory, executive functions, verbal and visual learning, and processing speed). Specifically, cognitive deficits are the strongest predictors of poor functional outcomes, as they are not responsive to existing treatments (Howes & Murray, 2014; McCutcheon, 2023), which in turn impact an individual's ability to work, live independently, and maintain social relationships (Stepien et al., 2018). Pharmacological treatment involves the use of antipsychotic drugs, which act on the dopaminergic system, and can be divided in two classes: first-generation antipsychotics—which blocks D2 receptor of the mesolimbic pathway causing a reduction of release of DA—and atypical antipsychotics—which exhibit a lower level of D_2 receptor affinity (Juckel et al., 2006; Nielsen et al., 2012; Schlagenhauf et al., 2008).

According to the aberrant salience hypothesis of SCZ, dysregulated DA signaling in patients with SCZ results in an over-attribution of meaning and motivational value (incentive salience) to rewarding and neutral stimuli, with antipsychotics ameliorating psychotic symptoms by normalizing these signals (Kapur, 2003; Heinz, 2002; Li et al., 2017; Robinson et al., 2020; Zald et al., 2017). Many fMRI studies in patients with psychosis and SCZ indicate blunted activity in the midbrain and striatum in response to rewarding and neutral outcomes (Corlett et al., 2007; Romaniuk et al., 2010; Juckel et al., 2006a, 2006; Schlagenhauf et al., 2009; Waltz et al., 2009, 2010; Gradin et al., 2013; Katthagen et

al., 2018; Nielsen et al., 2012), during both reward anticipation and outcome (Zeng et al., 2022). In addition, positron emission tomography (PET) directly quantified elevated DA synthesis and release capacity compared to HC, both in the striatum (Howes et al., 2012) and in the midbrain origin of the neurons (Howes et al., 2013; Tseng et al., 2018; Mizrahi et al., 2012; Abi-Dargham et al., 2009). However, several studies have indicated that these characteristics are not intrinsic to the disease, but they may be attributed to the effects of antipsychotics, which are associated either with the blockade of D_2 receptors (Juckel, 2016; Juckel et al., 2006), or the suppression of dopaminergic neuron firing (Bunney & Grace, 1978; Grace et al., 1997; Chiodo & Bunney, 1983). Indeed, Sportelli et al. (2024) found that a higher polygenic risk score for SCZ within the dopamine-system-enriched C80 gene set predicted greater striatal activation during reward anticipation in HC, consistent with past findings of positive correlations between polygenic risk for SCZ and striatal activation during the MID task in a large sample of healthy adolescents. The acute depletion of DA through the antipsychotics may concomitantly diminish striatal BOLD activity during reward anticipation in both patients (da Silva Alves et al., 2013) and controls (da Silva Alves et al., 2011).

Consequently, the current state of knowledge does not permit a definitive conclusion regarding the contribution of DA to the various stages of reward and their relevance to the neural phenotypes of patients with SCZ. Considering the relevance of both the complexity of symptoms and heterogeneous responses to antipsychotic treatments, understanding alterations in dopaminergic neurotransmission and salience attribution by studying the ST/GT model in this clinical population can be essential for further characterizing the disorder.

3.1.5. Hypothesis and aims of the study

Although researchers have sought to develop experimental paradigms capable of capturing human analogs of these distinct behavioral phenotypes (for reviews, see Anselme & Robinson, 2020; Colaizzi

et al., 2020), there is a need for more translationally relevant approaches to explore the underlying neural mechanisms that distinguish sign- and goal-tracking tendencies in humans (Schad et al., 2020; Schettino et al., 2024)—especially when overt dysfunction is absent, as is the case in healthy populations.

To extend the knowledge of individual differences in reward processes, based on the state of the art about the ST/GT framework in both animals and humans, the first aim of the current study was to assess whether ST/GT phenotypes emerge in humans in the absence of new learning while still enabling distinct evaluation of reward anticipation and outcome. Recognizing that behavior fundamentally arises from brain activity (Buzsáki, 2004; Fox et al., 2007; Buzsáki, 2010; Poil et al., 2012), here we focused on the neural substrates of reward processing. We used fMRI to examine brain responses during the MID task (Knutson et al., 2000). The MID paradigm used in human fMRI studies has the advantage of eliminating the learning component that characterizes the Pavlovian conditioning approach, since participants were explicitly instructed on cue–reward associations during a pre-scan training session, thereby allowing the isolation of the neural correlates of the salience attributed during the anticipation or reward delivery phase. Brain activity is measured during the reward anticipation phase—when cues signaled the potential for receiving a monetary reward—and the outcome phase—when participants received monetary reward feedback based on their performance in a simple discrimination task.

As our Discovery cohort, we utilized data from an open-access consortium of 890 healthy young adults to identify two distinct neurocognitive profiles analogous to the ST/GT phenotypes observed in rodent models. We hypothesized that individuals with ST tendencies would exhibit increased neural activity during reward anticipation, whereas those with GT tendencies would show greater activation during reward outcomes. Moreover, we predicted that these imaging phenotypes would differentiate behavioral tendencies, with ST demonstrating a more automatic or cue-driven

response style, while GT engaged in a more deliberate and controlled behavior (Enkel et al., 2019). To validate these findings across different populations, we analyzed an independent sample of 245 healthy participants, referred to as the Replication cohort, who completed a different version of the MID task and underwent additional neuropsychological assessments to test whether the cognitive and temperamental predispositions that distinguish ST and GT in rodents similarly characterize these human phenotypes. Because rodent ST/GT differences hinge on the persistence of incentive salience attribution, we also leveraged the large numbers of trials per condition in the Replication dataset to conduct a trial-by-trial analysis of ventral-striatal responses to test the hypothesis that sustained incentive salience attribution to reward-anticipation cues may differentiate ST from GT (Flagel et al., 2009; Schad et al., 2020).

Once the consistency and the replicability of this neurocognitive model were confirmed, the second aim of the study was to test whether ST/GT imaging phenotypes would predict differential responses to dopaminergic agents, motivated by preclinical evidence linking DA signaling to cue-driven behavior (Franken et al., 2005; Flagel et al., 2011; Bergamini et al., 2016; Kuhn et al., 2018; Colaizzi et al., 2023). In rodents, ST and GT exhibit distinct sensitivities to D_2 modulation, suggesting that cue-reactivity in ST depends on DA signaling (Morrow et al., 2011; DiFeliceantonio & Berridge, 2012; Saunders and Robinson, 2012; Gillis and Morrison, 2019), showing a greater reliance on DA-mediated reward prediction errors than GT (Schad et al., 2020). To address this translational question, we employed a double-blind, placebo-controlled, crossover design in healthy adults, comparing risperidone or haloperidol to placebo and assessing dose-response effects in the MID task (Hawkins et al., 2018, 2021), specifically in the Pharmacological Challenge cohort. We hypothesized that ST would exhibit selective reductions in cue-driven BOLD activity under D2 blockade, thereby providing a potential mechanistic basis for targeted therapeutic strategies in disorders characterized by maladaptive reward processing.

The third aim of the current study was to evaluate the potential clinical translation of the ST/GT framework. Prior investigations have consistently reported a blunted ventral striatal response to both reward anticipation and outcome in patients with SCZ (Radua et al., 2015), reflecting a global attenuation of incentive salience in this population. We therefore sought to determine whether, despite the overall attenuation of ventral striatal activation—attributable, at least in part, to antipsychotic treatment (Pessiglione et al., 2006; Hawkins et al., 2021)—meaningful interindividual variability in cue- versus outcome-driven neural responses analogous to ST and GT profiles could still be detected in medicated patients with SCZ and patients at their FEP. Finally, we aimed to investigate the predictive utility of these neurocognitive phenotypes to stratify patients according to treatment type and psychopathological profiles.

Studies investigating reward functioning in patients have a significant limitation: the comparison of healthy individuals with patients using a group-level approach. Moving beyond this methodological procedure, this study holds significant implications for mental health by demonstrating that interindividual variability in reward processing observed in healthy individuals may also be present in patients with SCZ, where reward abnormalities are also predictive of medication outcomes. Identifying distinct neural profiles associated with cognitive functioning could help delineate subgroups within the disorder, paving the way for more personalized and targeted therapeutic interventions based on specific functional alterations.

3.2. Materials and Methods

3.2.1. Participants

The study included fMRI data from 1,204 individuals (described in the Supplementary Information, Appendix B1), comprising a Discovery cohort of 890 healthy participants from IMAGEN (Schumann et al., 2010) (454 female, 436 male; mean age=22.1 years, SD=.7 years), a multi-site European study, an independent Replication cohort of 245 healthy individuals collected at the University of Bari Aldo Moro (UNIBA; 141 female, 104 male; mean age=25.7 years, SD=5.7), a Pharmacological Challenge cohort of 34 healthy males (mean age=26.9 years, SD=6.8), and a Clinical cohort of 34 individuals collected at UNIBA (24 males; mean age=24.7 years, SD=5.6), including 17 with FEP (11 with Psychotic Disorder NOS, 2 with Brief Psychotic Disorder, 1 with Brief Psychotic Disorder/Mood Disorder NOS, 1 with Paranoid Type Schizophrenia, and 2 with an unspecified diagnosis), and 17 with SCZ, confirmed by the Structured Clinical Interview for DSM Disorders (SCID). Inclusion criteria were the same as those for the healthy subjects enrolled in the independent Replication cohort (see Supplementary Information, Appendix B1). Demographic descriptions of four cohorts are provided in Table III.S1.

While the Discovery and Replication cohorts included only one time point per participant, the Pharmacological challenge cohort participated in a double-blind, placebo-controlled, three-period cross-over study. Specifically, data from two cohorts were considered here (Hawkins et al., 2018; Hawkins et al., 2021), in which participants were pseudo-randomly assigned to receive either risperidone (0.5 mg or 2 mg) and placebo (N=17) or haloperidol (3 mg), and placebo (N=17), undergoing three MRI sessions at one-week intervals (a detailed description of the sample and the design is provided in the Supplementary Information, Appendix B1).

3.2.2. Experimental design

All four cohorts completed an fMRI scan, performing a variation of the MID task (Knutson et al., 2000; Knutson and Greer, 2008). Additionally, a neuropsychological assessment was conducted in both the Discovery and Replication cohorts, encompassing evaluations of personality and cognitive function.

In the Pharmacological Challenge cohort, participants rated their momentary cognitive-affective state using 16 visual analog scales (VAS) anchored by opposing descriptors (e.g., Alert–Drowsy, Focused–Distracted, Calm–Stressed). Ratings were collected 90 minutes before and after drug or placebo administration, with higher scores indicating greater endorsement of the right-end descriptor on each scale.

Finally, in the Clinical cohort, clinical symptoms were assessed during a clinical evaluation preceding the MRI session using the PANSS.

A summary of the experimental design is reported in Figure III.5.

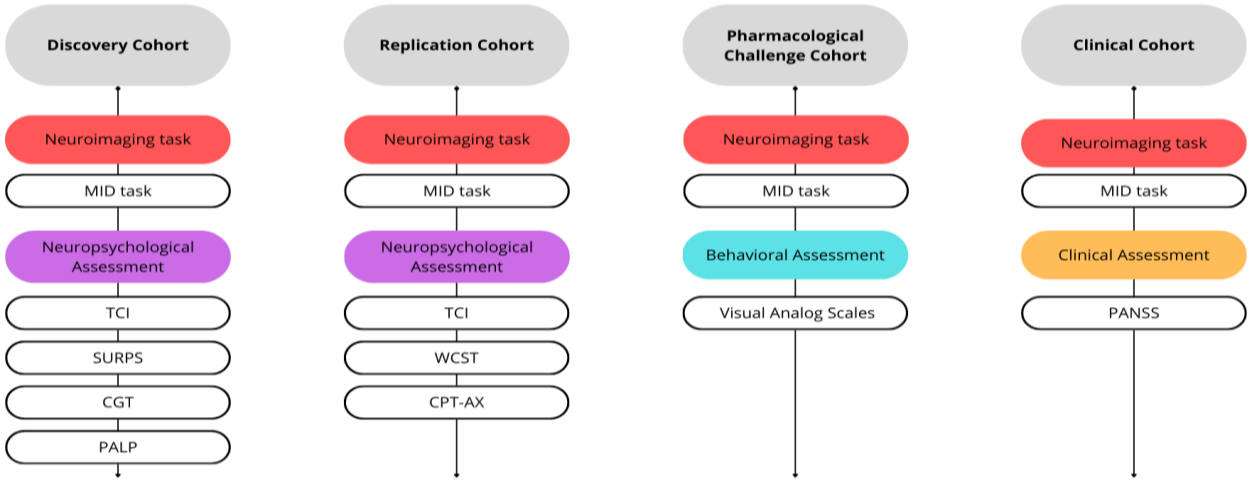

Figure III. 5 Experimental design. All four cohorts performed the MID fMRI task. Discovery and Replication cohorts include a neuropsychological assessment encompassing both personality (TCI-R, SURPS) and cognitive evaluation (CGT, WCST, CPT-AX), as well as reward sensitivity (PALP). In the Pharmacological Challenge cohort, participants performed a battery of 16 Visual Analog Scales 90 minutes before and after drug/placebo administration. In the Clinical cohort, the standardized clinical interview, i.e., Positive and Negative Syndrome scale (PANSS), were used to assess symptom severity. *Abbreviations*: Monetary Incentive Delay (MID), Temperament and Character Inventory-Revised (TCI-R), Substance Use Risk Profile Scale

(SURPS), Cambridge Guessing task (CGT), Continuous Performance test (CPT-AX), Passive Avoidance Learning Paradigm (PALP), Positive and Negative Symptoms Scale (PANSS).

3.2.2.1. Neuroimaging task

The four cohorts employed a slightly different variation of the MID task.

Discovery cohort

This version of the MID task (Schumann et al., 2010), consisted of 66 10-second trials in total and 22 trials per condition, in which small and large monetary gains were indicated by different cues. On each trial, an incentive cue indicating potential rewards was displayed on the left or right side of a black screen for 250 ms. There were three types of incentive cues (i.e., three within-subject conditions): a circle with two lines (large reward: 10 points), a circle with one line (small reward; 2 points), and a triangle (no reward: 0 points). After a variable delay (4,000-4,500 ms) of fixation on a white crosshair, participants were instructed to respond with a left/right button press as soon as the target appeared. Only responses given within the response interval were considered correct (i.e., reward hit). Using a tracking algorithm, task difficulty (i.e., target duration varied between 100 and 300 ms) was individually adjusted to produce a 66% success rate: the response interval was shortened if the success rate exceeded 66% (making the task more difficult) and lengthened if the success rate was below 66% (making the task easier). Then, feedback on whether and how many points were won during the trial lasted a total of 2000 ms. The inter-trial interval was 3,500–4,150 ms, during which a fixation cross was presented. A trial was considered a hit when the response was made while the target was on the screen. Participants had first completed a practice session outside the scanner (~5 minutes), during which they were instructed that for every 5 points won, they would receive one food snack in the form of small chocolate candies. See Figure III.6 for a depiction of the MID task.

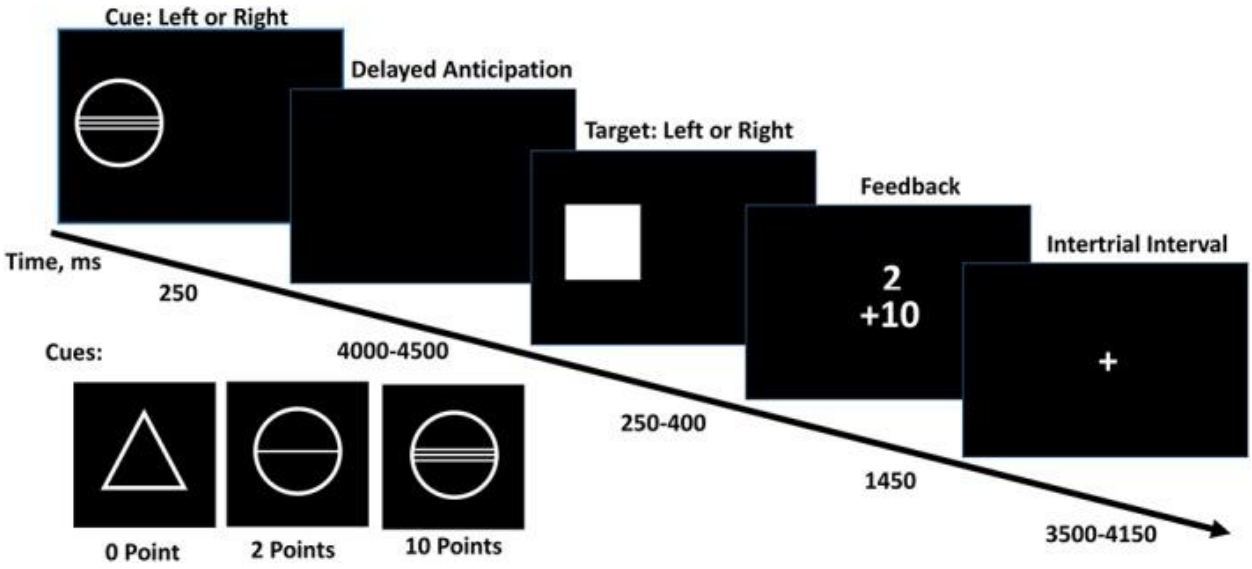

Figure III. 6 Version of the MID task used in the Discovery cohort. Cues signaling the task condition (no reward, small reward, large reward) were displayed for 4,000 to 4,500 ms. The response and feedback phase lasted a total of 2 s, during which target was displayed for 250–400 ms and feedback was displayed for 1,450 ms. trials were separated by 3,500–4,150 ms intertrial interval. Figure from Cao et al., 2018.

Replication cohort and Clinical cohort

The general layout of a trial in this version of the MID task (Sportelli et al., 2024) is divided into three phases: anticipation (2000 ms), target (up to 500 ms), and outcome (2000 ms). See Figure III.7 for a depiction of the MID task. Immediately prior to scanning, participants were instructed on the task to be performed in the scanner, including being explicitly informed of the meaning of each cue. The target presentation duration was calculated based on individual performance during the training phase to achieve at least 66% success rate. During anticipation, cue stimuli consisted of three different white geometric shapes: a full circle, representing a chance of gaining 100 points (reward condition), an empty circle, representing a chance of gaining 0 points (control condition) and a full square, representing a chance of losing 100 points (punishment condition). The target was the appearance of a white triangle. Immediately after the response, feedback appeared for 2000 ms documenting whether the participant had won or lost points as well as their cumulative total at that point. Participants were instructed to respond with a single button press with their right thumb when the target appeared to

gain or not lose points. A trial was considered a hit when the response was made while the target was on the screen. The outcome phase provided feedback on money earned from that trial and the total amount earned so far.

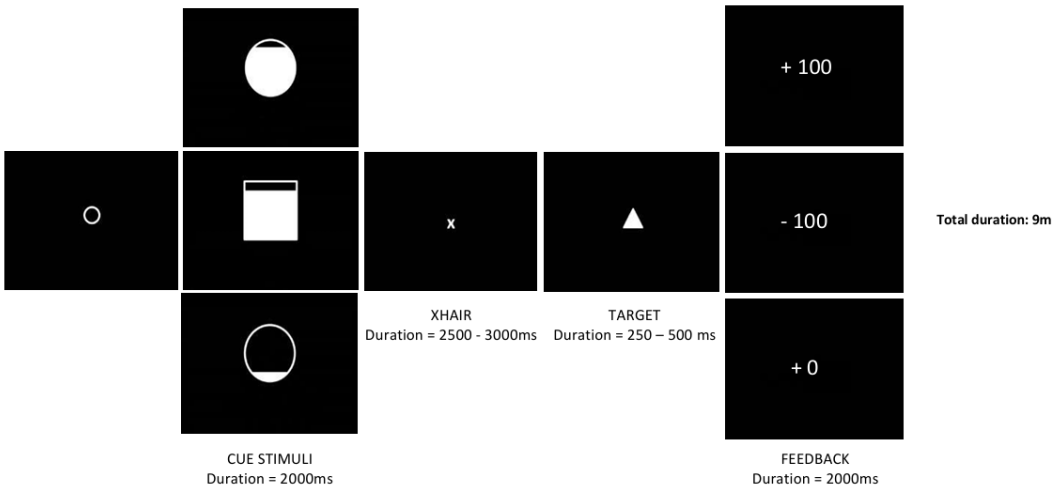

Figure III. 7 Version of the MID task used in the Replication and Clinical cohorts. Cues signaling the task condition (no reward, reward, punishment) were displayed for 2,000 ms. The target was displayed for 250–500 ms and feedback was displayed for 2000 ms.

Pharmacological challenge cohort

This version of the MID task (Knutson et al., 2001) employed four randomized trial types: three active reward-level cues (high (£2), low (£0.20), and control (£0)), and a passive trial that required no response. If the participant pressed the button during the presentation of the fixation cross or within 100 ms of the presentation of the target (an unrealistic reaction time), the trial would be set as a no-win. Each condition was presented 24 times within a total duration of approximately 13.8 minutes. The three active trial types were conducted within a fixed 10-second window, while the passive trial was a simple 'X' displayed for 4.25 seconds without the fixation, target, or feedback screens displayed in the active trials. Four separate 'playlists' with the 96 trials randomly arranged in each were created which participants randomly received on each visit according to a Latin square design, to ensure there were no learning effects from completing the same task on each visit. Performance-related criteria

were also set to ensure only data from participants who were actively and appropriately engaged in the task was included in the final analysis. During active trials, a hit was defined as a button press with a reaction time within ±3 standard deviations of the mean, and only participants with a button press rate exceeding 66% in the 500ms response window were included, ensuring active engagement. See Figure III.8 for a depiction of the MID task.

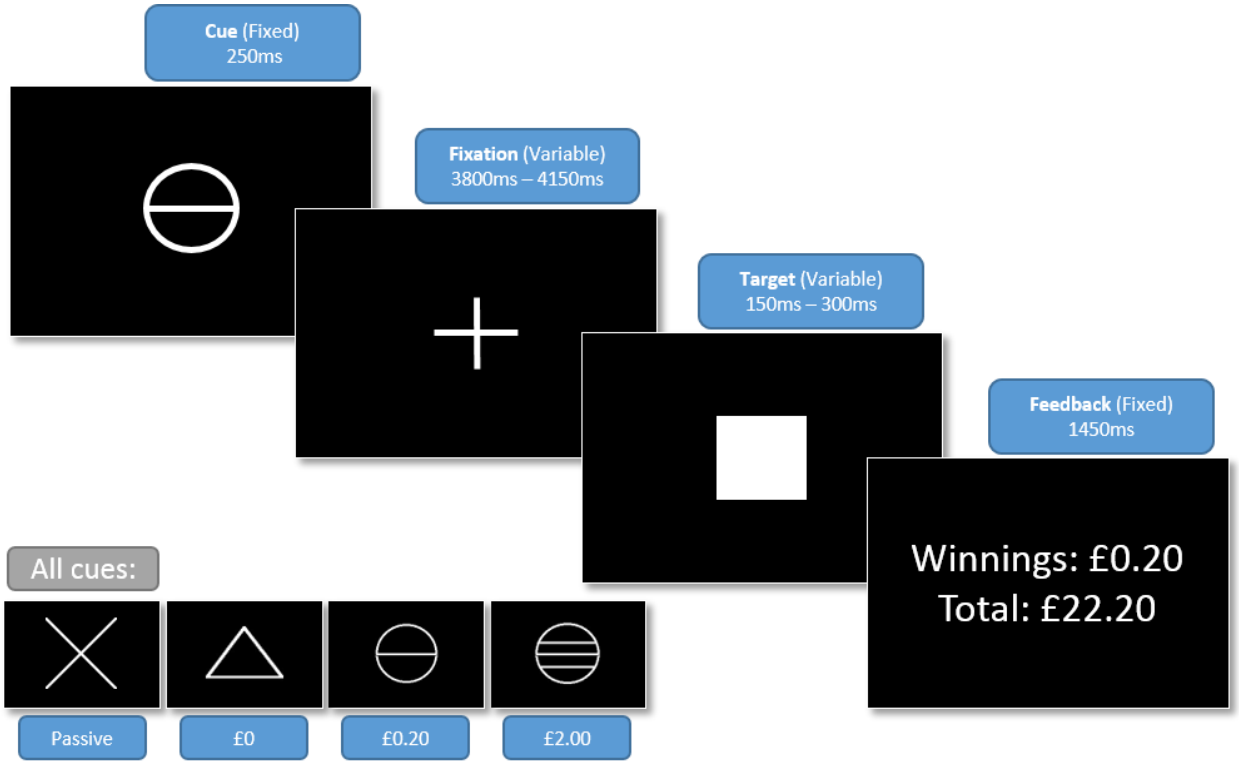

Figure III. 8 Version of the MID task used in the Pharmacological Challenge cohort. Figure from Hawkins et al., 2021.

3.2.2.2. Neuropsychological, behavioral, and clinical assessment

Discovery cohort

The neuropsychological assessment in the Discovery cohort included the Temperament and Character Inventory-Revised (TCI-R) (Cloninger et al., 1993; Farmer and Goldberg, 2008), which comprised the Novelty Seeking scale and its four temperamental subcomponents (exploratory excitability,

impulsiveness, extravagance, and disorderliness) to assess lower order trait dimensions more specifically related to disinhibitory psychopathology, and the Substance Use Risk Profile Scale (SURPS) (Woicik et al., 2009), investigating the role of main personality traits as potential risk factors for addiction and co-morbidity in psychopathology development.

- Specifically, the TCI-R is a self-administered dimensional questionnaire developed to evaluate 7 basic dimensions of personality as it develops and widely varies within individuals in the general population according to Cloninger's comprehensive biopsychosocial model (Cloninger et al., 1993). Cloninger's personality model includes 4 temperament dimensions (i.e., Novelty seeking, Harm Avoidance, Reward Dependence, and Persistence) and 3-character dimensions (i.e., Self-directedness, Cooperativeness, and Self-transcendence). In the IMAGEN data collection only the 35 self-reported TCI-R items evaluating the Novelty Seeking, as the tendency to respond impulsively to novel stimuli with active avoidance of frustration, were included to assess lower order trait dimensions more specifically related to disinhibitory psychopathology: participants rated each of the 35 items related to Novelty Seeking on a 5-point Likert scale, with 1='definitely false', 2='mostly false', 3='neither true or false', 4='mostly true', and 5='definitely true'. Thus, we could compute and select for analysis purposes only 5 personality summary scores, i.e., the 4 total scores for Exploratory excitability, Impulsiveness, Extravagance and Disorderliness and the overall Novelty Seeking score to which they contribute as personality sub-dimensions.
- The SURPS is a self-reported questionnaire largely used in epidemiological and longitudinal designs to investigate the role of 4 main personality traits as potential risk factors for addictive behaviors and co-morbid psychopathology development. Specifically, in the IMAGEN 23-items version of the questionnaire, 7 items concur in measuring hopelessness levels (as the tendency to develop bleak expectations about oneself and the future), 5 items concur in

measuring anxiety sensitivity levels (as the fear of anxiety-related physical sensations), 5 items concur in measuring impulsivity levels (as lack of premeditation and difficulties with response inhibition), and 6 items concur in measuring sensation seeking levels (as the need for intense and novel experiences). Participants rate each item on a 4-point Likert scale, with 1= 'strongly disagree', 2= 'disagree', 3= 'agree', 4= 'strongly agree'. Thus, we could compute and select for analysis purposes 4 summary scores, one for each assessed risk personality trait.

To assess cognitive performance across diverse cognitive domains, including executive function, decision-making and cognitive control, we selected scores from the Passive Avoidance Learning Paradigm (PALP) (Arnett and Newman, 2000) and from the Cambridge Guessing Task (CGT) included in the Imagen computerized neuropsychological battery (Cambridge Neuropsychological Test Automated Battery – CANTAB; https://www.cambridgecognition.com/cantab/).

- In the PALP, subjects must learn by trial and error to respond to "good" numbers for monetary reward and withhold response to "bad" numbers to avoid punishment (loss of money). The 9 outcome measures in the task were the number of commission errors, denoting too little behavioral inhibition (pressing the button on no-go trials), and the number of omission errors, denoting too much behavioral inhibition (not pressing the button on go trials), as well as the reaction times (ms), across each combination (PP=Punishment-Punishment; RP=Reward-Punishment; RR=Reward-Reward).
- The CGT task was developed to assess decision-making and risk-taking behavior outside a learning context. Relevant information is presented to the subjects 'upfront' and there is no need to learn or retrieve information over consecutive trials. On each trial, the subject is presented with a row of ten boxes across the top of the screen, some of which are red and some of which are blue. At the bottom of the screen are rectangles containing the words 'Red'

and 'Blue'. The subject must guess whether a yellow token is hidden in a red box or a blue box. In the gambling stages, subjects start with a number of points displayed on the screen and can select a proportion of these points displayed in either rising or falling order in a second box on the screen to gamble on their confidence in this judgment. A stake box on the screen displays the current amount of the bet. The subject must try to accumulate as many points as possible. To make the task shorter and more interesting for adolescents, the time between stakes is reduced from 5s to 2s. Stakes are displayed in ascending order first, and then in descending order. The 6 scores, including delay aversion, deliberation time, overall proportion bet, quality of decision-making, risk adjustment, and risk-taking, were computed and included in the between-group analysis.

Replication cohort

To replicate the neuropsychological and cognitive measures used in the discovery cohort, we employed 7 personality scores from the TCI (Cloninger et al., 1993) (Novelty Seeking, Harm Avoidance, Reward Dependence, Persistence, Self Directedness, Cooperativeness and Self-Transcendence), 4 cognitive scores from the Wisconsin Card Sorting test (WCST) (Heaton and Pendleton, 1981; Test, 1993), which assess executive functions and 7 scores from the Continuous Performance Test (CPT) (Lee and Park, 2006), assessing sustained attention and cognitive control. No measure of decision-making and passive avoidance learning was included in the data collection.

- In particular, the WCST is a well-established measure of general executive function evaluating several "frontal" lobe functions including strategic planning, organized searching, utilizing environmental feedback to shift cognitive sets, directing behavior toward achieving a goal, and modulating impulsive responding. The test uses stimulus and response cards (64-card scoring version) characterized by various forms, colors, and numbers. The participant is told to match the cards, but not how to match them; however, the subject is told whether a particular match

is right or wrong. As the test progresses, there are unannounced shifts in the sorting principle which require the subject to alter her/his approach. Performance data are recorded as the number of completed categories, number of trials, and number of perseverative and non-perseverative errors.

- The CPT–AX examines the attentional domain, specifically sustained and selective attention. During the task, single letters were presented sequentially on the screen, and participants responded when they saw an X after an A. Randomly, a non-cue letter could precede the target (e.g., B-X), or a distractor could follow a cue (e.g., A-Y), or a non-cue letter could precede a distractor (e.g., B-Y). The ratio of the target stimuli actually detected (HIT), the ratio of the false alarms, i.e. the inappropriate responses given to a distractor appearing after the letter "A" (FA) (Antonucci et al., 2020) and "B" (FAC), a sensitivity index calculated using the proportion of total false alarms (AP S), the proportion of AY errors (AY), individual's ability to discriminate signal from noise (DP S), and sensitivity of target detection in relation to the nature of the preceding stimulus (DPC S), were computed and implemented in the between-group analysis.

Pharmacological Challange cohort

In the Pharmacological Challenge cohort, participants completed a battery of 16 VAS designed to assess their momentary cognitive and affective state. Each scale consisted of a 100-mm horizontal line anchored by opposing descriptors, with participants indicating their current experience by marking a point along the line. Scores were recorded in millimeters from the left (0 mm) to the right (100 mm), such that higher values reflected greater endorsement of the right-side descriptor. The 16 VAS items included the following pairs: Alert–Drowsy, Energetic–Tired, Strong–Feeble, Motivated–Unmotivated, Determined–Aimless, Focused–Distracted, Clear-headed–Mentally foggy, Able to concentrate–Unable to concentrate, Calm–Stressed, Relaxed–Anxious, At ease–On edge, Confident–

Unsure, In control–Overwhelmed, Sociable–Withdrawn, Cooperative–Resistant, and Friendly–Irritable. Ratings were collected at two time-points: immediately before drug or placebo administration (pre-dose) and again 90 minutes later (post-dose), approximately 30 minutes before the MRI session.

Clinical cohort

In the Clinical cohort, treatment information was available for 23 out of 26 in the final sample on ST/GT. For these patients, the chlorpromazine equivalent was computed using Gardner et al. 2010. In this sample, clinical symptoms were assessed during a clinical evaluation preceding the MRI session using the standardized, clinical interview PANSS. It comprises 30 items on three subscales: 7 items covering positive symptoms (hallucinations and delusions), 7 covering negative symptoms (blunted affect and social withdrawal), and 16 covering general psychopathology (somatic concern, anxiety, and depression). Each item is scored on an item-specific scale ranging from 1 to 7. The psychometric properties of the PANSS have been well established within adult literature, and it is considered one of the most widely used research measurement tools.

3.2.3. fMRI Acquisition

Discovery cohort

Structural and functional MRI data were acquired at eight IMAGEN assessment sites with 3-T scanners of different manufacturers (Siemens, Philips, General Electric, and Bruker). The scanning variables were specially chosen to be compatible with all scanners. The same scanning protocol was used in all sites.

High-resolution T1-weighted 3D structural images were acquired for anatomical localization and coregistration with the functional time series. BOLD functional images were acquired with a gradient-echo echoplanar imaging (EPI) sequence and using a relatively short echo time to optimize reliable imaging of subcortical areas. 300 vol were acquired for each participant, and each volume

consisted of 40 slices aligned to the anterior commission/posterior commission line (2.4 mm slice thickness and 1 mm gap). The echo time was optimized (echo time = 30 ms; repetition time = 2200 ms) to provide reliable imaging of subcortical areas. The acquisition parameters per site and quality control procedures are described elsewhere (Schumann et al., 2010).

Replication cohort and Clinical cohort

SMRI and fMRI data were acquired with a 3 T Philips Ingenia scanner. High-resolution T1-weighted 3D structural images were acquired for anatomical localization and coregistration with the functional time series. BOLD functional images were acquired with gradient-recall echo-planar imaging with the following parameters: TR = 2000 ms; TE = 38 ms; flip angle = 90; 64 × 64 matrices; FOV = 240 mm; and 38 3.6 mm slices acquired with an interleaved order of slice acquisition.

Pharmacological challenge cohort

All scans were conducted on a GE MR750 3-Tesla scanner using a 12-channel receive-only head coil. Functional scans (MID and breath-hold) were carried out using a temporal series of Gradient Recalled Echo Planar Imaging (GE-EPI) whole brain scans, each comprising 38 near-axial slices, with an isotropic spatial resolution of 3.3 mm and the following parameters: TR = 2000 ms; TE = 28 ms; flip angle = 75; number of volumes = 414; FOV = 214 mm.

3.2.3.1. fMRI processing and task-activity modeling

Discovery, Replication and Clinical cohorts

We followed the same preprocessing pipeline for the Discovery, the Replication and the Clinical cohorts using the open source containerized Harmonized AnaLysis of Functional MRI pipeline (HALFpipe) version 1.2.2 (Waller et al., 2022) built using fMRIprep version 20.2.7 (Esteban et al., 2019) in combination with Magnetic Resonance Imaging Quality Control tool (MRIQC) (Esteban et al., 2017). Consensus steps for structural images included skull stripping, tissue segmentation, and

spatial normalization. For functional images, pre-processing included motion correction (and motion parameter extraction), coregistration, spatial normalization to MNI152 NLIN 2009c (asymmetric) space and resliced to 2mm3 and smoothing with a 6 mm FWHM Gaussian kernel. Data were high-pass filtered with a cutoff of 125 s to remove low-frequency drifts. Grand mean scaling was applied to the data. Participants were excluded if mean root-mean-square framewise displacement exceeded 0.5mm. MRIQC provides an interactive widget to rate the quality of each image (Esteban et al., 2019) for each site.

A first-level general linear model (GLM) was run to model BOLD responses to events of interest separately for each subject. GLM regressors describing the stimulus presentations for each experimental condition (i.e., cues by type, successful/unsuccessful outcomes, missing and error trials) are convolved with a double Gamma HRF and the overall model is fit for each voxel in the brain using FSL FILM.

Our contrasts of interest were created to capture the individual reward sensitivity by contrasting the BOLD signal of rewarded trials to control cue events. Thus, for the anticipation phase, we contrasted the brain activation corresponding to the successful trials preceded by reward cues to the brain activation corresponding to the successful trials preceded by control cue (Discovery cohort: Hit Large Reward > No-Reward; Replication and Clinical cohorts: Reward > No-Reward). Similarly, for the outcome phase, we contrasted the brain activation during the appearance of the successful feedback preceded by a reward cue during anticipation to the brain activation during the appearance of the successful feedback preceded by a control cue (Discovery cohort: Hit Large Reward > No-Reward; Replication and Clinical cohorts: Reward > No-Reward).

Pharmacological challenge cohort

Preprocessing was conducted in the Statistical Parametric Mapping (SPM) analysis suite, version 12, on Matlab 8.2.0.701, and included resetting of image origins, slice time correction, two-pass realignment, co-registration and normalization to MNI space using DARTEL (Diffeomorphic anatomical registration through exponentiated lie algebra) (Ashburner, 2007), and smoothing using an 8 mm FWHM kernel. The origin of the functional and structural images was reset to the anterior commissure-posterior commissure line. Functional images were slice time corrected (reference slice: 19). An initial between-session alignment was performed for each participant where the first volume of sessions two and three was aligned to the first volume of the first session prior to two-pass realignment within each session. All volumes were then realigned to the mean image of all three sessions. The T1-weighted image for each subject was then coregistered to the resampled mean functional image from the realignment step, using the normalized mutual information objective function in SPM, and a DARTEL (Ashburner, 2007) template created from the T1-w images. The realigned and coregistered functional volumes were resliced to original voxel sizes, and the DARTEL flow fields were applied to warp the data into MNI space. Normalized images were smoothed using an 8mm FWHM kernel. Motion and framewise displacement parameters estimated during the realignment process were added as regressors in the first-level design matrix (Siegel et al., 2014). Any volumes with a displacement of 1mm or more were flagged and marked with a 3-TR regressor (to include the volumes on either side) in the first-level design matrix (Power et al., 2012). Any scan that required more than 10% of the volumes of the full run being regressed out in this manner resulted in that participant being removed from the analysis. The realignment parameters were visually inspected, and any time series for which the maximum detected translation from the first volume was greater than the dimensions of one voxel, or those which indicated stimulus-correlated movement, were flagged for exclusion. Three participants were excluded based on these criteria. These participants

were identified and excluded parallel to data collection, allowing them to be replaced by new recruits to maintain a suitably powered study. Participants excluded due to head motion do not form part of the sample described above.

Contrasts of interest were set to explore main effect of anticipation of reward (High Cue > Neutral Cue), and main effect of receipt of reward (High Win > Neutral Cue). The MID was modelled as outlined elsewhere (Abler et al., 2007). Three cue regressors (high win, low win and neutral) were defined for the anticipatory period depending on the cue presented (as a result of temporal jittering; this period had a variable total amount of time between 4050ms and 4400ms). The target was defined by a single regressor of 500ms. Five feedback regressors (high win, low win, high no win, low no win, neutral feedback) were defined for the feedback period depending on the cue type and outcome (win or no win) and set for a fixed period (1450ms). The entire duration of passive trials was defined as a single event of 4250ms. Motion and framewise displacement parameters estimated during the realignment process were also added resulting in the model consisting of the ten task regressors above and seven movement-related parameters. Performance-related criteria were also set to ensure only data from participants who were actively and appropriately engaged in the task was included in the final analysis.

3.2.3.2. Trial-by-trial Analysis

Replication cohort

To investigate group differences in reward anticipation between ST and GT, we analyzed trial-by-trial variability across 28 reward and 28 no-reward trials in the Replication cohort. We employed AFNI's 3dDeconvolve tool (Cox, 1996) to perform deconvolution analyses, obtaining single-trial beta estimates via the beta-series regression method (Rissman et al., 2004).

Specifically, the *-stim_times_IM* option was used within *3dDeconvolve* to model each trial individually for both reward and no-reward conditions. This approach assigns a unique regressor to each trial, enabling the extraction of distinct beta coefficients that reflect the hemodynamic response associated with each event. The resulting beta estimates were then used to assess trial-by-trial variability and compare neural responses between ST and GT during reward anticipation.

3.2.4. Statistical Analyses

As illustrated in Figure III.9, the study followed a multi-step pipeline integrating neuroimaging, machine learning, neuropsychology, and pharmacological challenges with the aim of identifying imaging phenotypes of reward-related brain activity and validating them across independent cohorts. All the individual-level analyses were implemented using RStudio (2023.09.1+494).

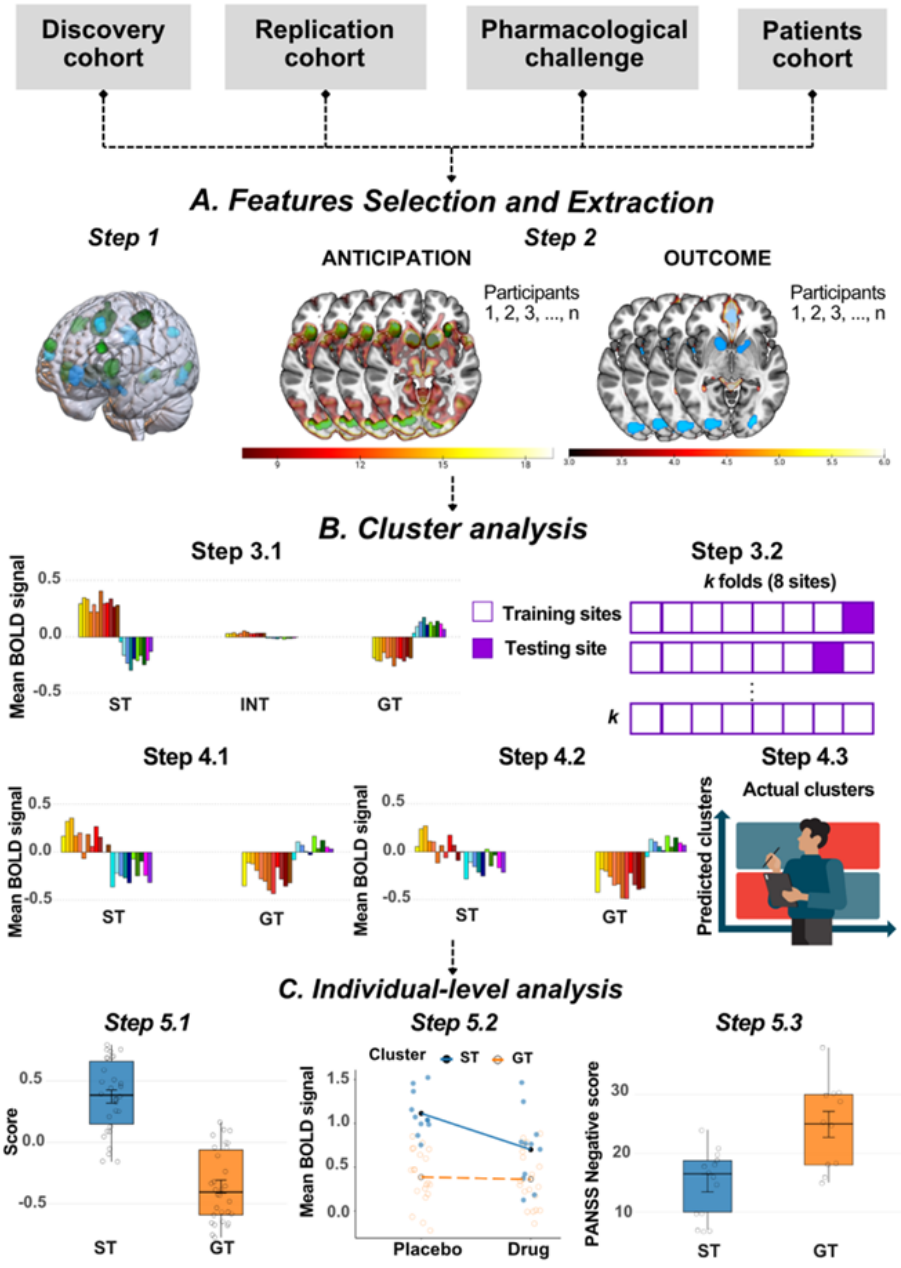

Figure III. 9 Pipeline for neuroimaging-based analyses and clustering framework. The figure illustrates the analytic pipeline from (**A**) Feature Selection and Extraction through (**B**) Cluster analyses and (**C**) Individual-level measures. In Step 1, meta-analytically derived regions of interest (ROIs) were defined for reward anticipation (bilateral VStr, thalamus, SMA, bilateral motor cortex, bilateral anterior insula, bilateral lateral occipital cortex, and bilateral middle frontal gyrus) and outcome (bilateral VStr, vmPFC, bilateral lateral occipital cortex, posterior cingulate cortex, bilateral superior parietal lobule, left middle frontal gyrus, and right inferior frontal gyrus). Next (Step 2), subject-level blood-oxygen-level-dependent (BOLD) signals were extracted from the above ROIs and used as features in a hierarchical k-means cluster analysis. The algorithm was first applied to the Discovery sample, identifying three data-driven clusters (Step 3.1), and validated in a leave-site-out framework (Step 3.2). To assess the generalizability of the clustering, the Discovery algorithm was applied to independent datasets, i.e., the Replication, Pharmacological Challenge and Clinical cohorts by projecting only the ST/GT centroids (Step 4.1), followed by a hierarchical clustering on the external datasets

(Step 4.2) to obtain actual data-driven clusters to be compared in a confusion matrix evaluating cross-tabulation consistency (Step 4.3). Finally, the clusters were examined in relation to neuropsychological measures (Step 5.1), effects of D_2 receptor antagonists on reward sensitivity in the VStr (Step 5.2), and symptoms assessed with the ANSS (Step 5.3). *Abbreviations:* ST = sign-trackers; INT = intermediate; GT = goal-trackers.

3.2.4.1. Machine Learning pipeline

Feature selection and extraction. Feature extraction began with the definition of meta-ROIs derived from a published meta-analysis of 81 studies that did not include the current datasets (Chen et al., 2022). These ROIs captured key regions implicated in reward processing, such as the ventral striatum during anticipation and the medial prefrontal cortex during outcome (Figures III.9, panel A, Step 1 and III.10). Using Marsbar (Brett et al., 2002), the mean BOLD signal was extracted for each individual from twelve ROIs during anticipation and ten ROIs during outcome (Step 2). This procedure was applied consistently across all cohorts.

Cluster analysis in the Discovery cohort. Clustering was first performed in the Discovery cohort to examine whether participants could be grouped based on reward-related activity (Figure III.9, panel B). Hierarchical k-means clustering (Peterson et al., 2018) was applied to the whole dataset, using Euclidean distance and Ward's linkage as implemented in the *factoextra* R package. Outliers were identified with Rosner's generalized extreme Studentized deviate test (Rosner, 1975) and removed iteratively. All features were scaled to a range of [–1,1], and fitting regression residuals accounted for site effects. The optimal number of clusters (k = 3) was determined using two complementary approaches: the *NbClust* R package and a hypothesis-driven approach. The latter is based on prior work on the ST/GT model 12,69, yielding the ST, GT, and INT phenotypes (Figure III.9, panel B, Step 3.1).

In our clustering, we hypothesized that we would find similar groups based on brain activity during the two separate phases of the task. Specifically, ST with elevated brain activity during the anticipation

phase, GT with elevated brain activity during the outcome phase, and INT with mixed brain activity during both phases.

Leave-site-out validation. To validate this solution and assess the clustering's stability, we implemented a leave-site-out framework in which the Discovery dataset was split into a training set comprising 7 sites and a test set comprising 1 site. Clustering was performed on the training set, and predicted memberships were assigned to the test set using only the ST/GT centroids. Validation was assessed using confusion matrices (Kuhn, 2008), comparing predicted cluster assignments with those obtained from independent clustering in the testing set (Figure III.9, pane B, Step 3.2). Balanced accuracy is computed from sensitivity and specificity and represents the stability of our clustering.

Generalizability of clustering. The Discovery model was then applied to the Replication, Pharmacological Challenge, and Clinical cohorts (Figure III.9, panel B, Steps 4.1–4.3). According to both animal (Fitzpatrick & Morrow, 2020) and human studies (Colom, 2023; Schad et al., 2020; Dinu et al., 2024), we projected only the two centroids – ST and GT – to obtain predicted clusters and reduce noise. Additionally, hierarchical k-means clustering with both k = 3 and k = 2 was used to examine different options. We compared the predicted clusters, derived from the projections of the Discovery centroids, with the true clusters obtained from a hierarchical k-means algorithm. The k = 3 solution continued to categorize individuals into groups reflecting ST, INT, and GT in the Replication cohort (Figure III.S13). Therefore, for the external three cohorts, we projected the two Discovery centroids (ST/GT), excluding the INT group to reduce noise. In each case, data were scaled to the same range, clustering was repeated, and predicted assignments were compared with independent clustering results via confusion matrices. Sensitivity, specificity, and balanced accuracy were computed, using balanced accuracy as the primary measure of replicability. Individuals misclassified by the Discovery model were excluded from subsequent individual-level analyses.

More detailed information about the Machine Learning pipeline is reported in the Supplementary Information.

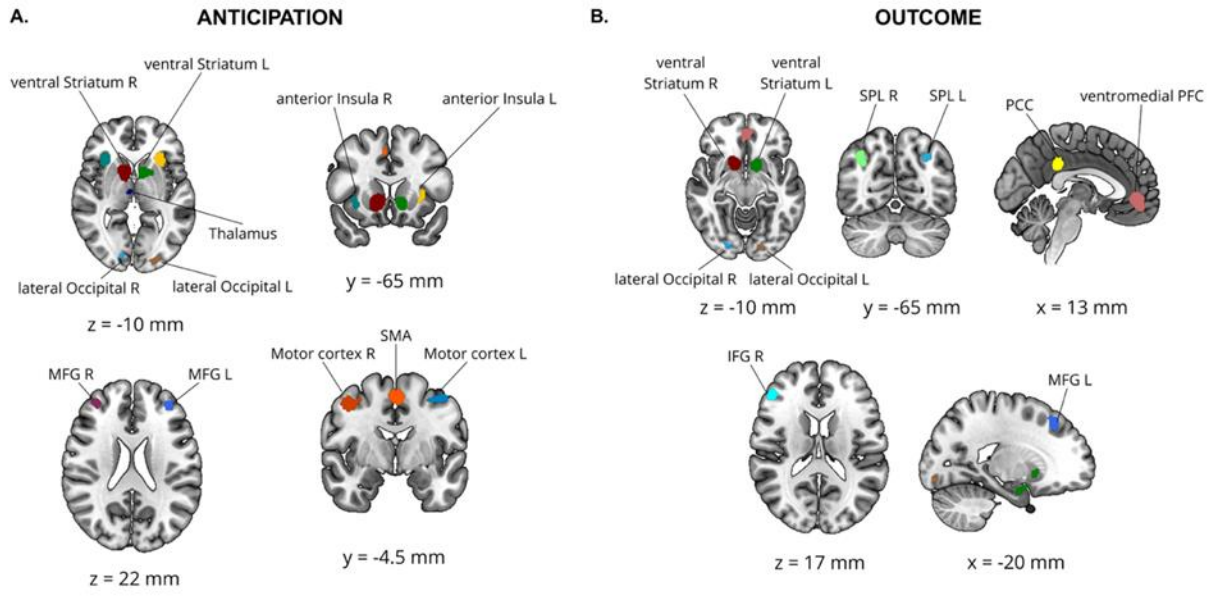

Figure III. 10 Regions of interest (ROIs) employed in the current study, derived from meta-analytic maps. (**A**) Twelve ROIs utilized for feature extraction of individual BOLD signals during the anticipation phase of the MID task. (**B**) Ten ROIs used for feature extraction during the outcome phase of the MID task. *Abbreviations:* ROIs = regions of interest; BOLD = blood-oxygen-level-dependent; MID = monetary incentive delay task; R = right; L = left; MFG = middle frontal gyrus; SMA = supplementary motor area; SPL = superior parietal lobe; PCC = posterior cingulate cortex; PFC = prefrontal cortex; IFG = inferior frontal gyrus.

3.2.4.2. Group-level task activity across clusters

In both the Discovery and Replication samples, as well as in the Clinical cohort, we evaluated task-related activity at the group level for each cluster, i.e., ST and GT, by performing separate one-sample t-tests for each task phase. These tests were conducted by entering the individual contrast maps described above using SPM12. All statistics were non-parametrically corrected for multiple comparisons using the Threshold-Free Cluster Enhancement (TFCE) approach (Smith & Nichols, 2009). Significance was set at α <0.05, using the family-wise error rate (FWE) for multiple comparison correction.

3.2.4.3. Trial-by-trial Analysis

In the Replication cohort, we implemented trial-by-trial analyses using the beta-series method (Rissman et al., 2004) to assess whether ST and GT differentially attributed incentive salience to reward anticipation cues throughout the experiment (Gillis & Morrison, 2019). The models included only individuals correctly identified as GT or ST.

We used a linear mixed-effects model (LME) to assess trial-by-trial dynamics to predict ventral-striatal BOLD responses. Following best practices for confirmatory hypothesis testing (Barr et al., 2013), we initially attempted to fit a maximal random effects structure justified by the experimental design, which included a random intercept and random slope for Trial for each subject. As this model failed to converge, we implemented a more parsimonious no-random-correlations (NRC) model that excluded the correlation parameter. This is a recommended approach to achieve a stable solution in such cases while retaining the critical random effect variance. All models were fitted using restricted maximum likelihood (REML). Additional information about the analysis is reported in Supplementary Information, Appendix B2.

3.2.4.4. Individual-level analyses

3.2.4.4.1. Association with personality and cognitive measures

In both the Discovery and Replication cohorts, we used the Wilcoxon rank-sum test (equivalent to the Mann–Whitney U test) to test the differences in personality and cognitive functioning between ST and GT (Figure III.9, panel C, step 5.1). Each variable in both the Discovery and Replication cohorts was scaled (0-1 range) to remove the differences between features. To remove any confounding effects due to the site in the Discovery cohort, we used the residuals of the regression models computed with the function *adjust* implemented in R.

We first removed outliers with the Rosner test (Rosner, 1975). In the Replication analyses, we excluded subjects who were misidentified in the cross-tabulation of cluster assignments to reduce noise in the data (final sample, N = 202). The WCST was included in the Replication cohort to reproduce the PALP processes observed in the Discovery cohort. To replicate the PALP results, showing ST to be more efficient than GT in learning shifted contingencies, we used a one-tailed Wilcoxon rank-sum test to test the specific hypothesis that ST would outperform GT on the WCST and assess the same effect direction observed in the PALP results.

To determine effect size (r equivalent), we used the rstatix R package (Tomczak and Tomczak, 2014) by dividing the Wilcoxon Z-transformed values by the square root of the number of observations (Rosenthal and Rubin, 2003). In each cohort, results were corrected for multiple comparisons grouped by task, setting statistical significance at $\alpha < 0.05$, FDR-corrected (Benjamini, 1995).

3.2.4.4.2. Placebo vs. Drug

An LME was implemented to investigate the effects of a D2 antagonist dose on reward-related responses during both the anticipation and outcome phases of the MID (Figure III.9, panel C, step 5.2). The analysis was performed using R (version 2023.09.1+494), employing the *lme4* package for model fitting and *lmerTest* for significance testing. The *clubSandwich* package was used to compute robust standard errors, accounting for potential heteroskedasticity and model assumptions violations, thereby ensuring valid standard errors under model misspecification. Robust standard errors were computed using a heteroskedasticity-consistent covariance matrix estimator, specifically CR2, which adjusts for within-subject correlations and potential non-sphericity in the data.

The dependent variable used in the model was the BOLD signal, averaged over the anticipation ROIs or the outcome ROIs. Additional information about the preliminary analyses is

reported in the Supplementary Information, Section B3. The primary model included reward phase (anticipation vs. outcome), drug (placebo vs. drug), and cluster (Sign-Trackers [ST] vs. Goal-Trackers [GT]) as fixed effects, along with their full factorial interactions. Age was included as a second-degree polynomial term to model potential nonlinear effects. Individual differences in risperidone or haloperidol levels were accounted for using a binomial variable as a covariate. To account for repeated measures, a random intercept was included for each participant.

Model assumptions were checked using the *performance* package in *R*, examining normality of residuals, homoscedasticity, and potential influential data points. Likelihood ratio tests were performed using the drop1 function to assess whether the inclusion of higher-order interactions significantly improved the model fit. Given that a significant three-way interaction was observed, two additional models were fitted separately for anticipation and outcome phases, using the same specifications described above.

To follow up on significant interactions, estimated marginal means (EMMs) were computed using the *emmeans* package in *R*. Post-hoc comparisons were performed for the effect of drug within each cluster, the difference between ST and GT within each drug condition, and the interaction between drug and phase within each cluster. To control for multiple comparisons, Tukey's method was applied, providing strong control over family-wise error rate while maintaining adequate power for pairwise contrasts.

To further test the drug effect, a classification approach based on iterative leave-one-out cross-validation was used. A hierarchical k-means clustering (Peterson et al., 2018) was employed to partition the placebo dataset into two clusters ($k = 2$) using Euclidean distance and Ward's linkage. The resulting cluster assignments were stored as the actual labels for subsequent classification analyses. An iterative classification with leave-one-out cross-validation was conducted separately for the placebo and drug conditions. For each iteration, a single observation was removed from the dataset as the test sample,

while the remaining data were used for training. Hierarchical k-means clustering (Peterson et al., 2018) was performed on the training set with $k = 2$, and the cluster centroids were determined. The observation in the testing set was assigned to the closest cluster centroid by computing the Euclidean distance from the cluster centers. The predicted cluster assignment was then compared with the actual label to assess classification accuracy. The classification performance was assessed using several metrics. Accuracy was calculated as the proportion of correctly classified test samples across all iterations. Receiver Operating Characteristic (ROC) curves were generated to evaluate classification performance, and the Area Under the Curve (AUC) was calculated as a measure of discriminative ability. Additionally, a confusion matrix was generated to compare predicted and actual classifications. To evaluate whether cluster structures identified in the placebo condition generalized to the drug condition, a hierarchical k-means clustering model (Peterson et al., 2018) was trained on the placebo dataset with $k = 2$. Each subject from the drug dataset was then assigned to a cluster based on its proximity to the placebo cluster centers. Classification accuracy was calculated by comparing the predicted cluster assignments in the drug dataset with the original placebo clusters. A contingency table was generated to assess the agreement between predicted and actual labels.

3.2.4.4.3. Dose-response effect in the reward anticipation phase

Since the Placebo vs. Drug analysis showed an interaction between cluster and drug only for the reward anticipation condition, the dose-response effect was tested only on this reward phase in the 17 subjects included in the risperidone study. An LME was used to investigate the effects of placebo, low-dose risperidone, and high-dose risperidone on response values during the anticipation phase of a reward-related task. In the LME analysis, drug condition was treated as a categorical fixed effect with three levels (placebo, low-dose risperidone, high-dose risperidone), cluster (Sign-Trackers [ST] vs. Goal-Trackers [GT]) was treated as a fixed effect factor, age was modeled as a second-degree

polynomial, and participants were included as a random effect to account for repeated measures within subjects.

Given the three-level nature of the drug factor, different model specifications were tested to determine both categorical and continuous dose-response relationships, including linear and quadratic trends. In addition to the categorical model used in the primary analysis, to further explore potential dose-dependent effects, additional models were fitted, including an ordinal model testing for a linear trend across drug conditions, a quadratic model testing for a curvilinear relationship, and a stepwise model testing the contrast between high-dose risperidone and all other conditions combined. Model selection was performed using likelihood ratio tests (LRTs) comparing each model against the null model, with significance assessed using Chi-square tests. Post-hoc analyses were conducted using EMMs, with pairwise comparisons between drug conditions adjusted using Tukey's method and FDR corrections.

In addition, to assess whether variance differences between ST and GT did not confound the observed drug effects, we performed Levene's test for homogeneity of variance under placebo and drug conditions. Finally, Bayesian analyses (BF_{01}) served to eventually test for support in favor of the null hypothesis when no significant differences were detected.

3.2.4.4.4. Behavioral Analysis

In the Pharmacological Challenge cohort, we examined ST–GT differences in treatment-related changes in behavioral outcomes using three self-report factors—Energy, Calm, and Focus—derived from confirmatory factor analyses of the full set of 16 VAS items.

The psychometric properties of the VAS items were evaluated in a two-step process using *RStudio*. First, an Exploratory Factor Analysis (EFA) was conducted on the 16 original items from the Placebo Pre-dose condition to identify the underlying factor structure in a baseline state, free from

any pharmacological intervention. Following standard practices, we used maximum likelihood estimation with an oblimin (oblique) rotation to allow for correlated factors. The number of factors to retain was determined by a convergence of criteria: the Kaiser criterion (eigenvalues > 1), a scree plot inflection point, and parallel analysis based on 100 permutations.

Second, based on the EFA results and an iterative process of model refinement to improve fit and parsimony, a 10-item, 3-factor model was specified for Confirmatory Factor Analysis (CFA). This refined model was composed of Factor 1 (Energy), which included the items Alert_to_Drowsy (loading=0.85), Strong_to_Feeble (loading=0.89), and WellCoordinated_to_Clumsy (loading=0.92); Factor 2 (Calm), comprising Antagonistic_to_Friendly (loading=0.97), Troubled_to_Tranquil (loading=0.82), and Tense_to_Relaxed (loading=0.63); and Factor 3 (Focus), consisting of MentallySlow_to_QuickWitted (loading=0.90), Incompetent_to_Proficient (loading=0.92), Attentive_to_Dreamy (loading=-0.59), and Muzzy_to_ClearHeaded (loading=0.92). The observed loadings are high, indicating strong associations between items and their respective factors. The score of Factor 1 was reversed to maintain a consistent scaling of low to high score. Notably, our 3-factor solution exhibits both similarities and distinctions compared to foundational factor analyses of similar visual analogue scales from the 1970s. Bond and Lader (1974) identified three factors: 'alertness', 'contentedness', and 'calmness'. Their 'alertness' factor encompassed items related to mental clarity and physical vigor, aligning with items in both our Energy and Focus factors. Their 'calmness' factor (specifically comprising 'Calm' and 'Relaxed' items) strongly corresponds to our Calm factor. Herbert et al. (1976) identified two primary factors, 'alertness' and 'tranquillity'. His 'alertness' factor similarly combined aspects of mental and psychomotor performance, while the 'tranquillity' factor merged elements of contentedness and calmness. Our model's differentiation of the broader historical 'alertness' domain into distinct Energy and Focus factors provides a more granular perspective on these dimensions.

The adequacy of this 3-factor structure was formally tested using the *lavaan* package. The model was fit separately to the data from four distinct conditions: placebo predose, 90 minutes post-dose in the placebo condition, predose in the drug condition, and 90 minutes post-dose in the drug condition. Model fit was assessed using standard indices, including the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). To justify the comparison of factor scores across conditions, a multi-group CFA was performed to test for measurement invariance (configural, metric, scalar, and strict) across the three conditions that were not used in the EFA. Finally, for each participant at each timepoint, factor scores for Energy, Calm, and Focus were estimated and saved using the *lavPredict* function for subsequent analysis. As a final step, factor scores for Energy were reversed so that higher values consistently reflected greater energy.

The derived factor scores were analyzed using LMEs to investigate the effects of Time (prepose, post-dose), Cluster (GT, ST), and Phase (Placebo, Drug). Analyses were conducted in R using the *lme4* and *lmerTest* packages. To align with the study design, separate analyses were conducted for the Placebo and Drug phases. For each of the three factors within each phase, an LME was fitted with the factor score as the dependent variable. In the applied model, cluster, Time, and their interaction were included as fixed effects to test our primary hypothesis. Participant age, its quadratic term, and drug cohort (risperidone, haloperidol) were included as covariates. A random intercept for each subject was included to account for the non-independence of repeated measures. P-values for the fixed effects were obtained using Satterthwaite's degrees of freedom approximation. To control for multiple comparisons across the three factors within each phase, the p-values for the Cluster * Time interaction terms were adjusted using the FDR correction. Significant interactions were followed up with post-hoc pairwise comparisons using a Tukey correction for multiple comparisons.

The final 10-item, 3-factor model demonstrated an improved and adequate fit for the data compared to the initial 16-item EFA solution. In the Predose condition during pre-dose, fit indices were acceptable. Fit remained adequate across the other conditions, with CFI values of .886 in the predose condition during drug and .858 for post-dose condition during drug, and SRMR < .09 in all conditions. Because of the small sample size, RMSEA values were elevated (> .21). Multi-group CFA supported measurement invariance across the post-dose condition during placebo, the predose condition during drug, and post-dose condition during. The criteria for configural, metric (ΔCFI = .003, p = .223), and scalar invariance (ΔCFI = .002, p = .253) were met, indicating that the factor structure, item loadings, and item intercepts were equivalent across conditions. This justifies the direct comparison of latent factor scores. Strict invariance was not supported, suggesting that residual variances differed between conditions.

3.2.4.5. Clinical analysis

In the Clinical cohort, we integrated antipsychotic receptor binding profiles with clinical dosing parameters to compute dose-normalized receptor affinity ratios. Equilibrium dissociation constants (Ki values) for DA receptors D1–D5 were obtained from the dataset published by McCutcheon et al. (McCutcheon et al., 2023) via their public GitHub repository, while clinical data comprised patient-level medication regimens, individual dosages, and chlorpromazine-equivalent doses (CPZ). Dose-to-affinity ratios were calculated as CPZ equivalents divided by the respective Ki values (nM) for each receptor subtype, yielding dimensionless indices of receptor occupancy potential at therapeutic doses. These indices quantitatively measure the relative blockade capacity for each receptor subtype under clinically relevant dosing conditions and were generated using a custom R pipeline. The final analytic sample consisted of 24 patients, after excluding individuals with missing data or misidentification during clustering. Group differences in VStr BOLD signal between patients and HC from the UNIBA

cohort were tested with independent-sample t-tests. In contrast, ST–GT differences in positive and negative symptoms were examined using Wilcoxon rank-sum tests (N = 25, with PANSS data unavailable for one patient). Correlational analyses further assessed associations between receptor affinity indices, VStr activity, and clinical symptom dimensions.

3.3. Results

3.3.1. Cluster analysis

Figure III.11A illustrates that the three-cluster solution in the Discovery cohort delineated ST (N = 161), who exhibited greater BOLD activity in anticipation-related ROIs rather than outcome, GT (N = 319) with increased activity in outcome-related ROIs than anticipation, and an INT group (N = 410) displaying relatively low brain activation in both phases. There was no evidence of sex differences based on the two results. First, the distribution of males and females did not differ between ST (M=89, F=72) and GT (M=149, F=170), $\chi^2_{(1,\ N=480)}=2.81$, p=0.094, and second, very similar imaging phenotypes emerged when the cluster analysis was conducted separately for males and females (Figure III.S1). Despite inter-site variability, the leave-site-out cross-validation analysis produced an overall accuracy of 86.7% and a balanced accuracy of 84.9%, based on sensitivity (87.8%) and specificity (82%) (Figure III.11E). Barplots showing the BOLD signal in each ROI across both projected and true clusters for each test site are reported in Figures III.S2-S9.

Cluster generalizability in the independent Replication cohort showed good clustering performance, with 82.4% overall accuracy, perfect specificity (100%), but lower sensitivity (57.4%) and 78.7% balanced accuracy. While all GT were correctly identified, some ST (N = 43) were misidentified (mis-ST), in line with rodent models where a subset of individuals does not fall into distinct categories (Figure III.11B). The mis-ST were excluded from further analyses.

Moreover, cluster replicability in the Pharmacological Challenge cohort was high (balanced accuracy of 92.3%), with 94.1% accuracy, perfect specificity (100%), and high sensitivity (84.6%), misidentifying only 2 out of 13 ST (Figure III.11C) that were excluded from subsequent analyses.

Finally, cluster replicability in the Clinical cohort showed good clustering performance (Figure III.11D), with 77.1% overall accuracy (sensitivity: 100%; specificity: 66.7%) and 83.3% replicability

based on sensitivity and specificity. While all GT were correctly identified, eight ST were misidentified (mis-ST) and were excluded from subsequent analyses.

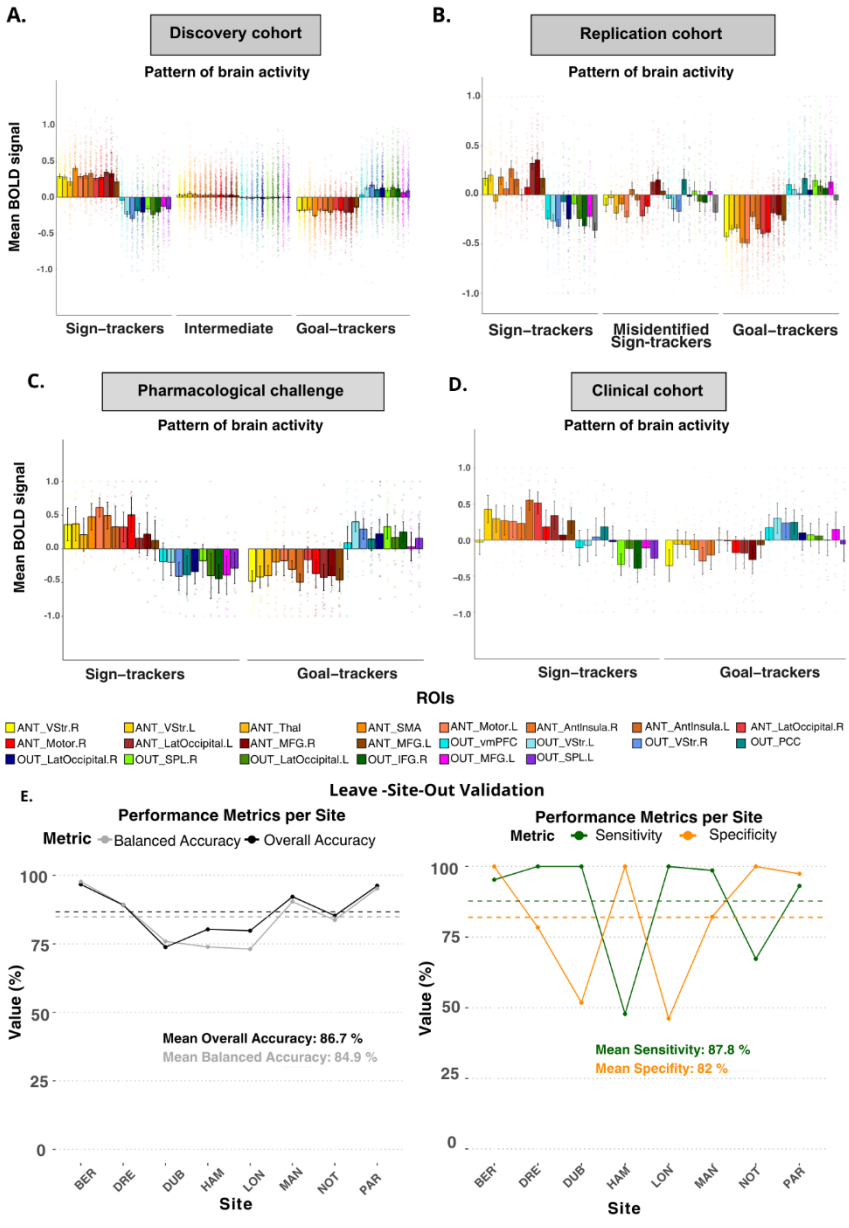

Figure III. 11 Development, validation, and characterization of three reward-related phenotypes. (**A**) A three-cluster solution identified sign-trackers (ST), goal-trackers (GT), and an intermediate group. Bars represent 95% CI. (**B**) Generalization in the Replication cohort showed high classification accuracy (82.4%), by projecting the ST and GT centroids. (**C**) Pattern brain activity of ST and GT in the Pharmacological Challenge cohort and in the (**D**) Clinical cohort. (**E**) Leave-site-out validation results. The clusters were highly reproducible across eight imaging sites (86.7% accuracy, 84.9% replicability).

3.3.2. Group-level analysis

In the Discovery cohort, whole-brain voxel-wise analyses showed significant and widespread activation in ST during the anticipation phase, and in GT during the outcome phase ($p_{TFCE\text{-}FWE} < 0.05$; Figure III.12A; Table III.S2). Minimal activation was observed in the alternate phase for each group. This pattern was replicated in the Replication cohort, where ST again exhibited extensive activation during anticipation, and GT during outcome ($p_{TFCE\text{-}FWE} < 0.05$; Figure III.12B; Table III.S2), with negligible activation during the alternate phase.

Similarly, in the Clinical cohort, whole-brain voxel-wise analyses ($p_{TFCE\text{-}FWE} < 0.05$) showed robust activation in ST during anticipation and in GT during outcome (Figure III.12C), with no significant activation during the alternate phase (Table III.S2). An independent sample t-test showed no significant difference in chlorpromazine equivalents between ST and GT (data available for 24 of the 26 patients), $p = 0.500$.

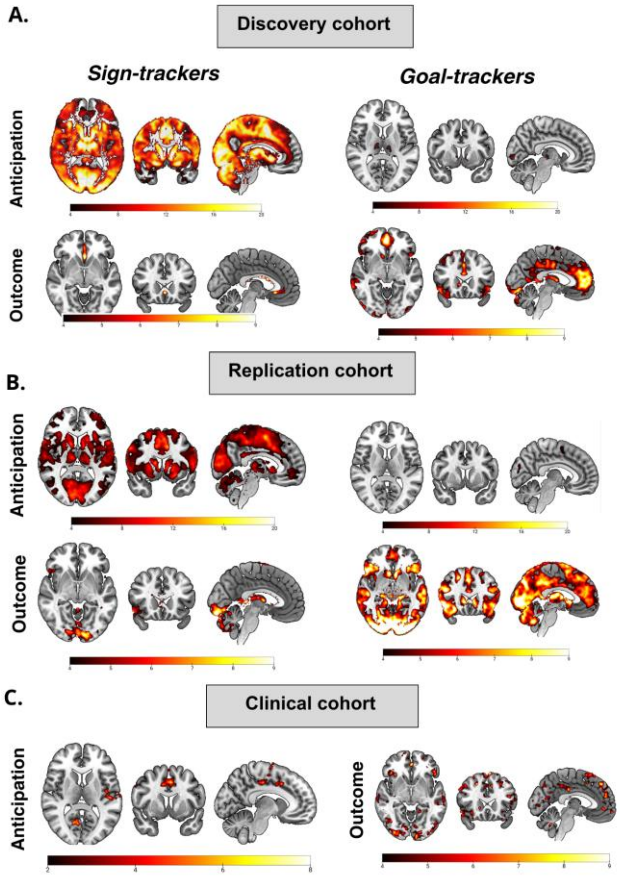

Figure III. 12 Group-level activation maps in Discovery (A), Replication (B), and Clinical (C) cohorts. Whole-brain analyses confirmed strong anticipatory BOLD activation in the ST and outcome activation in the GT, with minimal engagement during the alternate phase. All statistics were corrected for multiple comparisons using the Threshold-Free Cluster Enhancement (TFCE) approach (Smith and Nichols, 2009), with significance set at a family-wise error rate (FWE) of $\alpha < 0.05$.

3.3.3. Trial-by-trial analysis

We tested the hypothesis that ST and GT attribute differential incentive salience to the reward anticipation cues throughout the experiment using trial-by-trial dynamics. As we predicted, these trial-by-trial group differences would be specific to the reward condition, so we modeled reward and no-reward trials in separate analyses. We used a stronger VStr response across trials as an index of incentive salience and implemented linear mixed-effects models predicting ventral-striatal BOLD responses with fixed effects for Group (ST vs. GT), Trial (continuous), and their interaction. The

models included random intercepts by subject and random slopes for trial by subject, without estimating the correlation between these two terms. This no-random-correlations structure was chosen to balance model complexity and avoid potential convergence issues (Barr et al., 2013). The trial-by-trial analysis for the no-reward condition showed no differences in the VStr activity between ST and GT (β=−0.0076, SE = 0.0276, t(245.3)=−0.276, p=0.783). The Trial effect in the VStr responses indicates a slight decrease in BOLD activity over trials (β=−0.00181, SE = 0.00091, t(243.0)=−1.983, p=0.048), with no group differences in habituation rates (β=0.00160, SE = 0.00142, t(243.0)=1.126, p=0.261). In the reward condition, ST showed significantly higher VStr responses than GT (β=0.0860, SE = 0.0287, t(243.0)=2.990, p=0.003). There was no overall response decrease over trials (β=−0.00048, SE = 0.00091, t(241.9)=−0.526, p=0.599). However, a significant Trial × Group interaction (β=0.00275, SE = 0.00141, t(241.9)=2.540, p=0.008) indicated that ST maintained a stronger VStr response over trials, compared to GT (Figure III.13).

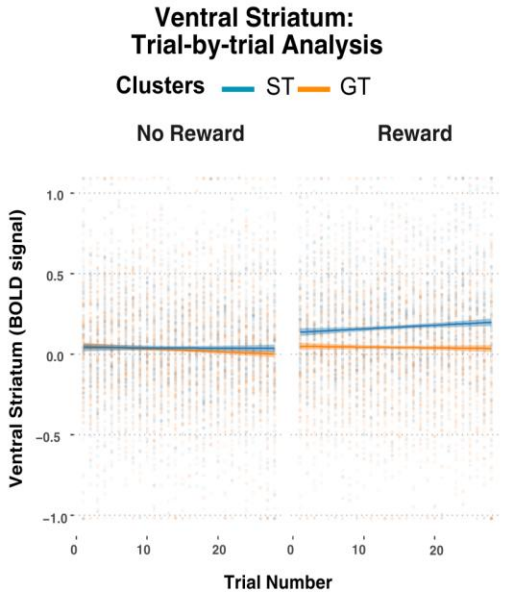

Figure III. 13 Pattern of activity during the MID experiment in the Replication cohort. Trial-by-trial analysis showed ST had consistently higher VStr activity than GT during reward trials, with no difference during no-reward trials.

3.3.4. Personality and cognitive assessment of ST and GT

For both Discovery and Replication cohorts, mean (SD) values for ST and GT, along with full statistical details of group effects, are reported in Table III.S3.

In both the Discovery and Replication cohorts, no significant differences were observed between ST and GT in personality characteristics, as assessed by the TCI in both samples and SURPS in the Discovery cohort.

In the Discovery cohort, we found a significant difference in decision-making processes between groups, as assessed through the CGT, with ST scoring higher than GT on risk adjustment (pFDR = 0.002), which is defined as the extent to which a participant adjusts their bet based on their perceived probability of winning. Significant differences emerged between groups in reward-related behavior (PALP), with GT performing more poorly than ST. In the absence of contingency switches, GT performed more poorly than ST. They made significantly more omission errors (nonresponses to 'go' trials, indicating too much behavioral inhibition) and commission errors (responses to 'no-go' trials, indicating too little behavioral inhibition) in both the reward-reward (Omission: p_{FDR}=0.01; Commission: p_{FDR}=0.002) and punishment-punishment (Omission: p_{FDR}=0.001; Commission: p_{FDR} < 0.001) conditions. This suggests that GT individuals struggle with a lack of behavioral inhibition (commission errors) and an excess of it (omission errors), indicating a general impairment in cognitive control and behavioral flexibility in a task involving reward and punishment.

In the Replication cohort, no significant ST-GT differences were found in CPT measures of sustained attention, impulsivity, and vigilance, as well as for personality traits (TCI). Instead, in the WCST, measuring executive functions, ST outperformed GT, using significantly fewer trials (p_{FDR}=0.026) and making significantly fewer perseverative and non-perseverative errors (p_{FDR}=0.042) to complete more categories (p_{FDR}>0.05, p=0.13) than GT. Results are depicted in Figure III.14.

Altogether, these results suggest that while ST and GT do not differ on measures of personality or sustained attention, the ST phenotype is associated with a distinct cognitive profile characterized by superior executive functions.

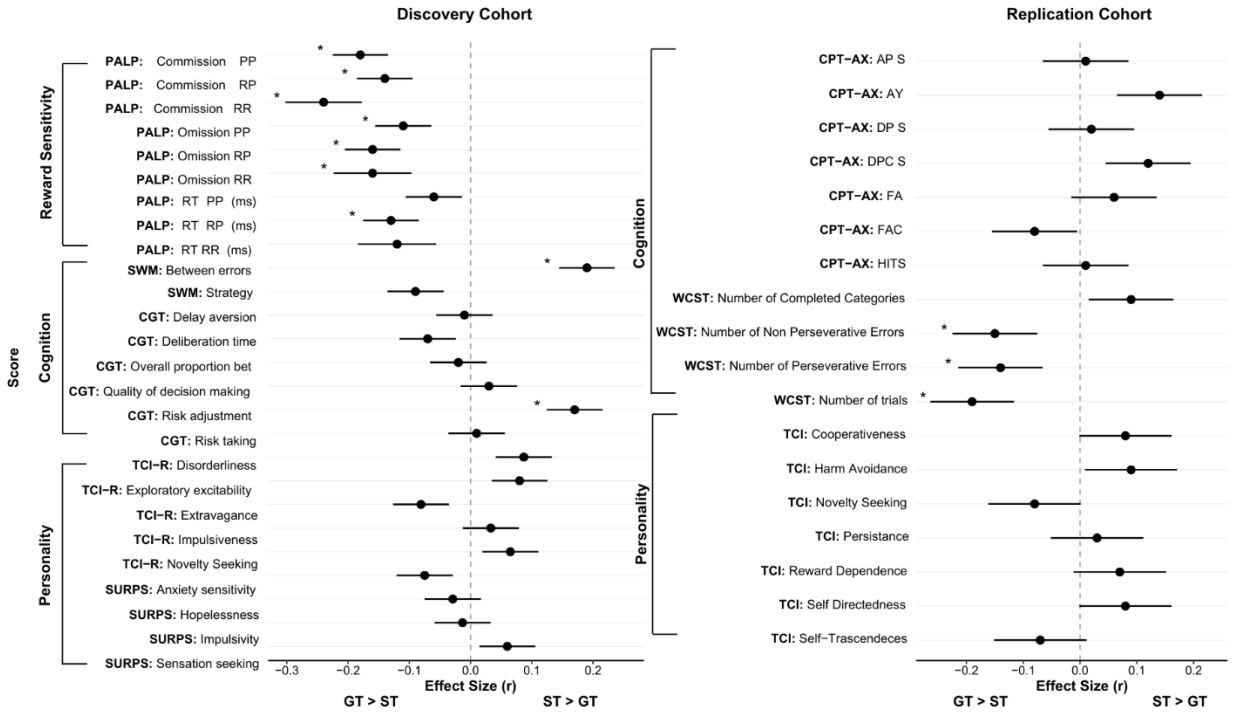

Figure III. 14 Neuropsychological results in both the Discovery and Replication cohorts. Histograms depict the effect size r of the Wilcoxon rank-sum tests performed for the neuropsychological scores grouped by domain, i.e., reward sensitivity, cognition, personality, and cohorts.

3.3.5. Placebo vs. Drug effect

Drug effects on the bilateral VStr BOLD signal were tested using linear mixed-effects models with Phase (anticipation vs. outcome), Drug condition (placebo vs. drug), and Cluster (ST vs. GT) as fixed effects, including their interactions, and with subject as a random intercept; F-tests assessed omnibus interactions and post-hoc contrasts examined estimated β-coefficients. Results are depicted in Figure III.15A. Analyses restricted to the bilateral VStr yielded a significant three-way Phase × Drug × Cluster interaction ($F(1, 90)=6.905$, $p = 0.010$), and in the anticipation phase, a Drug × Cluster interaction ($F(1, 30)=6.676$, $p = 0.015$). Under placebo, ST showed higher BOLD responses than GT ($\beta=0.697$,

SE=0.128, t(53.3)=5.433, p<0.001), and drug (risperidone or haloperidol) significantly reduced ST responses (β=-0.514, SE=0.131, t(30)=-3.935, p<0.001) but had no effect in GT (β=-0.097, SE=0.095, t(30)=-1.030, p=0.311). There was a significant Drug × Cluster interaction in the outcome phase, F(1, 30)=4.952, p=.034. Post-hoc tests showed that the drug effect (drug-placebo) was not significant in GT (β=0.387, SE=0.200, t(30)=1.934, p=0.062), but it was significant in ST, where drug increased BOLD response compared to placebo (β=1.148, SE=0.277, t(30)=4.147, p<0.001). Using whole-network estimates (in reward-related regions), a leave-one-out cross-validation framework was employed to assess classification performance between ST and GT individuals. In the Placebo condition, this classification achieved a replicability of 97.62%, considering GT as the positive class (sensitivity for ST=100%; specificity for ST=95.24%), with AUC for predicting GT of 0.976, confirming clear group separation. When this classification approach (using Placebo-based labels) was applied to drug data, the replicability in distinguishing ST from GT individuals dropped to 75.09% (AUC for predicting GT, as positive class=0.75; sensitivity for ST=69.23%; specificity for ST=80.95%). This result suggests that while the classification is less clear under drug conditions, ST individuals are still identified with sufficient sensitivity despite marked reductions in their BOLD signal. Similar results, regarding drug effect on BOLD activity, were obtained using estimates from the reward anticipation and outcome networks (Supplementary Information, Appendix B4, Figure III.S10).

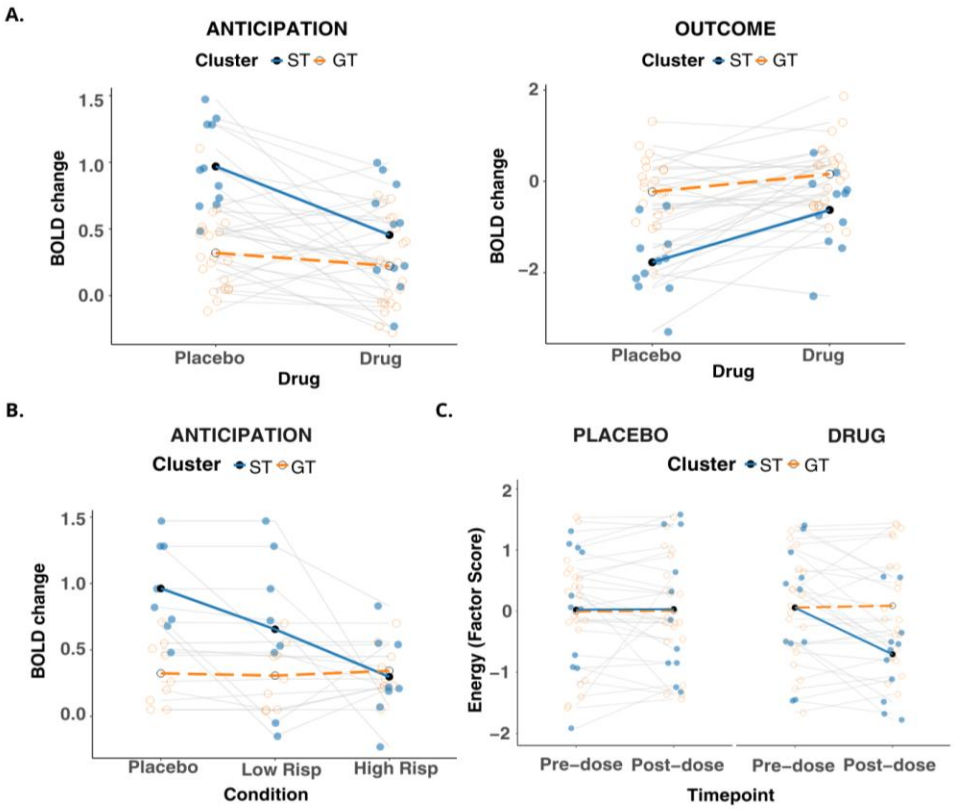

Figure III. 15 Individual-level analyses in the Pharmacological challenge cohort. (**A**) A three-way interaction between phase, drug, and cluster showed that, during anticipation, ST had higher BOLD responses than GT under placebo, and D2 antagonists reduced this activity. The drug-by-cluster interaction was not significant during the outcome phase. (**B**) Dose-response analyses indicated that ST anticipatory responses linearly decreased with increasing risperidone doses, while GT showed no change. (**C**) Behavioral analysis showed that ST had a lower Energy Factor Score after the drug dose.

3.3.6. Dose-response effect in the reward anticipation phase

The dose–response effects, available only in the risperidone dataset (N=17), were tested in the anticipation phase, where a Cluster × Drug interaction had been reliably identified in the VStr and in the whole-network analysis (see Supplementary Information, Appendix B5) and were further tested in the bilateral VStr. Results are shown in Figure III.15B.

In ST, a significant overall Drug effect was observed ($\chi^2(2) = 10.829$, $p = 0.005$) with a robust linear trend ($\chi^2(1) = 10.800$, $p = 0.001$), indicating a dose-dependent decrease in BOLD response; the

quadratic term was non-significant ($\chi^2(1) = 0.029$, $p = 0.865$). A stepwise comparison confirmed that high-dose risperidone significantly reduced ST responses relative to placebo and low-dose ($\chi^2(1)=7.825$, $p=0.005$). Post-hoc contrasts showed a significant difference between placebo and high-dose ($\beta=-0.665$, SE=0.181, $t(14)=3.677$, $p_{FDR}=0.008$), whereas the low-dose effect was not significant ($\beta=-0.3075$, SE=0.1808, $t(14)=1.700$, $p_{FDR}=0.111$). For GT, no significant drug effects emerged ($\chi^2(2) = 1.978$, $p = 0.372$), with non-significant linear ($\chi^2(1) = 0.536$, $p = 0.464$) and quadratic ($\chi^2(1) = 1.442$, $p = 0.230$) trends. Post-hoc contrasts showed no differences between placebo, low-dose, and high-dose (all $p_{FDR} \geq 0.532$). Bayesian analysis strongly supported the null hypothesis, showing no risperidone effect on anticipatory responses in GT (Bayes Factor, BF_{01}=5.47, 5.93, 3.95, respectively). The BF assesses the evidence that the data provide an effect of being present or absent. Specifically, "1" corresponds to H1 and "0" to H0. Generally, values of BF01 larger than 3 indicate evidence favoring H0 (i.e., no effect).

Analyses of dose-response effect in the anticipation phase across the whole-network estimates are reported in the Supplementary Information (Appendix B5) and depicted in Figure III.S10B.

3.3.7. Behavioral differences in ST and GT

The mixed-effects analysis of the three behavioral factors (Energy, Calm, and Focus) showed a significant three-way interaction for the Energy factor only ($p=0.043$). Follow-up analyses showed a significant Time × Cluster interaction in the drug phase ($\beta=-0.794$, SE= 0.175, $t(30)=-4.525$, $p<0.001$), indicating that the change over time differed significantly between the ST and GT groups. Post-hoc comparisons confirmed that this was driven by a significant decrease in Energy from pre- to post-dose for ST ($\beta=0.764$, SE=0.142, $t(30)=5.373$, $p<0.001$), while there was no significant change for GT ($\beta=-0.030$, SE=0.103, $t(30)=-0.293$, $p=0.997$). Results are shown in Figure III.15C.

The analysis found no significant Time × Cluster interaction in the placebo phase (p=0.980). Furthermore, there were no significant three-way interactions for the Calm (p=0.444) or Focus (p=0.949) factors.

3.3.8. Clinical assessment

When comparing patients to neurotypical controls (NC) derived from the Replication cohort (collected with the same imaging protocol at UNIBA), ST patients showed a lower BOLD signal in the ventral striatum during anticipation, t(71)=7.12, p<0.001, Cohen's d=2.06 (Figure III.16A). In contrast, no differences were found between GT patients and NC, t(153)=1.02, p=.3084, Cohen's d=0.32. Nonetheless, within patients, ST showed larger BOLD activity in the VStr during reward anticipation compared to GT, t(24)=2.72, p=.0121, Cohen's d=1.08, in line with what was observed in NC, t(200)=16.02, p<.0001, Cohen's d=2.49. Furthermore, to investigate the relationship between antipsychotic medication and neural activity, we calculated a patient-specific D_2 dopaminergic affinity index by integrating individual medication dosages with established receptor binding affinities (McCutcheon et al., 2023). Consistent with our hypothesis, this index of D_2 receptor blockade was significantly and negatively correlated with striatal BOLD activity during reward anticipation (r=-0.462, p=0.023, Figure III.16B), after accounting for age and sex. Affinity for other DA receptors yielded no significant associations with striatal BOLD activity in anticipation (D_1: p=0.417; D_3: p=0.105; D_4: p=0.323; D_5: p=0.245).

Differences in clinical symptoms, assessed through the PANSS, were tested in association with BOLD signal and D_2 dopaminergic affinity. We found that the BOLD signal in the VStr during anticipation was significantly and specifically associated with negative symptoms (r=−0.579, p=0.002, Figure III.16C), while no association was found with either positive symptoms (r=0.03, p=0.901) or general psychopathology items (r=−0.15, p=0.486). Consistent with these findings, GT individuals

demonstrated significantly higher negative symptom severity than ST individuals (Wilcoxon W=118, p=0.006, Figure III.16D). Furthermore, a significant positive association was found between D_2 dopaminergic affinity index of the administered therapy and negative symptom severity (r=0.471, p=0.020, Figure III.16E). No other significant associations were detected between striatal BOLD activity in anticipation and dopaminergic affinity (D_1: p=0.336; D_3: p=0.244; D_4: p=0.140; D_5: p=0.320). Together, these results suggest that patients administered antipsychotic treatments characterized by stronger D_2 blockade also show a GT-like imaging phenotype and have a higher negative symptom burden.

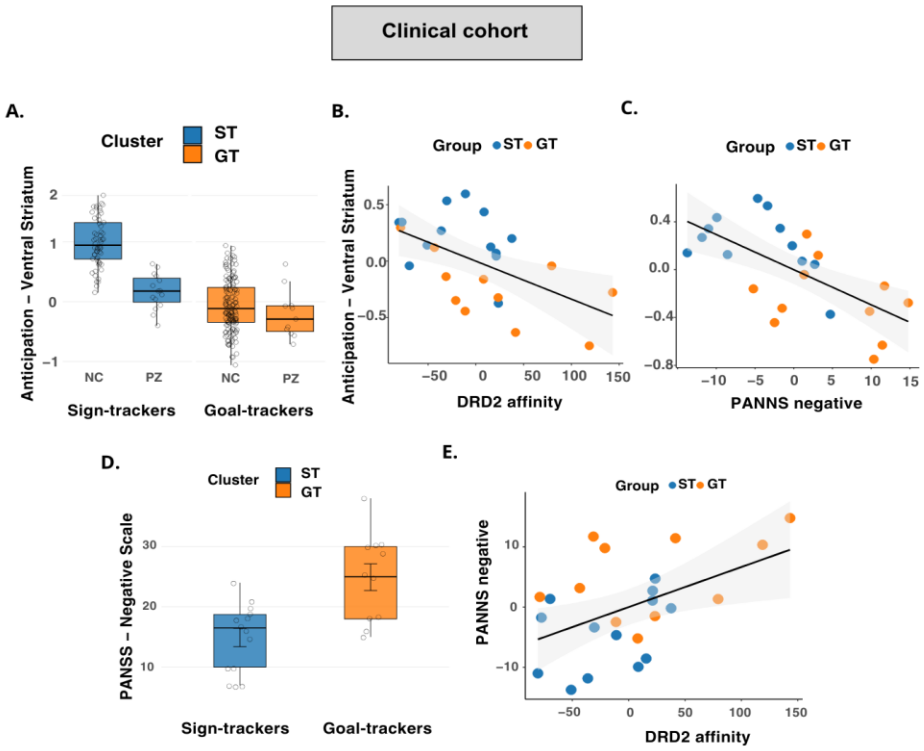

Figure III. 16 Individual-level analyses in the Clinical cohort. (**A**) BOLD signal in the VStr during anticipation, separately illustrated for neurotypical controls (NC) from the Replication cohort and patients (PZ) for both ST and GT. (**B**) Scatterplot shows negative correlation between the BOLD signal in the VStr during anticipation and DRD2 affinity. (**C**) Scatterplot showing negative correlation between the BOLD signal in the VStr during anticipation and the negative symptoms measured with the PANSS. (**D**) Boxplots representing ST vs. GT mean difference in relation to the negative symptoms. (**E**) Scatterplot depicting the positive correlation between negative symptoms and DRD2 affinity. *Note.* All scatter plots display residualized data, adjusted for age and sex.

3.4. Discussion

Drawing from a well-characterized behavioral phenotype in rodents, we translated the ST/GT framework to humans using neuroimaging to investigate individual differences in reward processing and antipsychotics with high affinity for D_2 receptor modulation. While previous human studies have explored ST/GT analogs with different approaches (Colaizzi et al., 2020; Heck et al., 2025), our study offers a translationally relevant validation by leveraging the well-established MID task to identify reproducible imaging phenotypes analogous to ST-like and GT-like profiles across four independent cohorts.

Our results highlight four main findings. First, the ST/GT model is highly replicable in humans even without new learning, as BOLD signal clustering consistently delineated ST-like and GT-like profiles across all cohorts. Second, ST attribute high incentive salience to reward-anticipation cues, showing amplified ventral striatal and whole-network responses during anticipation, whereas GT show preferential activation during outcome processing. Third, D_2 blockade selectively attenuated anticipatory BOLD activity in ST, with a corresponding decrease in self-reported Energy, while leaving GT's neural and subjective responses unchanged. Finally, in patients with SCZ and FEP, those treated with strong dopaminergic antagonists displayed the most blunted ventral-striatal anticipation signals and also had the highest negative-symptom burden. Together, these findings demonstrate a translational pathway from preclinical models to healthy humans and ultimately to clinical populations, establishing a neuroimaging biomarker of incentive salience, and offer a phenotype-driven strategy for tailoring D_2-targeted interventions.

3.4.1. Clustering on the brain activity during the MID task is a reliable approach to translate the ST/GT animal model into humans

A critical concern in human ST/GT research is ensuring alignment between the definitions of sign-tracking behavior used in humans and those established in animal models. One major problem with human experimental paradigms is the difficulty in implementing a direct measure of behavioral approach, which in animal studies is the time of latency before physical contact or the number of lever presses. Consequently, in most paradigms, the approach is inferred from proximal measures (Saunders & Robinson, 2013), specifically, the visual attention allocated to a CS. This perspective sounds legitimate in the sense that sign-tracking is akin to the "attentional biases" observed, for instance, in response to drug cues compared to neutral cues in humans (Anselme & Robinson, 2020; Field & Cox, 2008).

Heck et al. (2025) classified the human ST/GT literature into four main categories: (1) physical PCA tasks, (2) virtual PCA tasks, either alone or embedded within PIT paradigms, (3) value-modulated attentional capture (VMAC) tasks, with or without eye-tracking, and (4) alternative tasks, which include game-like paradigms, neuroimaging, and real-world or unclassified approaches.

For instance, in VMAC tasks, approach behavior is inferred from attentional metrics, particularly visual fixation on reward-associated cues (Saunders & Robinson, 2013). Most studies exploring ST behavior in humans have relied on computerized reaction-time tasks (Albertella et al., 2019a, 2019b; Liu et al., 2021), some of which incorporating eye-tracking indices to enhance measurement precision (Garofalo & di Pellegrino, 2015; Schad et al., 2020). Another procedure sometimes used in the ST/GT literature is the PIT task, which induces a cue-reward association through Pavlovian conditioning to alter instrumental responses to the cue (Cartoni et al., 2016; Garbusow et al., 2022). In altering the motivational influence of the reward predictive cue, PIT has

much in common with sign-tracking (Garbusow et al., 2014; Garofalo & di Pellegrino, 2015). The PIT effect has also previously been used to quantify incentive salience attribution (Sebold et al., 2021; Wassum et al., 2013), although the propensity to attribute incentive salience to the CS is not directly tested. The standard PIT procedure, therefore, cannot easily identify ST/GT phenotypes.

Achieving consensus on experimental parameters and data analysis strategies is crucial for advancing human ST/GT research. The ST/GT model in rats is derived from the observation of a characteristic behavior emerging in a Pavlovian conditioning paradigm, which implies a learning component, supported by ventral striatal DA signaling relevant for sign-tracking behavior, but not for goal-tracking (Flagel et al., 2007, 2008, 2009). However, in a real context, stimuli also have predictive value of reward without forming new associations; rather, it may be acquired through direct experience, learning, or explicit instructions. In humans, the MID task allows us to study both cue-reactivity and reward-related responses, starting from the investigation of the brain activity underlying reward sensitivity—and not the opposite—thus evaluating the value of reward itself in two distinct temporal phases: the anticipation phase, when cues trigger reward prediction, and the outcome phase, when feedback is evaluated (Knutson et al., 2001).

In this study, we introduce a new method for examining cue-reward sensitivity in humans by combining the MID task with the ST/GT framework to address their respective limitations. Since the MID is an instrumental task, it allows us to assess the "reward value" and the "incentive salience" of reward cues without the variation caused by individual learning styles in the ST/GT model. Therefore, applying the ST/GT framework enabled us to complement the phase-based approach of the MID task, thereby emphasizing individual differences in cue-reward learning and integrating both reward anticipation and outcome phases.

Far from adding more heterogeneity to the set of approaches used for studying reward and ST/GT behavior, this study sheds light on a novel approach that may be preferable to the more

complex paradigms reported above, which are both difficult to implement and replicate results throughout the literature.

3.4.2. Hierarchical k-means clustering identifies reproducible ST- and GT-like imaging phenotype across diverse cohorts

In the ST/GT animal model, individual variability in the tendency to develop sign-tracking behavior allowed researchers to categorize ST and GT (see Flagel et al., 2007, 2008, 2009), suggesting that these categories reflect stable phenotypes (Gillis & Morrison, 2019).

When translating this framework in humans, an additional source of variability is the construct validity, and consequently, the categorization of participants into ST and GT groups (Heck et al., 2025). The nature of the "response bias score" from which the categories were defined is highly heterogeneous. In general, few studies included a third (intermediate) category for individuals whose behavior alternated between sign- and goal-tracking. For each category of paradigms—physical and virtual PCA, VMAC, and alternative—Heck et al. (2025) presented studies that implemented participant profile classification. In both physical and virtual PCA studies, individuals are grouped according to their tendency to develop sign-tracking, established from the number and latency of physical contacts with (or gazes to) the CS versus the reward location (e.g., Colaizzi et al., 2023; Cope et al., 2023; Dinu et al., 2024). While contemporary PCA paradigms typically classify individuals into ST, GT, or INT phenotypes (Colaizzi et al., 2023; Colom, 2023, Chapter 9; Cope et al., 2023; Joyner, 2019), earlier studies often did not. Virtual PCA task studies have all identified different groups of participants based on their sign- or goal-tracking behavior, although Lehner et al. (2017) used alternative terminology to ST and GT, based on their predominant eye fixation on the CS or the reward location (e.g., "high eye index" vs. "low eye index"; Lehner et al., 2017). VMAC tasks, by contrast, usually assess attentional bias on a continuous scale and do not typically lead to a ST/GT

categorization (exceptions include Duckworth et al., 2022; Liu et al., 2021). Among studies using alternative tasks, such as computerized or real-world tasks, only a few categorized individuals into ST or GT groups (Nelson et al., 2022; Schembre et al., 2016; Versace et al., 2016). Finally, other studies define behaviors akin to sign- and goal-tracking in human participants in very vague ways, such as "abnormal responses to reward cues" (Wardle et al., 2018) and "markers of impulsivity and attentional problems focusing on individual differences" (Serrano-Barroso et al., 2022).

In our study, hierarchical k-means clustering was used to stratify healthy individuals based on their brain activity in the reward-related network, considering an arbitrary number of clusters (i.e., k = 3). It was determined based on the existing literature on ST/GT behavior in animal models (Flagel et al., 2008, 2009, 2011; Flagel & Robinson, 2017; Saunders & Robinson, 2013; for a review, see Felix and Flagel, 2024). Thus, in the Discovery cohort, we included the INT group, as it more accurately reflects the population distribution without differentiating the two extreme phenotypes. However, a model with k=2 for the Discovery cohort still identified ST and GT, supporting the idea that a third group, which shows a mixture of brain patterns, better reflects the heterogeneity of the population, leading to more precise stratification of individuals (Figures III.S11-12). In fact, for the subsequent analyses, only the ST and GT clusters were used, despite a decrease in sample size. For the generalization of the algorithm in the Replication cohort, we applied hierarchical k-means clustering with k = 3 and k = 2 solutions. We compared the predicted clusters, derived from the projections of the Discovery centroids, with the true clusters obtained from a hierarchical k-means algorithm. The k = 3 solution continued to categorize individuals into groups reflecting ST, INT, and GT in the Replication cohort (Figure III.S12). However, for the purpose of obtaining more precise profiles, we focused on the projection of the two centroids from the Discovery cohort as the main analysis. Therefore, in the other three cohorts, we used the two Discovery centroids corresponding only to ST and GT, excluding the INT group to reduce noise. We found stable and reproducible ST and GT

imaging phenotypes across all cohorts, consistent with previous studies in humans, with the INT group being excluded from the analyses (Colom, 2023, Chapter 9; Schad et al., 2020; Dinu et al., 2024). To further demonstrate that selecting k = 2 would not have altered the results, even in the subsequent analysis, Figure III.S14 compares the results from neuropsychological analyses performed with k = 2 and k = 3 with INT dropped.

This result demonstrates that the initial subdivision into three groups can be generalized to diverse populations, including a clinical one. This interpretation is supported by the high accuracy (balanced accuracy > 78.7% across three cohorts), as demonstrated in the confusion matrices, which compare predicted clusters from the Discovery algorithm with actual data-driven cluster membership. Inconsistencies in conceptual interpretation and operationalization contribute to heterogeneity across studies, potentially limiting the validity of cross-species comparisons.

So far, the various human paradigms have greatly differed in their attempts to identify distinct ST and GT phenotypes. Our results indicate that stratification into three groups enhances ecological validity as it better defines the reward-cue responsivity of the real population. Thus, future studies may rely on the stability of our classification to address the categorization problem in their research on ST and GT.

3.4.3. Neurocognitive characterization of ST and GT in humans is consistent with animal models

On a neurobiological level, several studies support the distinction between "top-down" and "bottom-up" pathways responsible of the GT and ST behavior, respectively (Flagel et al., 2011; Haight, Fuller, Fraser, & Flagel, 2017; Sarter & Phillips, 2018). Whereas GT are thought to rely primarily on "top-down" cortical mechanisms to guide their goal-directed behaviors, ST are believed to have enhanced

"bottom-up" processing, as a function of increased activity in subcortical regions, including the striatum, amygdala, midline thalamus, and hypothalamus.

This different level of control may be reflected in cognitive functioning. GT are considered to have greater top-down attentional control than ST, and it has been postulated that a "deficit" in this top-down control contributes to the sign-tracking phenotype (Sarter & Phillips, 2018). In support of this view, GT perform better than ST on high-demand cortical control tasks, including those associated with impulse control (Flagel et al., 2010; Lovic et al, 2011) and sustained attention (Paolone et al., 2013).

In our study, we have found either that ST outperformed GT in reward-related tasks and executive functions, or no differences in personality and attentional resources emerged between the two imaging phenotypes. These results can be interpreted in light of various studies, including both animal and human research.

For example, Enkel et al. (2019) demonstrated that ST and GT rats show task-dependent differences in executive function: although ST rats modestly outperformed GT rats on a place-to-cue set-shifting task, they exhibited poorer inhibitory control in a conditional responding paradigm and delayed alternation, illustrating that inferences about "top-down" control hinge critically on the specific cognitive assay (Enkel et al., 2019). In contrast, our human ST showed superior performance on an executive functions test—fewer perseverative and non-perseverative errors—and more efficient contingency switching during reward-based tasks. These outcomes are more akin to the enhanced set-shifting seen in ST rodents but at odds with their deficits in response inhibition. Rather than attributing these discrepancies solely to cross-species neurobiology, we suggest they reflect divergent task demands and contextual factors: human tests often emphasize explicit rule abstraction and verbal strategies, whereas rodent tasks rely on implicit learning and Pavlovian–instrumental transfer (Pitchers et al., 2018). Finally, ST rats have been found to be more prone to maladaptive behaviors, i.e., addiction,

due to their lack of control and inhibition of reward-related stimuli, even if it is worth noting that these rats were in a food-deprived condition. Clearly, human studies cannot replicate this status, although a different deprived condition, such as money, may be relevant in the study of variability in the attribution of incentive salience to rewarding stimuli related to inappropriate behaviors.

Therefore, recognizing how task architecture shapes observed phenotypes is essential for translating ST/GT frameworks across species and potentially selecting human paradigms that more closely model the underlying neurocognitive mechanisms. In humans, research on ST/GT differences in cognition is currently in its infancy, and future studies should more precisely characterize cognitive profiles to clarify both the parallels and differences from animal models.

3.4.4. Uncovering the striatal involvement in the reward processing in humans

The involvement of the striatum in reward processing has been widely shown (Cardinal et al., 2002; Delgado, 2007; Wang et al., 2016). Findings have highlighted its role in the computation of reward prediction errors (O'Doherty et al., 2003) and in the anticipation of reward delivery (Knutson et al., 2001).

Studies conducted in rodents (Corbit & Balleine, 2011; Berridge & Kringelbach, 2015)—and more recently in humans (Pool, Tord et al., 2022)—provided evidence of dissociable contributions of the VStr subregions to the motivational and the hedonic component of the affective processing of reward. The motivational mechanisms determine how much effort an individual mobilizes to obtain a reward, whereas the hedonic mechanisms determine the degree of pleasure an individual experiences during reward consumption (Berridge & Kringelbach, 2015). Animal studies have outlined the critical role of the VStr, particularly in pavlovian-triggered motivation (Corbit & Balleine, 2011; Wassum et al., 2013), using the PIT paradigm. A growing number of studies have extended animal findings

regarding the role of the VStr in Pavlovian motivation to humans by adapting the PIT paradigm for use in an MRI scanner (Talmi et al., 2008; Chen et al., 2020; Schad et al., 2020).

These studies have played a pivotal role in the formulation of the incentive salience hypothesis, which postulates that under some particular circumstances the motivational (i.e., "wanting") and hedonic (i.e., "liking") components of the affective processing of reward can be dissociated, thereby making organisms work for a reward that they will not necessarily like once obtained—a key feature of compulsive reward-seeking behaviors such as addiction (Berridge & Kringelbach, 2015). The ST/GT model offers a unique opportunity to dissociate the neurobiological mechanisms underlying the motivational and hedonic components of reward processing. Indeed, our study employed the MID task to investigate the anticipation and outcome phases separately within a ST/GT framework. Current findings indicate that persistent striatal activation in ST occurs even without a learning paradigm, suggesting that heightened ventral striatal responses in ST do not depend on active learning. This aligns with findings from studies on cue-induced eating, where ST/GT-like neurobehavioral profiles were identified, even when cue-reward contingencies were explicitly instructed rather than learned through trial and error, yet still predicted the amount of food consumed (Versace et al., 2019). Previous studies have linked ST to model-free learning and insensitivity to outcome devaluation (Morrison et al., 2015; Smedley and Smith, 2018). Moreover, a recent human study has shown that ST rely on reward prediction error signals (Schad et al., 2020). However, these studies involved reinforcement-based tasks, leaving open the question of whether striatal engagement in the ST/GT model requires learning.

By showing that ST maintain elevated striatal activity in a task where reinforcement contingencies do not affect behavior, this study rules out a dependence on trial-by-trial learning. These findings reinforce the idea that ST behavior is characterized by sustained striatal engagement.

3.4.5. Beyond the dopamine role in "wanting"

The implementation of pharmacological challenges, combined with high-resolution fMRI protocols, is a useful tool for scrutinizing the modulation of reward-related activity (Ogawa, Lee, Kay, & Tank, 1990). Several studies in healthy human volunteers have reported alterations in the reward-related BOLD signal in both subcortical and cortical regions following the pharmacological manipulations which alter DA signaling, such as the dopamine D2 receptor antagonists amisulpride (Admon et al., 2017), haloperidol (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), olanzapine (Abler, Erk, & Walter, 2007), lurasidone (Wolke et al., 2019) and sulpiride (Diederen et al., 2017); the indirect dopamine agonists dextroamphetamine (Knutson et al., 2004) and amphetamine (O'Daly et al., 2014); and the dopamine precursor L-DOPA (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006).

For instance, olanzapine is a DA receptor antagonist with high affinity for D1-, D2-, and D4-receptors, and, in line with other atypical antipsychotics, appears to selectively affect the mesolimbic and mesocortical but not the dorsal striatal dopaminergic system (Bymaster et al., 1997; Moore et al., 1997). It induces an elevation of extracellular DA levels in vivo in the NAc and PFC in rats (Koch et al. 2004; Zhang et al. 2000). In healthy subjects, Abler et al. (2007) discovered that reward-related brain activation in the VStr, ACC, and inferior frontal cortex was reduced on olanzapine compared to placebo. The antipsychotic effects of olanzapine have been attributed to the dampening of increased mesolimbic dopaminergic stimulation, which reduces the assignment of salience to important stimuli (Kapur 2003; Spitzer 1997).

Knutson et al. (2004) and Abler et al. (2007) demonstrated the acute effects of dopaminergic drugs on brain activation related to a reward paradigm. For instance, using amphetamine as an indirect DA agonist in a within-subject, double-blind, placebo-controlled study, Knutson et al. (2004) found that the effect of psychostimulants on brain activity, measured with fMRI, was most evident in the VStr and psychologically enhanced mood and attention. These findings align with those of Abler et

al. (2007), who used olanzapine, and were replicated in a recent study by Hawkins et al. (2021), in which risperidone suppressed reward-anticipatory signals in striatal areas. One explanation for the reduction in the striatal signal during reward anticipation following risperidone administration, as proposed by Hawkins and colleagues (2021), is that D2 blockade on striatal postsynaptic membranes results in the suppression of the postsynaptic potential and the associated BOLD signal (Menon et al., 2007; Schott et al., 2008). Therefore, that manipulating dopaminergic signaling in humans—via D_2-receptor targeting compounds—alters BOLD responses during reward anticipation is by now well established.

In our study, the novelty lies in the fact that, in humans, ST exhibit a selective response and dose-dependent attenuation in BOLD activity after challenge with antipsychotics that have high affinity for D_2 receptor antagonism, accompanied by a parallel decrease in self-reported Energy. The effect on subjective reports is especially noteworthy for the potential of fMRI-mediated ST-GT discrimination in predicting treatment response and side effects of dopaminergic agents. This specific dopaminergic sensitivity in ST mirrors findings from rodent work showing that cue-driven incentive salience in ST rats depends critically on DA transmission (Flagel et al., 2009). By contrast, GT rats acquire cue–reward associations with comparatively less dopaminergic involvement. Together, these results underscore the central role of DA in modulating incentive salience (Robinson & Berridge, 1993; Heinz, 2002) and demonstrate that ST individuals, who attribute greater motivational value than GT to cues, are particularly sensitive to compounds with high affinity for the D_2 receptor antagonism at both the neural and subjective levels.

Observations of ST-specific dopaminergic sensitivity in modulating incentive salience may inform our understanding of conditions described by the hypothesis of aberrant salience in psychosis. This hypothesis proposes that dysregulated DA signaling leads to inappropriate significance attribution in psychosis, with antipsychotics ameliorating psychotic symptoms by normalizing these

signals (Kapur, 2003). Importantly, identifying ST-like salience attribution in drug-naïve patients could guide personalized treatment: those with pronounced cue-driven BOLD responses might benefit most from D_2-targeted agents, whereas individuals with GT-like profiles might require alternative strategies to avoid unnecessary DA blockade. Such phenotype-driven approaches hold promise for optimizing therapeutic efficacy and minimizing side effects in disorders marked by dysregulated salience processing, including SCZ. Indeed, the results in our patient cohort suggest that GT might be more likely to develop secondary negative or more severe primary negative symptoms when treated with strong D_2 blockade.

3.4.6. Blunted activity in the ventral striatum of patients with schizophrenia

In our Clinical cohort, hierarchical k-means clustering again identified ST and GT profiles despite the administration of antipsychotic medications. Patients exhibited reduced striatal activity during reward anticipation compared to HC, in line with a consistent body of evidence demonstrating blunted ventral striatal activation during anticipation in pharmacological challenges (Hawkins et al., 2021), as well as in unmedicated patients with SCZ performing the MID (Esslinger et al., 2012; Nielsen et al., 2012). Indeed, studies of drug-naïve and chronic patients (Juckel et al., 2006; Grimm et al., 2014; Esslinger et al., 2012), along with first-episode individuals, and unaffected first-degree relatives (Grimm et al., 2014; Lancaster et al., 2016) are important to disentangle intrinsic neurobiological traits from medication-induced alterations.

One compelling hypothesis for the reduced striatal reward activation in SCZ suggests that the increased baseline DA tone in the striatum in patients (Abi-Dargham et al., 2000; Fusar-Poli & Meyer-Lindenberg, 2013; Howes et al., 2012) results in a dampening of the phasic signals associated with

rewarding stimuli or reward-predicting cues. Consequently, these critical signals fail to elicit significant changes in the BOLD response and are effectively attenuated (Heinz & Schlagenhauf, 2010).

In our Clinical cohort, ST-like patients showed markedly blunted ventral-striatal responses during reward anticipation compared with healthy ST. In contrast, GT-like patients did not exhibit the same blunted activation profile compared to healthy GT individuals. This finding is noteworthy because it suggests that the classic "blunted" striatal response, often described in SCZ (Radua et al., 2015), is not a universal characteristic of the disorder, but rather a variable feature that our ST/GT framework may help explain. This perspective reinforces the argument that we should expect different treatment responses based on phenotypes.

Furthermore, while conventional antipsychotics primarily target positive symptoms (Kapur, 2003), the fact that our ST-like/GT-like phenotyping was linked to an effect on negative symptom burden—a dimension often resistant to treatment—is particularly impressive. Particularly, GT-like patients, with a lower striatal activation than ST-like patients, also showed more severe negative symptoms. In fact, reduced sensitivity of DA receptors during neuroleptic treatment has been associated with increased negative symptoms that may be related to a dysfunctional response to rewarding stimuli in patients (Schmidt et al. 2001). Furthermore, a consistent body of evidence showed alterations in the reward-related signaling in the striatal and cortical activity (Juckel et al., 2006; Simon et al., 2010), with ventral striatal hypoactivation predicting negative symptoms, and medial PFC deficits correlating with positive symptoms (Radua et al., 2015; Oldham et al., 2018). This result highlights the power of a phenotype-driven strategy to uncover a neurobiological link to an important symptom dimension, even if it was not the primary target of the medication.

When dealing with blunted striatal activity, disentangling trait-like salience attribution from drug-induced effects is challenging. According to the aberrant salience attribution theory, hyperdopaminergic dysregulation within the midbrain-striatal-prefrontal pathway is relevant for the

onset of psychosis (Grimm et al., 2014), and it causes specific impairments in the brain's response to motivationally relevant cues (Esslinger et al., 2012). Further studies supported this model, showing increased DA synthesis in the striatum also in healthy first-degree relatives of schizophrenic patients (Fusar-Poli & Meyer-Lindenberg, 2013; Huttunen, Heinimaa, Svirskis et al., 2008), suggesting that deficits in the striatum during anticipation are not merely consequences of chronic illness or pharmacological treatment but may reflect a core neurobiological vulnerability associated with SCZ. In our results, on the one hand, patients classified as GT based on their BOLD signature may truly reflect goal-tracking phenotypes; on the other, long-term antipsychotic exposure itself remodels striatal function. Benjamin et al. (2022) reported that chronic D_2 antagonism drives extensive transcriptional changes in the human caudate, including alterations in *DRD2*-related pathways. Additionally, Osugo et al. (2025) showed that seven days of D_2/D_3 receptor blockade induce both attenuated striatal BOLD responses to reward and emergent motivational deficits reminiscent of the negative symptom dimension. By contrast, our single-dose challenge in healthy ST yielded immediate, dose-dependent reductions in anticipatory BOLD activity and subjective Energy—effects that do not result from underlying molecular adaptations. Together, these findings imply a two-phase impact of D_2 antagonists: an acute suppression of DA-mediated incentive salience in ST, followed by a slower, gene-expression–mediated dampening of reward processing under chronic treatment.

The blunted anticipatory BOLD response observed in our GT-like patients, in conjunction with their higher negative symptom burden and a positive correlation between D_2 affinity and negative symptoms, suggests that the observed neural and behavioral deficits may be a consequence of long-term effects of drugs with high affinity for the D_2 receptor blockade. This highlights the potential of the ST/GT framework to provide an objective biomarker for treatment-emergent side effects, helping to distinguish core pathology from drug-induced symptoms. This is further supported by the

significant negative correlation between D_2 affinity and ventral striatal activity, which points to a drug-mechanism link in this population.

3.4.7. The ST/GT model as a promising framework for studying impulsivity

The ST/GT framework offers a valuable lens for investigating individual differences in personality traits and their neurobiological underpinnings. In rodent models, the ST-GT distinction has proven to be a powerful predictor of impulsivity and vulnerability to addiction (Flagel & Robinson, 2017). Preclinical evidence has consistently reported higher impulsivity and cue-reactivity in ST compared to GT, suggesting that individual differences in reward-driven behavior may partly reflect stable personality dimensions (Colaizzi et al., 2020; Flagel et al., 2009; Flagel et al., 2008, 2010; Garofalo & di Pellegrino, 2015; Robinson & Flagel, 2009; Tomie et al., 2000).

Impulsivity is a heterogeneous construct (Dalley & Robbins, 2017) but can be broadly defined as a lack of behavioral inhibition that may lead to unplanned, premature, and often risky behaviors (Belin et al., 2016; De Wit, 2009). Importantly, the manifestation of impulsivity in ST and GT rats varies depending on the task. In tasks measuring impulsive action, like the Differential Reinforcement of Low Rates of Responding, which requires animals to wait a set time before responding to earn a reward, ST rats exhibit greater impulsivity (Lovic et al., 2011; Campus et al., 2016). They struggle to control their responses and tend to press a lever quickly, demonstrating challenges in inhibiting responses—even when doing so is counterproductive—suggesting a motor control deficit driven by the strong pull of reward cues (Flagel et al., 2010; Lovic et al., 2011). Conversely, in tasks like delay discounting, which assess impulsive choice by having animals select between a smaller immediate reward and a larger delayed reward, GT rats tend to make more impulsive choices, opting for the smaller, immediate reward. The better decision of ST rats indicates that their impulsivity is more

related to a heightened sensitivity to reward cues that override motor control, rather than a general difficulty in valuing future outcomes.

In humans, translating this model has yielded promising but inconsistent findings. Some studies report higher self-reported impulsivity and cue reactivity in ST-like individuals (Garofalo & di Pellegrino, 2015), while others, including Study 2 of this thesis, found no significant personality differences. In Study 2, impulsivity, measured by the SURPS, showed no difference between ST and GT imaging phenotypes. This null result may be due to the limited sensitivity of self-report questionnaires in capturing state-dependent or task-specific expressions of impulsivity, or the influence of the fMRI paradigm design. The MID task is adjusted based on individual performance by altering the time interval needed for a correct response. More broadly, current human research suffers from methodological heterogeneity in defining ST/GT and largely relies on self-report assessments rather than behavioral measures of impulsivity.

Future studies should combine behavioral and computational measures of impulsivity, such as tasks that evaluate response inhibition or delay discounting, with neuroimaging measures of both task activity and connectivity in the prefrontal cortex and striatum. These multimodal methods could clarify whether the neural signatures differentiating ST and GT reflect distinct mechanisms of impulse control and how motivational salience is attributed. Ultimately, this research could help explain how trait impulsivity interacts with cue-driven reward sensitivity, advancing a mechanistic understanding of vulnerability to impulse regulation disorders, including addiction and ADHD (Paolone et al., 2013).

3.4.8. Limitations

Despite the robust replicability of our clustering approach and the clear correspondence between our behavioral and neural findings and those from animal models, several limitations warrant consideration. First, the absence of subjective reports on whether participants were aware of the

treatment condition during each session, although the observed ST-specific reduction in energy suggests that drug effects were behaviorally meaningful despite this uncertainty. Second, the lack of observational or clinically-informed measures to supplement the self-reported data in the pharmacological cohort is a limitation. Third, our modest patient sample, which included both first-episode and chronic schizophrenia cases, limits our ability to definitively disentangle trait-like salience attribution from treatment effects. The challenge of separating intrinsic phenotypes from drug-induced brain adaptations is particularly relevant here, as long-term antipsychotic exposure can remodel striatal function. This leaves open the compelling question of whether patients exhibiting a GT-like activation profile are those receiving ongoing therapy, rather than reflecting a distinct underlying phenotype.

3.4.9. Conclusion

In summary, this study identifies robust ST and GT imaging phenotypes in humans, which mirror those observed in rodent models of reward processing. ST exhibit heightened anticipatory striatal activity and DA-sensitive cue reactivity, while GT preferentially engage in the reward outcome phase. These profiles were consistently detected across tasks/cohorts, drug conditions, and clinical status, linking them to both subjective drug response and symptom severity in psychosis. By translating preclinical models into a replicable human framework, our findings provide a neuroimaging biomarker of incentive salience and offer a promising lead for fMRI-based clinical predictions.

Chapter 4

4. General Discussion

The primary objective of this doctoral thesis was to demonstrate that shifting from group-level averages to an inter-individual perspective provides a more accurate and predictive framework for understanding brain-behavior relationships. This work also aimed to demonstrate that utilizing ML techniques to characterize individual variability enables the detection of meaningful neurobiological patterns that are typically obscured by population averaging.

To address this overarching objective, the research was conducted across two distinct case studies, namely memory and reward.

Study 1. The first study aimed to overcome the limitations of standard anatomical atlases by mapping individual-specific patterns of cortico-thalamic connectivity and memory task-related activity. The goal was to test whether mapping the individual spatial overlay of thalamic subdivisions in the FPNs and DMN could improve the prediction of interindividual differences in memory performance. In this study, a core methodological framework has been established by adopting a novel individual-level approach to studying memory, specifically the association between rs-fMRI and task-based brain measures extracted from non-normalized spatial maps and individual memory performance. This analysis is grounded in the hypothesis that individuals may recruit different neural resources supporting their learning abilities. Interindividual variability in thalamic recruitment before, during, and after memory encoding has been analyzed using a native individual-space approach. Rigorously validated across two independent datasets, this method achieves substantial interindividual variability in the corticothalamic functioning involving the anterior and medial thalamus in associative memory encoding. This study shows that different individuals may recruit distinct corticothalamic resources, which, in turn, are associated with learning success.

Study 2. Once the methodological approach has been validated in the first study as a proof of concept, the second study aimed to apply a neuroimaging-based stratification approach to the domain of reward processing. By applying a clustering algorithm—an unsupervised technique used to group similar data points together into clusters based on their characteristics or attributes—to four different independent cohorts, the aim was to investigate the interindividual variability in the response to rewarding stimuli. The primary aim was to identify human neuroimaging analogs of ST/GT phenotypes during reward anticipation and outcome, without requiring learning. The second aim was to test whether these phenotypes predict differential neural and subjective responses to dopaminergic modulation. Third, their clinical translational value was evaluated by assessing whether ST/GT-like profiles can be detected in patients with psychosis and linked to clinical features. The strong generalizability of these phenotypes across four different cohorts, exhibiting a distinct pattern of brain activation during reward anticipation and outcome, suggests that this model is highly reliable and stable. Individuals can be stratified based on their attribution of incentive salience to rewarding stimuli, even in a clinical population. The ultimate goal of this study was to put forward a translational neuroimaging biomarker to develop personalized treatment strategies for psychiatric disorders involving reward dysfunction.

Together, the studies developed in this thesis advocate for the methodological shift proposed in the General Introduction: prioritizing interindividual variability in addition to seeking commonalities across individuals by averaging participant data. Rather than treating between-subject differences as 'noise' to discard through averaging individual data (Kanai & Rees, 2011), it posits that cognitive functions, such as memory and reward, reflect different brain functioning between individuals. Demonstrating that everyone is unique in cognitive processing sheds light on the heterogeneity that characterizes clinical conditions.

A Transformative Shift Toward Personalized and Biologically Informed Models

In the General Introduction, the challenges inherent to studying the brain-behavior relationship are outlined, moving from group- to individual-level approaches. In neuropsychology, the traditional focus has been on how specific brain regions relate to their functions, creating a structured cognitive taxonomy. The initial scientific evidence for the relationship between brain structure and behavior in humans came from lesion studies (Genon et al., 2018). These studies demonstrated associations between localized brain lesions and specific behavioral deficits, thereby supporting the notion that certain brain regions are associated with specific behavioral functions. For example, the relationship between the hippocampus and episodic memory was highlighted in the foundational work of Scoville and Milner (1957). However, the initial one-to-one mapping concept is primarily species-specific and does not account for the unique differences between individuals. Thus, in recent years, it has been revisited in favor of a many-to-many perspective, especially following the rise of functional and structural neuroimaging studies in healthy populations.

The existence of interindividual variability, meaning that each person has a unique model of functioning within the physiologically viable limits, is not immediately apparent and must be empirically demonstrated. For instance, if a medication is strongly linked to a behavior—such as falling asleep—it is reasonable to assume that the majority of people will exhibit that behavior, although the dosage required may vary from person to person. Similarly, when solving a cognitive task, there is an interval of variability in which different individuals use their brains in different ways. Unlike studies that assume a one-to-one mapping between cognitive functions and brain structure applicable to everyone, this perspective allows for the identification of subgroups within the larger population. In recent decades, cognitive neuroscience has sought to investigate the intriguing variability of human cognition. This exploration holds significant importance, especially in clinical contexts, as it unveils the potential for a multitude of innovative solutions tailored to individual needs. This perspective

acknowledges the multivariate nature of brain-behavior relationships and supports generalization, depicting the relationships between brain and behavioral data using latent dimensions of interindividual variability (Genon et al., 2022).

The next section discusses the methods and findings of Studies 1 and 2 in light of their significance to the field using individualized approaches. The final sections will address the limitations of the studies and propose future research directions emerging from these results.

4.1. From Averaging to Individualized Prediction: Evidence from Two Case Studies

As outlined in the General Introduction, heterogeneity in task-based and FC has gained more attention in recent years. Building on the latest advancements in the interindividual variability approach in cognitive neuroscience, this thesis presents two case studies that investigate learning and memory, as well as the affective domain of reward, through this pioneering perspective. This approach is evident not only in the theoretical framework of the studies but also in the statistical methods employed. Collectively, findings from this thesis highlight the pivotal importance of individual differences in enriching our understanding of cognition and behavior, making a compelling case for a transformative shift toward personalized and biologically informed models in the field of neuroscience.

Traditional group-average studies largely rely on spatial normalization to account for morphological differences between individuals (Friston et al., 1995), but normalization does not address the anatomical and functional heterogeneity often present within a given brain area (Fedorenko and Blank, 2020), potentially leading to inaccurate group estimates (Fedorenko, 2021). To tackle the challenges associated with the variability in brain function patterns, researchers have developed analytical methods and study designs aimed at separating functional signals from specific neuroanatomical structures, such as hyperalignment and localizer tasks. Specifically, hyperalignment

is a high-dimensional mapping technique that projects individual brain activity patterns into a common representational space based on functional response profiles rather than anatomical coordinates, effectively rotating individual voxel spaces to maximize cross-subject correlation (Haxby et al., 2011; Guntupalli et al., 2016; Feilong et al., 2018). Complementarily, localizer tasks are brief functional paradigms designed to identify specific ROIs in individual participants based on their unique functional response to a cognitive contrast (e.g., faces vs. objects). This 'functional ROI' approach ensures that the analyzed signals originate from areas that are functionally equivalent across individuals, regardless of their precise macro-anatomical location (Kanwisher et al., 1997; Fedorenko, 2021). While hyperalignment optimizes the common representational space across the whole cortex, localizer tasks provide a targeted solution for identifying functionally homologous nodes in individual brains (Fedorenko, 2021).

In this paragraph, I will outline the Methods and Results of Studies 1 and 2, which support the preference for an interindividual variability approach in the study of cognitive neuroscience.

Are individual-level analyses better than group-level analyses? A proof-of-concept study

Study 1 addresses a long-standing challenge in memory research: the functional heterogeneity of the thalamus. Traditional group-level studies frequently fail to find robust thalamic associations with memory performance, often attributing this to the small size of the nuclei or low SNR. However, these findings suggest a different interpretation: the information loss occurs during spatial normalization.

By conducting specular analyses at both group and individual levels, this study diverges from traditional MNI-based pipelines (Friston et al., 1995) and highlights their limitations. While group-level one-sample t-tests identified significant clusters within the DMN and FPN networks, the signals

extracted from these averages were not behaviorally predictive. This converges with recent critiques suggesting that group-level functional ROIs often overfit the mean and fail to generalize to individuals (Fedorenko, 2021). In contrast, by mapping thalamic recruitment in native space, this study unveiled associations between medial and posterior subdivisions and memory indices that were previously invisible at the group level. This aligns with and extends the work of Gordon et al. (2017) and Keller et al. (2023), who argued that individual functional topography is more behaviorally relevant than macro-anatomical overlap. Notably, our finding that the proportion of significant voxels (a spatially informed metric) is more sensitive than FC parallels the findings of Kong et al. (2019), suggesting that the precise "where" of an activation is often more informative than the magnitude of the signal itself.

Stratifying individuals to capture differences across a range of variability in the reward processing

In Study 2, a multi-level approach has been developed to assess ST/GT phenotypes in humans without requiring new learning, while still allowing for the distinct evaluation of neural substrates underlying reward anticipation and outcomes. Replicated in three independent cohorts of healthy individuals and one cohort of patients with psychosis, ST profiles showed greater activation during anticipation and blunting in the outcome, whereas GT exhibited the opposite pattern. The results converge with preclinical models (Flagel et al., 2011; Robinson & Berridge, 2008), confirming that DA-dependent incentive salience (ST) and top-down goal-directed behavior (GT) are neurobiologically distinct and reproducible phenotypes in humans. Crucially, while previous human studies have reported inconsistent results due to complex behavioral designs, the current use of unsupervised clustering on imaging data provides a more robust solution for identifying these profiles.

ST/GT phenotypes in humans unveil stable, biologically based individual differences in incentive salience attribution. The identification of distinct subgroups, each displaying unique patterns of neural activity across different cohorts, underscores the power of individualized approaches to

enrich our understanding of interindividual variability. This insight holds significant promise for clinical applications, as ST and GT phenotypes have been observed even among patients receiving medication. The specific dopaminergic sensitivity linked to ST enhances the appeal of this model as a transformative framework for personalized interventions in psychiatric disorders associated with reward dysfunction. This framework not only aligns with varying antipsychotic treatment patterns but also sheds light on the severity of negative symptoms, paving the way for more effective and tailored therapeutic strategies.

To summarize, this thesis highlights the crucial role of individual-level analyses in capturing the variability in brain function. By demonstrating that individualized approaches can uncover predictive patterns where group-level analyses fail, this work supports a broader movement in the field. This movement seeks to replace the "one-size-fits-all" anatomical template with a framework that respects the unique functional landscape of each brain. Ultimately, integrating inter-individual variability into the core of neuroscientific inquiry is not just a methodological choice, but a requirement for the development of truly translational and clinically relevant models.

4.2. Limitations

Since the research question of this thesis aims to ambitiously shift the focus from "group average" to "individual variability" in the study of cognitive neuroscience research, the limitations discussed here relate to this goal, considering the findings from Studies 1 and 2. Therefore, the limitations in this General Discussion should not be viewed as weaknesses, but rather as critical reflections that help better frame the scope of the results.

Generalizability across the functional hierarchy and cognitive domains. A limitation of this research concerns the extent to which the findings in memory and reward can be generalized to the

whole functional architecture of the human brain. Although these two domains were chosen for their significant variability between individuals and clinical relevance, they mainly reflect specific levels within the cognitive hierarchy, primarily involving associative and limbic circuits. Evidence indicates that interindividual variability is not uniformly distributed across the cortex; for instance, higher-order heteromodal regions (such as the frontoparietal networks studied here) typically show significantly higher topographical variance than primary sensory and motor regions (Mueller et al., 2013; Noble et al., 2019). Therefore, the success of the individualized mapping and stratification approaches demonstrated in this thesis may be domain-dependent. It remains to be established whether these precision models retain the same predictive power when applied to more "constrained" systems, such as basic perception or motor control, where the signal-to-noise ratio and anatomical-functional alignment may differ. Future research should aim to integrate these individualized frameworks across a broader spectrum of the cognitive taxonomy to determine whether a "universal" model of variability can be achieved or whether different systems require specialized analytical pipelines.

MRI spatial resolution. A central challenge in understanding inter-individual differences is the inherent trade-off between fMRI spatial resolution and the mesoscopic scale of brain organization. While neuroimaging methods are often optimized for the cerebral cortex, where functional signals span larger surface areas (Hermes et al., 2012; Raimondo et al., 2021), applying these techniques to small subcortical structures poses significant methodological challenges. This thesis, particularly Study 1, highlights that standard fMRI—typically with voxel sizes of 2-3mm—may produce 'mixed signals' or partial volume effects when examining small, densely packed regions, such as thalamic nuclei (Pergola et al., 2013b; Phillips et al., 2021). This issue is exacerbated by substantial anatomical variability in subcortical regions across individuals (Johansen-Berg et al., 2005; Georgescu et al., 2020), which may not be fully captured by personalized mapping when the signal-to-noise ratio is inadequate. This 'scaling problem' indicates that the accuracy of inter-individual models is limited spatially: we can

reliably map differences in broad cortical networks, but resolving the detailed functional topography of subcortical 'hubs' remains constrained by current hardware and sequence limitations. Although the results are consistent across datasets, the advent of ultra-high-field MRI (7T and higher) and advanced signal enhancement techniques (Alkemade et al., 2022) suggests that current findings may be conservative estimates of variability. Future studies should investigate whether principles of individual functional organization in the cortex can be directly applied to subcortical regions or if significant modifications to spatial acquisition methods are necessary to accurately distinguish biological signals from measurement noise.

Different MRI acquisition parameters and varying experimental paradigms. In Study 1, two fMRI datasets were acquired at different resolutions. Notably, those findings displayed remarkable consistency across both datasets. However, recent breakthroughs in enhancing signal quality within subcortical regions, particularly the thalamic nuclei (Alkemade et al., 2022), prompt us to approach these results with caution. Therefore, while I am encouraged by the robustness of our findings, I recommend careful interpretation, considering the evolving landscape of neuroimaging technology. In Study 2, a similar issue was observed not only between different cohorts but also within the Discovery cohort, which had fMRI data collected from eight different sites. Although a leave-site-out cross-validation was performed, confirming the model's validity despite variations in acquisition parameters, this aspect should still be taken into account.

Therefore, the experimental paradigms also differed from one another. In Study 1, despite the distinct retrieval task demands of the two memory tasks—one involving single items and the other picture pairs—the remarkable similarity in brain patterns observed during retrieval suggests that both demands activate thalamocortical pathways linked to associative memory as captured by fMRI (Vatansever et al., 2021; Ambrus, 2024). In contrast, the experimental tasks in Study 2 included different variations, such as changes in the shape of the cue during anticipation and the introduction

of a punishment condition. While these differences are notable, they are relatively minor in the broader context of the research, as the same contrast — i.e., Reward vs. No-Reward — could be created across all cohorts. One limitation to acknowledge is that although the MID task offers an established and reliable method for examining reward-related activity, it only captures a small part of the ecological complexity of real-world motivational behavior and cue–reward interactions. These studies support the idea that commonalities across different data collection modes suggest that common processes emerge regardless of parameter differences.

Limited sample size. A critical constraint shared by both studies—and pervasive across the neuroimaging field—is determining an adequate sample size to ensure sufficient statistical power. Historically, many fMRI studies have relied on heuristics or practical constraints rather than formal, *a priori* power analyses to determine cohort size (Poldrack et al., 2017). This omission is often due to the inherent difficulty in pre-specifying effect sizes for complex, multivariate individual-level mapping, for which established benchmarks are still lacking (Cremers et al., 2017; Button et al., 2013).

The implications of this limitation are twofold. In Study 1, clustering analysis was performed on a large cohort of 148 participants, showing notable variability in membership across different brain hemispheres. This variability indicates that there are clear differences in how memory processes are lateralized in the brain. To better understand the interaction between the two hemispheres in supporting various memory functions, it is important to replicate these clustering results in larger and more diverse groups. This is particularly relevant as recent large-scale studies have demonstrated that brain-wide associations with behavioral traits often require thousands of subjects to produce reproducible effect sizes (Marek et al., 2022).

In Study 2, the clustering was validated in a cohort of 890 individuals, providing a more robust dataset for analysis. However, it is essential to acknowledge that the sample size used for specific analyses, particularly those related to the Pharmacological Challenge and the Clinical cohort—which

included only 34 participants each—creates a notable limitation. It should be emphasized that a sample of 34 participants is generally considered within the standard range for detecting drug effects in a pharmacological fMRI study, particularly given its demanding longitudinal design involving three separate assessment timepoints. However, power limitations become a more significant concern during the transition from whole-sample to subgroup-level analyses. When individuals are stratified into ST/GT imaging phenotypes and misidentified subjects are removed from the analysis, the resulting effective sample size is significantly reduced. This reduction limits the statistical robustness of group comparisons and increases the risk of inflated effect sizes. Therefore, while these findings provide crucial proof of concept, they should be interpreted as preliminary, highlighting the need for larger, longitudinal, multi-site trials to validate the stability of these neurobiological subtypes. Additionally, including patients at different illness stages—both FEP and chronic SCZ—restricts the ability to disentangle medication-induced effects from inherent, trait-like neural characteristics.

Representativeness of the sample and generalizability (The "WEIRD" Problem). Research on individual differences requires very large and diverse samples to accurately distinguish between what is "normal" and what is "unique". As a direct consequence of the previous limitations, it comes to mind that many neuroimaging studies are biased toward "WEIRD" populations—Western, Educated, Industrialized, Rich, and Democratic. Using a homogeneous sample, such as college students, can lead to an underestimation of the true variability and may not reflect the broader population or other cultures. Henrich et al. (2010), along with Marek et al. (2022)—famous for their work on BWAS (Brain-Wide Association Studies)—highlight that studying individual differences effectively requires thousands of participants to prevent findings that cannot be replicated.

Cross-sectional design and longitudinal implications for the stability of neural phenotypes. A crucial consideration in interpreting the current findings in the literature on "brain fingerprinting" is

whether the identified neural profiles represent a stable characteristic of the individual (a trait) or are influenced by state-dependent factors and individual strategies. While the clustering approaches identified reproducible and behaviorally meaningful profiles—such as distinct corticothalamic recruitment patterns in memory (Study 1) and sign-tracker/goal-tracker–like phenotypes in reward processing (Study 2)—the cross-sectional designs prevent definitive conclusions about their temporal stability. Thus, we suggest framing our findings cautiously.

In both studies, our findings shed light on the complex brain-behavior relationships among individuals by using a tailored approach that emphasizes the rich variation in biological processes across individuals, shaping diverse phenotypes. Although these studies identify individual patterns, without longitudinal data (repeated scans of the same subjects over months or years), it is difficult to determine how permanent these "signatures" are. Gratton et al. (2018) demonstrated that functional networks are dominated by individual traits, but cognitive states can still alter their topography, complicating prediction. Indeed, neural signatures captured with fMRI may reflect a combination of relatively stable predispositions (e.g., dopaminergic sensitivity, network architecture) and more flexible, context-dependent processes, including task strategies, motivational states, or recent learning history.

Moreover, this limitation is particularly relevant in light of preclinical evidence showing that sign- and goal-tracking behaviors, although partially trait-like, can be modulated by environmental contingencies and pharmacological manipulations (e.g., dopaminergic antagonists selectively altering incentive salience attribution in sign-trackers but not goal-trackers). Consistently, the Pharmacological challenge results in the present thesis suggest that dopaminergic modulation can transiently reshape reward-related neural responses, highlighting the dynamic interplay between trait vulnerability and state-dependent neuromodulation. Similarly, in the memory domain, variability in thalamocortical engagement may reflect not only stable individual differences but also differences in cognitive

strategies (e.g., encoding vs. retrieval-oriented approaches) that are not directly controlled in the current design.

In this context, it is crucial to distinguish between replicability and test-retest reliability. The findings presented in this thesis demonstrate high replicability: the ST/GT phenotypes in Study 2, for instance, were consistently identified across independent cohorts with high accuracy (e.g., up to 94% in the Discovery cohort), suggesting that these features are robust across different populations, or even within a leave-site-out cross-validation framework. However, as noted by Demidenko et al. (2024), high group-level replicability does not automatically guarantee high individual-level reliability. While the former confirms the existence of a phenomenon, the latter—often measured via Intraclass Correlation Coefficients—quantifies the stability of an individual's score across different time points.

Recent multiverse analyses of the MID task have highlighted that individual BOLD estimates often exhibit low to moderate test-retest reliability, even when group-level effects are robust (Demidenko et al., 2024; Holiga et al., 2018). Several factors may account for this discrepancy. First, biological maturation plays a significant role, particularly in studies involving adolescents or young adults, where the dopaminergic system undergoes substantial remodeling (Baranger et al., 2021). These maturational changes, alongside fluctuations in dopamine expression, can naturally alter the neural substrates of reward and memory over extended periods. Second, environmental factors and psychosocial stressors can impact the brain's "state", introducing noise into reliability estimates if the assessment window is too broad. As suggested by Baranger et al. (2021), while reward-related activation may show longitudinal stability at the group level, individual-specific signatures may be more volatile due to the brain's plasticity in response to experience.

Consequently, future perspectives for these studies should more directly address the trait–state distinction through longitudinal and multi-session designs, test–retest reliability assessments, and experimental manipulations of strategy and context. Future studies may deepen understanding of these

insights and consider intraindividual designs. Some developmental perspectives are discussed in the Section 4.3.3. below. This concept of intraindividual generalizability emphasizes that for a "brain fingerprint" to be clinically useful, it must ensure that these results are not only consistent across the population but also remain stable over time within each individual, or change in predictable ways. Combining fMRI with computational modeling, ecological momentary assessment, and pharmacological interventions may further help to characterize the stability, plasticity, and causal mechanisms underlying these profiles. Ultimately, understanding the tension between stable traits and dynamic states will be essential for moving toward personalized models that are not only consistent across populations but truly reliable for the individual. By exploring these personal dimensions more thoroughly, we can enhance our understanding of the complex connection between human behavior and biology.

Summary. Together, these limitations do not weaken the potential of studying interindividual variability and the developed methodologies; instead, they define their current limits. It is essential to acknowledge these challenges in order to enhance the application of personalized approaches in understanding brain-behavior relationships, particularly during clinical transitions. The next section highlights opportunities for future research to complement and expand these studies.

4.3. Future perspectives

Future research should aim to further disentangle the dynamic contributions of thalamic subdivisions to memory and the differences in reward processing by integrating multimodal imaging, longitudinal designs, and computational modeling. For instance, complementing these studies with longitudinal designs may be useful for better understanding the stability over time and potential flexibility of these patterns and phenotypes, especially in response to medication or behavioral interventions. For example, investigating how thalamic activation patterns change during development and aging, as well

as how they respond to cognitive or neuromodulatory interventions, could aid in the development of personalized strategies for cognitive improvement.

Moreover, expanding Study 1 to include clinical populations—such as those with SCZ or thalamic stroke—will be essential for validating the translational relevance of thalamic biomarkers. Likewise, expanding the validation of ST/GT phenotypes to larger and diagnostically diverse clinical populations, such as individuals with affective or substance use disorders, may further elucidate the transdiagnostic relevance of these profiles. Identifying subgroups of patients with unique thalamocortical connectivity patterns could inform personalized psychiatry, enabling targeted treatments based on individual neurocognitive profiles. Combining multiple methods, including resting-state and task-based FC, dopaminergic PET scans, and genetic and transcriptomic data, would provide a deeper insight into the biological factors that cause differences in reward processing between individuals. Also, future studies should integrate fine-grained, ecologically valid assessments of affective states, motivation, and reward sensitivity—such as momentary experience sampling, laboratory-based decision-making paradigms, or computational modeling of reinforcement learning—to better characterize the psychological signatures corresponding to these neurobiological phenotypes. Thus, integrating task-based and rs-fMRI with diffusion tractography, electrophysiological measures, biological and clinical phenotyping may identify distinct patient subgroups characterized by specific connectivity signatures and cognitive profiles.

Ultimately, advancing this framework toward clinical translation will necessitate the development of predictive models, potentially utilizing ML, to categorize individuals based on cue-driven versus outcome-driven reward processing tendencies. Such personalized, phenotype-based approaches could ultimately inform precision psychiatry strategies that aim to tailor treatment selection and optimize therapeutic outcomes based on individual neurocognitive signatures. A deeper understanding of individual variability in cortico-thalamic dynamics and reward processing will pave

the way for personalized diagnostics and targeted therapies in neuropsychiatric and neurodegenerative disorders.

4.3.1. A clinical translation: Toward biologically informed models of schizophrenia

As outlined in this Chapter and previously in Paragraph 2.4.6 of Chapter 2, the individualization of neural signatures in the domains of memory (Study 1) and reward (Study 2) provides a powerful framework for clinical translation, particularly in complex disorders such as schizophrenia (SCZ). Traditionally, SCZ has been treated as a monolithic clinical entity, yet it is characterized by profound heterogeneity in both cognitive performance and affective symptoms. The findings of this thesis suggest that mapping inter-individual variability is not merely a descriptive exercise, but a prerequisite for identifying the biological underpinnings of this clinical variability.

In the domain of memory (Study 1), the delineation of cortico-thalamic pathways associated with memory performance within our healthy sample offers a valuable framework for understanding the heterogeneous cognitive profiles characteristic of SCZ, a disorder marked by disrupted thalamo-prefrontal connectivity (Passiatore et al., 2023; Pergola et al., 2015; Giraldo-Chica & Woodward, 2017). In this context, memory dysfunction—encompassing working memory, episodic encoding, and retrieval—should not be regarded as a secondary manifestation of psychosis but as a core feature that strongly predicts functional outcomes and everyday adaptive capacity (Green et al., 2019). Identifying individual-specific patterns of thalamic engagement enables categorization of patients based on their neurofunctional profiles and tailoring cognitive remediation or neuromodulatory interventions accordingly. Importantly, targeting thalamo-prefrontal through neuromodulatory and behavioral interventions offers a promising path for translational research. For example, individuals showing reduced medial thalamic recruitment or impaired fronto-thalamic coupling could particularly benefit from interventions aimed at enhancing network synchrony, such as transcranial magnetic stimulation

(TMS), transcranial direct current stimulation (tDCS), or real-time fMRI neurofeedback (Sitaram et al., 2021; Lo et al., 2022). Additionally, combining these techniques with cognitive remediation or memory training may enhance neural plasticity and improve functional outcomes.

Parallel to this, the application of the Sign-Tracker/Goal-Tracker (ST/GT) model to a Clinical cohort enabled the stratification of patients based on their reward-processing profiles. Patients with SCZ and first-episode psychosis predominantly exhibited a GT-like imaging phenotype. Crucially, this phenotype was significantly associated with antipsychotic treatments characterized by stronger D2 blockade and was linked to a higher burden of negative symptoms. This suggests that the GT-like profile in SCZ may serve as a neurobiological marker of reduced incentive salience attribution, potentially exacerbated by medication-induced dopamine antagonism. By parsing the "noise" of clinical heterogeneity into these meaningful, dopamine-dependent signatures, we can gain a better understanding of why certain patients remain refractory to standard treatments or experience more severe motivational deficits.

The integration of these findings suggests a transformative path for personalized interventions. Categorizing patients based on their unique neurofunctional profiles could enable targeted treatment. While Study 1 informs more neuromodulatory strategies, Study 2 guides the tailoring of pharmacological and behavioral approaches. Identifying a patient's ST/GT profile could guide the choice of antipsychotics—prioritizing agents with different receptor profiles for GT-like patients to avoid further blunting of reward circuits—or inform "cue-exposure" therapies designed to de-escalate the salience of environmental stimuli for ST-like individuals to manage their high sensitivity to reward-predictive cues.

To further refine this translational framework, future perspectives must involve the integration of multimodal imaging data. While fMRI provides high spatial resolution of functional activity, it remains a proxy for underlying neurochemistry. The incorporation of PET data would be a critical

step to directly assess dopaminergic availability, such as presynaptic dopamine synthesis capacity or D2/D3 receptor occupancy. Combining PET and fMRI data would confirm whether the ST/GT phenotypes identified in this thesis are truly driven by the same molecular mechanisms established in preclinical models.

Ultimately, developing a general framework for inter-individual variability enables us to determine whether neurofunctional signatures—whether linked to memory or reward—serve as stable trait markers of vulnerability or as reversible, state-dependent changes. By combining personalized task-based fMRI with clinical phenotyping, we can shift from an average diagnostic approach to a precision psychiatry model that accounts for each patient's distinct neurobiological profile.

4.3.2. Precision imaging as a complementary approach to study interindividual variability

Although this thesis concentrated on interindividual variability, there are additional complementary approaches that could enrich our understanding and help us generalize these findings at the intra-individual level. Precision imaging reverses the traditional strategy of maximizing sample size by instead concentrating on the intensive scanning of single individuals (Mattoni et al., 2025). This approach aligns with our objectives, aiming to fully characterize the neural functioning of each individual and to eliminate inaccuracies that can arise from group averages and standardized anatomical atlases (Fedorenko, 2021; Seitzman et al., 2019). Precision imaging studies scan individuals for longer periods of time, typically on multiple occasions providing hours of data, providing valuable insights into brain functioning, such as the amount of data necessary for reliable FC estimation, consistency of individual brain network differences, and the relationship between task activation and resting state networks (Flournoy et al., 2024, Gordon et al., 2017, Gratton et al., 2018).

Precision imaging is an ideal approach to studying the individual for several reasons. First, longer scanning times and multi-echo imaging reduce noise in functional signals to provide more reliable characterization of individual-level brain functioning (Gordon et al., 2017, Lynch et al., 2020). Second, precision imaging addresses anatomical heterogeneity by localizing brain activation or networks using BOLD patterns that are specific to each individual. For instance, functional localizer tasks can increase the precision of individual-level BOLD responses (e.g., Fedorenko et al., 2010), and precision functional mapping can better characterize individual-level network functioning (Gordon, Laumann, Gilmore, et al., 2017). Third, repeated sampling enables the longitudinal assessment of dynamic changes that are central to an idiographic approach (the study of the individual) in science. For example, Lynch et al. (2024) found that expansion of the salience network was a reliable between-person correlate of depression across multiple precision imaging and consortia datasets. However, in an individual scanned 62 times across 1.5 years, salience network expansion was not associated with depressive symptoms. Instead, FC between the NAc and ACC predicted depressive symptoms, suggesting that depression and FC patterns have distinct interindividual and intraindividual relationships.

Moreover, trial-level analysis (Chen et al., 2021) of task-based studies can provide an intensive time series of neural data and behavioral responses. These time series can enable intraindividual study and be more practical than numerous repeated scans for certain research questions. For example, trial-by-trial modeling has been used to study various processes, including reinforcement learning, working memory, and cognitive control (Cho et al., 2022; Colas et al., 2022; Pessoa et al., 2002). In Study 2, it was relevant for assessing the attribution of incentive salience throughout the fMRI experiment, which distinguished ST from GT. Precision imaging enables deep characterization of the individual over time and context, providing an ideal study design for intraindividual processes such as cognitive mechanisms, development, or within-person brain-behavior associations.

Due to the advantages of individual-level precision, several commentaries have highlighted the use of precision imaging for clinical translation (Gratton et al., 2020; Kraus et al., 2023; Laumann et al., 2023). Building on findings from Study 1, exploring how thalamic recruitment patterns evolve across development and aging, and how they respond to cognitive or neuromodulatory interventions, could inform precision strategies for cognitive enhancement. Future investigations should pursue longitudinal and multimodal designs to examine how thalamic functional organization evolves across illness stages and treatment trajectories.

Ultimately, this evidence advances the broader goal of precision psychiatry and cognitive neuroscience. Moving beyond group-level averages to intra-individual studies, interventions are tailored to each person's unique neurobiological architecture. By integrating high-resolution imaging, personalized analytical pipelines, and behaviorally relevant biomarkers, future studies may bridge the gap between basic neuroscience and clinical application, paving the way for more personalized and effective strategies to improve cognition and quality of life in SCZ.

4.3.3. Toward a developmental framework of individual variability

A longitudinal perspective would be particularly valuable to clarify how individual differences in both cognitive and affective neural responses evolve over time and interact with cognitive maturation, environmental exposure, and clinical status. The individual functional topographies identified in this thesis—whether related to corticothalamic memory pathways (Study 1) or sign- and goal-tracking reward profiles (Study 2)—should not be viewed as static snapshots, but as products of a dynamic maturation involving the refinement of subcortical-cortical communication.

In Study 1, mapping thalamocortical circuits showed that individual connectivity patterns predict associative memory performance. From an evolutionary perspective, it is well established that FC between the thalamus and associative areas (such as the PFC and the DMN) increases significantly

during adolescence, shifting from a "sensory-centric" configuration to a complex hierarchical organization (Fair et al., 2010; Hwang et al., 2022). Longitudinal studies suggest that strengthening the thalamo-prefrontal tracts is a critical prerequisite for the development of working memory and cognitive flexibility (Petersen et al., 2024). Therefore, the inter-individual variability observed in the thalamic mapping of Study 1 likely reflects different "stages" of maturation of these loops, where greater cortical integration marks the shift toward more efficient and stable mnemonic strategies (Alcauter et al., 2014).

This trajectory converges with observations regarding reward processing (Study 2). Recent studies exploring sign- and goal-tracking in children and adolescents using adaptations of the PCA paradigm show that, despite variability, there is a developmental bias toward sign-tracking behaviors during childhood, likely due to the still-maturing PFC and limited top-down control (Colaizzi et al., 2023; Joyner et al., 2017; Flagel et al., 2008; Casey et al., 2000). Neuroimaging studies show that children with fewer sign-tracking behaviors exhibit greater activation in the inferior parietal lobe during reward anticipation, indicating better attentional control. Individuals with more sign-tracking behavior exhibit stronger amygdala activation (Colaizzi et al., 2023), indicating a shift from cue-driven, subcortical behavior to goal-oriented, cortical control. This bias could be explained by changes in reward sensitivity that occur during adolescence, leading to impulsive decisions and risk-taking behaviors, such as substance use (Jollans et al., 2016; Karoly et al., 2015; Weiland et al., 2013).

A longitudinal approach encompassing both domains would allow us to determine the growth curves of the neural fingerprints identified in this thesis, which emerge precisely from the consolidation of these thalamo- and striato-cortical connections. Charting how ST/GT-related neural patterns evolve from adolescence into adulthood might unveil whether persistent sign-tracking tendencies indicate a delay in the maturation of executive control or represent a unique neurobiological route associated with increased risk for psychopathology. For instance, persistent "immature"

thalamic connectivity or abnormal developmental trajectories of reward-related circuits could represent early vulnerability markers for neuropsychiatric disorders such as SCZ, Alzheimer's disease, and substance use disorders (Ma et al., 2016; Scheres et al., 2007; Marek et al., 2022). Ultimately, charting these individual trajectories would allow us to distinguish between developmental flexibility and permanent biological traits, providing a predictive framework for mental health across the lifespan.

4.4. Memory and reward: a combination of theoretical concepts to study interindividual variability

In this thesis, memory and reward processing have been studied separately, adopting a common methodological approach to investigate individual differences that characterize these processes.

However, numerous reports have investigated the relationship between memory and reward. On the one hand, encoding and retaining valuable information in long-term memory is a highly adaptive process, as it enables us to shape our future behavior through improved choices and actions (Shohamy & Adcock, 2010). On the other hand, reward processing promotes associative learning by helping organisms learn which actions and stimuli are beneficial, as it links actions to positive outcomes. Over time, cues associated with a reward become as motivating as the reward itself, allowing for more efficient pursuit of resources (Bhanji & Delgado, 2014). However, as also demonstrated by our Study 2, the ST/GT framework suggests that individual differences exist in the motivational value assigned to cues anticipating reward, with only ST showing incentive salience to reward cues, mediated by striatal dopamine signaling.

In real life, there are often situations where information that is important to learn is presented simultaneously with less important information. Information designated as important is more likely to be remembered successfully than unimportant one (Ngo et al. 2019; Cohen et al. 2022). This tendency

benefits memory function, implying that the brain probably has several mechanisms that enhance memory for important items.

In laboratory and functional neuroimaging studies of memory, researchers have investigated the effects of importance by manipulating the value of items to be remembered and presenting reinforcers during learning, such as offering monetary rewards (Adcock et al., 2006; Shigemune et al., 2014) or arbitrary point values (Castel, 2008). They found that this association elicits an increased activity in the mesolimbic dopaminergic system during successful memory encoding when a relatively high reward value is anticipated (Shohamy & Adcock, 2010). This dopaminergic response (Wolosin et al., 2012), which modulates neuroplasticity (Düzel et al., 2010) and thereby increases the memory strength of the studied items (Lisman et al., 2011; Lisman & Grace, 2005), plays a key role in prioritizing memories (Wittmann et al., 2008), presumably via enhanced encoding (Knowlton and Castel, 2022) and increased consolidation (Bialleck et al., 2011). Dopamine modulates plasticity in the hippocampus, a key structure for episodic memory (Lisman & Grace, 2005), determining which information enters long-term memory via the hippocampal-VTA loop (Murty et al., 2011). However, a more recent study (Cohen et al., 2020) found that when anticipating the possibility of gaining an extrinsic reward, activation of the dopaminergic reward system improves memory storage in a non-selective manner, as it strengthens memory for both the important information and for irrelevant aspects of the situation.

Recent studies have investigated how reward prediction errors (RPEs) contribute to the formation of memories of rewarding events. RPEs are quick signals that indicate the difference between expected and actual outcomes, also affecting DA levels. Specifically, DA increases when rewards exceed expectations and decreases when they fall short (Schultz & Dickinson, 2000; Schultz, Dayan, & Montague, 1997). Larger RPEs improve both recognition memory for items and learning from rewards, but these effects may involve different brain processes (Rouhani et al., 2018).

Additionally, recent studies have confirmed that memory performance is enhanced for neutral items that predict a later reward compared to those without a reward association (Lloyd & Nieuwenhuis, 2024; Nagel et al., 2024). This reward anticipation manipulation is thought to induce a heightened state of arousal (e.g., pupil size), leading to stronger encoding and greater subsequent memory success. (Lloyd & Nieuwenhuis, 2024).

Future research may benefit from integrating memory and reward while focusing on individual variability. For instance, it would be intriguing to investigate whether individuals differ in their memory performance based on varying levels of DA release within the mesolimbic circuit. It is possible that certain individuals, identified as ST and displaying unique cortico-thalamic functioning, may exhibit enhancements in memory performance. This hypothesis is supported by emerging evidence suggesting that individual differences in striatal-hippocampal coupling determine the extent to which reward cues prioritize memory encoding (Mason et al., 2017; Wimmer & Shohamy, 2012). By examining the interplay between these processes while focusing on individual variability, we can gain deeper insight into the complex mechanisms that underlie incremental learning and memory shaped by reward cues.

Beyond theoretical interest, this compelling approach rated approach holds profound implications for clinical prediction. By shifting the focus toward the interplay between these domains, we can move beyond descriptive biomarkers toward predictive models of clinical trajectory. In patients with SCZ, dealing with memory and reward difficulties, an "integrated biotype"—combining a patient's ST/GT profile with their individualized thalamic connectivity map—could serve as a powerful tool for clinical forecasting. It could lead to the discovery of novel translational biomarkers, aiding in the development of personalized interventions for this clinical population. For example, machine learning models trained on these multi-domain features could predict an individual's response to cognitive remediation or their likelihood of experiencing severe negative symptoms. This multi-layered characterization would enable clinicians to move away from reactive treatments and adopt a

proactive, personalized strategy that predicts and addresses the specific neurobiological needs of the patient before functional decline occurs. Such a leap in predictive accuracy is essential for the next generation of precision psychiatry, transforming how we diagnose and manage the complex heterogeneity of neuropsychiatric disorders.

Conclusions

These future directions aim to deepen our understanding of how brain activity relates to behavior, enriching our knowledge of these connections. In cognitive neuroscience, traditional methods that average brain data have been essential for developing general theories that identify which brain regions or systems support specific cognitive functions. However, it is essential to acknowledge that individuals differ in their perceptions, thoughts, and actions. This thesis highlights this variability, viewing it as a new way to examine brain function through the lens of individual differences in cognitive processes. This approach is not merely a methodological choice; it offers a perspective for understanding the neurobiological and functional mechanisms that link brain activity to traits. Improving these methods, expanding their use, and integrating diverse data types are essential for unlocking their full potential in both healthy and clinical contexts. By applying personalized methods alongside unsupervised ML techniques, we identified subgroups of individuals with distinct patterns of association between brain activity, functional connectivity, and behavior, such as memory (Study 1) and reward (Study 2).

Multi-modal and longitudinal studies are essential for ensuring intra-individual generalizability, helping researchers gain a detailed understanding of individual differences over time. The studies shown here not only lay a strong foundation for future research but also highlight the complexities of human behavior and mental processes. By combining different methods and examining trajectories over time, these studies could advance the fields of translational neuroscience and precision psychiatry.

Going forward, these efforts can enhance our ability to tailor interventions and treatments to meet each person's unique needs, ultimately leading to improved mental health outcomes.

Appendix A

A Supplementary Information (Study 1)

A.1. Comparison across connectivity metrics: FC vs. the proportion of significant voxels across subdivisions

With the hypothesis that FC could flatten the interindividual variability that is highlighted when using the spatial informed metric (H0: $\sigma^2A = \sigma^2B$ vs H1: $\sigma^2A > \sigma^2B$), we compare the connectivity metrics previously modeled, i.e., the proportion of significant voxels and the FC across the thalamic subdivisions during resting-state. Since the two metrics are on different scales and thus not directly comparable, we used the min-max scaling (de Amorim et al., 2023), which scales each distribution individually by its minimum and maximum absolute value, not centering the data, thereby preserving any sparsity. To test this hypothesis, we used Levene's tests to compare the variance between the distributions of FC and the proportion of significant voxels grouped by cohort, network, and hemisphere. Additionally, we examined the correlation between the two metrics across thalamic subdivisions to assess the discriminatory power of the measurement used through the Kendall-Tau correlation. Significance was set at $\alpha<0.05$.

A.2. The proportion of significant voxels may better discriminate the thalamic recruitment across subdivisions than FC

Levene's test comparing the differences in terms of variance distribution across metrics, i.e., the proportion of significant voxels and FC, demonstrated significant variance differences between the distribution of the proportion of voxels and FC in the right medial subdivision in the RFPN ($F(30)=$

3.68, p=0.03), whereas no differences were reported within the LFPN and DMN (p>0.05; Figure II.S1, panel A) in RUB.

In UNIBA, Levene's test showed significant variance differences between the distribution of the proportion of voxels and FC in the right medial subdivision in the RFPN (F(62)= 9.18, p=2e-03), and the DMN (F(68)= 17.07, p=5.02e-05) at baseline resting state, whereas statistical trends were reported for the comparison between the distribution of the proportion of voxels and FC in the left medial subdivision in the LFPN (F(54)= 2.1, p=0.07) and DMN (F(62)= 1.69, p=0.09; Figure II.S1, panel B).

In both cohorts, the correlation across metrics within and between thalamic subdivisions showed a higher correlation on average within FC signals during the baseline resting-state (averaged Kendall's rho ~0.2), compared to the proportion of significant voxels count that showed a lower correlation (averaged Kendall's rho ~0.09), as depicted in Figure II.S1, panel C.

These results suggest that the proportion of significant voxel count metrics may be more able to discriminate the contributions of individual thalamic subdivisions (Figure II.S1, panel C).

A.3 Data and code accessibility

Data from RUB and UNIBA cannot be shared at the individual level because of ethical restrictions based on the protocol approved by the relative institutional ethics committees to protect the privacy of the participants.

All R code, along with detailed group statistics needed to evaluate the conclusions in this work, is publicly available at: https://github.com/robertapassiatore/thalamus-memory.git. The THOMAS toolbox is publicly available at: https://github.com/thalamicseg/thomas_new.git. The Group ICA toolbox is publicly available at the following link: https://trendscenter.org/software/gift. The

Presentation® scenarios used for the associative memory tasks are available at the following link: https://doi.org/10.5281/zenodo.14374018.

A.4 Supplementary Figures

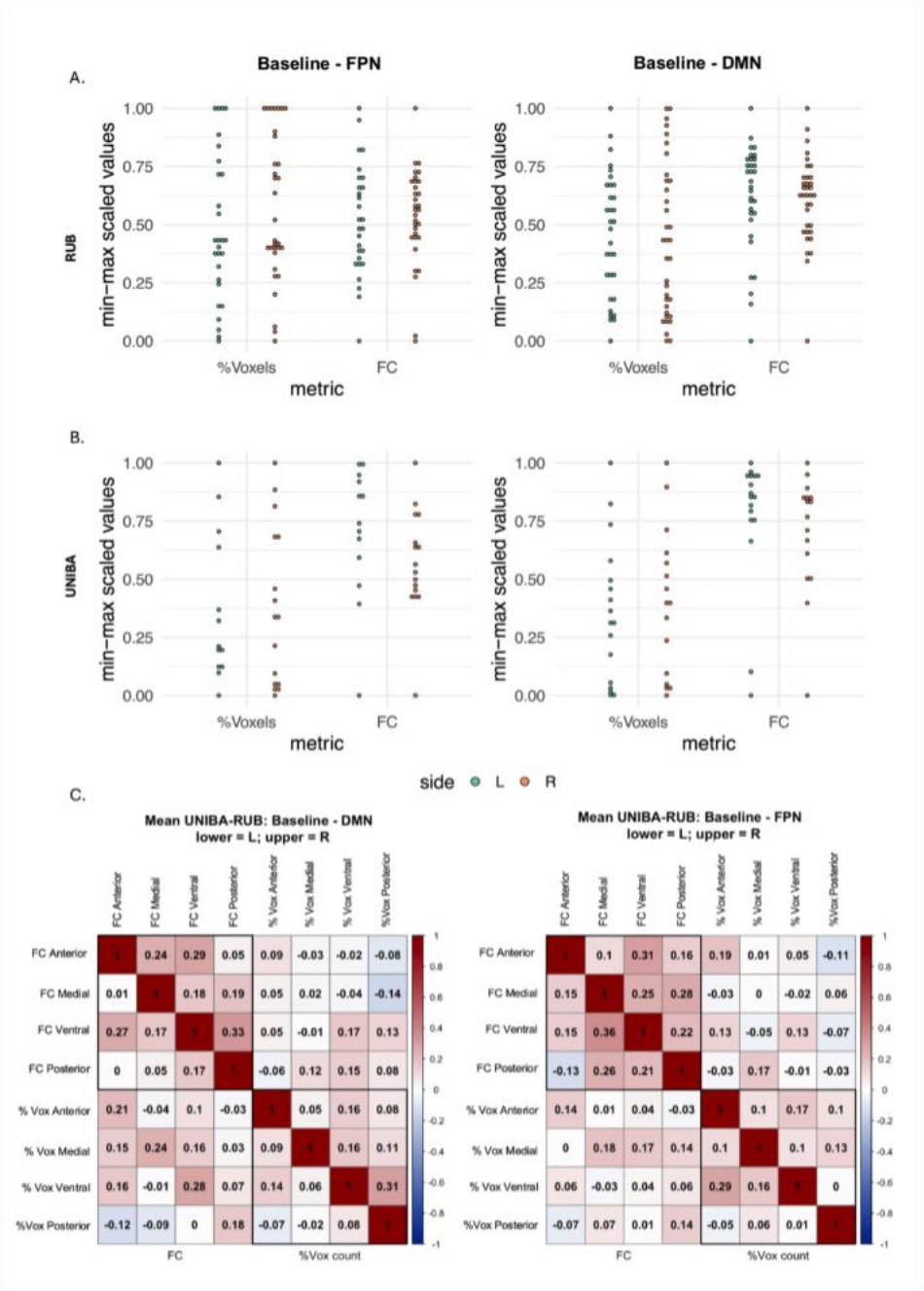

Figure II.S 1 Metric comparison between FC and the proportion of significant voxels. Dot plots show the differences in the distribution of the proportion of voxels and averaged FC extracted from the medial thalamic subdivisions within the left and right FPN and the DMN during baseline (**A**) in the RUB and (**B**)

UNIBA datasets. C, Averaged correlation matrices across datasets show Kendall's rho within and between the FC and the proportion of significant voxel count signals across thalamic subdivisions for the DMN and FPN at the baseline resting state. *Abbreviations*: R, right; L, left; DMN, default mode network; FPN, frontoparietal network; FC, functional connectivity; %Vox, proportion of significant voxels.

A.5 Supplementary Tables

		MNI coordinates				
Network	Side	*x*	*y*	*z*	Cluster extent (voxels)	*Z* value
RUB						
DMN	R	6	−24	8	11	5.71
DMN	L	−14	−16	8	19	5.67
LFPN	L	−4	−12	14	22	5.77
LFPN	L	−10	−4	−10	10	5.31
RFPN	R	6	−12	14	27	5.75
UNIBA						
DMN	R	12	−25	8	240	8.15
DMN	L	−9	−22	11	203	8.00
LFPN	L	−6	−13	11	119	8.05
RFPN	R	9	−16	14	213	7.84

Table II.S 1 Thalamic functional connectivity during resting state scans in the RUB and the UNIBA samples. Peak activations were detected through one-sample t-tests on the individual networks' spatial maps from Iraji et al. (2019). All peaks are shown at pFWE.

Scans	Network	Side	Anterior	Medial	Ventral	Posterior
Correlated voxels						
RUB						
Baseline	DMN	L	5	17	16	13
	LFPN	L	8	13	14	9
	RFPN	L	2	5	4	1
	DMN	R	3	15	8	8
	LFPN	R	1	2	7	2
	RFPN	R	10	16	13	12
Postencoding	DMN	L	1	13	12	10
	LFPN	L	9	8	10	13
	RFPN	L	3	4	6	2
	DMN	R	5	17	16	9
	LFPN	R	3	5	5	2
	RFPN	R	18	15	17	11
UNIBA						
Baseline	DMN	L	8	33	20	27
	LFPN	L	2	32	6	18
	RFPN	L	1	13	5	1
	DMN	R	17	36	30	30
	LFPN	R	0	5	2	1
	RFPN	R	16	37	21	14
Anticorrelated voxels						
RUB						
Baseline	DMN	L	4	1	1	4
	LFPN	L	4	3	2	7
	RFPN	L	2	7	5	9
	DMN	R	1	0	1	0
	LFPN	R	5	6	8	7
	RFPN	R	1	3	1	5
Postencoding	DMN	L	1	2	3	2
	LFPN	L	2	3	4	8
	RFPN	L	4	12	6	13
	DMN	R	1	2	2	1
	LFPN	R	7	10	12	12
	RFPN	R	3	0	3	6
UNIBA						
Baseline	DMN	L	5	2	1	3
	LFPN	L	2	7	2	3
	RFPN	L	2	2	4	32
	DMN	R	2	1	0	3
	LFPN	R	4	7	8	28
	RFPN	R	1	1	0	3

Table II.S 2 Number of non-zero observations in thalamic subdivision recruitment. The table 1277 reports the number of individuals presenting significant active correlated and anti-correlated voxels 1278 within resting state networks' spatial maps, i.e., DMN, LFPN, RFPN, grouped by hemisphere, 1279 scans, and thalamic subdivisions. We set the cutoff for sufficient observations to conduct two-part 1280 model regressions with memory performance at N≥10 observations. Abbreviations: DMN = default 1281 mode network; LFPN = left frontoparietal network; RFPN = right frontoparietal network; L = left; 1282 R= right.

Appendix B

B Supplementary Information (Study 2)

B.1 Participants

Discovery cohort

The IMAGEN study (Schumann et al., 2010) is a multisite, multinational longitudinal project that was carried out in eight European sites, namely Berlin, Dresden, Dublin, Hamburg, London, Mannheim, Nottingham, and Paris. The study involved adolescents recruited from high schools, and to obtain a diverse sample in terms of socio-economic status, emotional and cognitive development, private, state-funded, and special units were equally targeted. All participants underwent four types of assessment for different data domains (biological samples, brain imaging, clinical characteristics, and functioning data). These assessments were conducted at baseline (14 years of age) and followed up at 16, 19, and 22 years (i.e., follow-up 1, follow-up 2, and follow-up 3, respectively). Exclusion criteria included: the presence of overt neurological conditions such as epilepsy, brain tumors, bacterial infections of the CNS, muscular or myotonic dystrophy; cerebral trauma with loss of consciousness of more than 30 min; developmental issues such as major neurodevelopmental disorders, nutrition and metabolic diseases, uncorrectable visual or auditory deficits, IQ < 70; treatment for SCZ or bipolar disorder; presence of medical condition such as type 1 diabetes, systemic rheumatologic disorders, malignant tumors requiring chemotherapy, congenital heart defects or cardiac surgery, aneurysms; pre/perinatal issues such as maternal diabetes during pregnancy, excessive alcohol use of the mother during pregnancy, premature birth < 35 weeks and/or detached placental, hyperbilirubinemia requiring transfusion; MRI contraindication such as the presence of metal or electronic implants and severe claustrophobia. A detailed description of recruitment and research procedures has been previously reported (Schumann et al., 2010).

Clinical and behavioral assessments were performed using Psytools software (Delosis Ltd, London, UK) via its Internet-based platform. The battery of questionnaires and cognitive tasks was self-administered both at participants' homes and at neuroimaging facilities.

For the current study, the discovery sample was composed of a large sample of 1004 healthy young adults within Follow-up 3 (FU3) with MID (Knutson and Greer, 2008) task data available. Data from 114 participants were excluded because of poor behavioral performance at the MID task (accuracy below chance). No additional participants were excluded after the MRI quality check (Esteban et al., 2017), e.g., if the mean root-mean-square framewise displacement exceeded 0.5mm (see below in the fMRI Processing section). The final discovery sample included a total of 890 participants (age: M = 22.1, SD = 0.7 years; 454 females, 436 males). Demographic details grouped by sites are described in . Participants and their parents provided informed consent. The IMAGEN study protocol was approved by the KCL (King's College London) College Research Ethics Committee CREC/06/07-71 and by local ethics research committees at each site. Parents and adolescents gave written consent and verbal assent, respectively.

Replication cohort

We recruited an independent sample of 252 healthy adults who underwent an fMRI scan, performing a modified version of the MID task (Sportelli et al., 2024) and a neuropsychological and cognitive assessment. One subject was removed at his/her request. After removing six subjects for MRI data quality (e.g., excessive motion) and poor behavioral performance at the MID task (accuracy below chance), the final replication sample included 245 participants (age: M=25.7, SD=5.8; 141 F, 104 M). Participants signed an informed consent form complying with the Declaration of Helsinki after fully explaining all procedures approved by the local ethics committee. All subjects were Caucasians from the region of Apulia, Italy. Inclusion criteria were the absence of any lifetime psychiatric disorder, as

evaluated with the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders IV, NP (Non patient version), of any significant neurological or medical condition showed by clinical and magnetic resonance imaging evaluation, of history of head trauma with loss of consciousness, and of pharmacological treatment or drug abuse in the past year.

Pharmacological challenge cohort

Forty-two healthy right-handed males were pseudo-randomly assigned to one of two parallel groups in a double-blind, placebo-controlled, fully counterbalanced three-period crossover study design. All subjects were scanned three times, with 7 days separating each scan. Scanning was conducted at the same time of day per visit, and participants performed a modified version of the MID task (Hawkins et al., 2021). During each visit, each volunteer received a single capsule given orally with water. In one group the capsule contained either a single dose of risperidone 0.5 mg or risperidone 2 mg or placebo (herein referred to as group RIS-H/L); while in the other group participants received either a single oral dose of olanzapine (7.5 mg) or haloperidol (3 mg) or placebo (herein referred to as OLAN/HAL). Within-group treatment order was randomized using a Williams square design. Two subjects were discounted from the HAL/OLAN group due to DESPOT protocol unavailability, while two additional subjects were removed due to structural or functional image artifacts, leaving 17 subjects in this group (age: M=28.8, SD=6.3). In the RIS-H/L group, four subjects were removed for poor behavioral performance at the MID task (accuracy below chance), leaving a final sample of 17 subjects (age: M=27.6, SD=6.9 years). The study was approved by the London (Brent) Human Research Ethics Committee (REC reference: 13/LO/ 1183). Inclusion criteria required normal ECG, standard laboratory blood screens and urinalysis, and alcohol consumption within the recommended guidelines at the time of the study (<21 units per week). Exclusion criteria included a history of neurological of psychiatric illness, physical illness, and positive drugs of abuse or alcohol breath test on the screening or study days. Three volunteers from each group reported smoking one cigarette per day. However,

smoking was not permitted on the study days, so it is unlikely the potential acute effects previously reported would be a factor here (Franklin et al., 2015). A total of 34 subjects (age range 19–42 years, mean ± SD=26.9±6.8) were included in the final pharmacological challenge cohort.

Clinical cohort

We also recruited 35 individuals, comprising 17 with first-episode psychosis (FEP) and 17 with schizophrenia (SCZ) in the region of Apulia, Italy (12 F 22 M; age: M=24.7, SD=5.6). Recruitment procedures were conducted in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki), and approval was given by the local ethics committee ("Comitato Etico Indipendente Locale - Azienda Ospedaliero-Universitaria Consorziale Policlinico di Bari"). Diagnosis of SCZ was made using the Structured Clinical Interview for the DSM-5, Axis 1 disorders by board-certified psychiatrists. Exclusion criteria were: a significant history of drug or alcohol abuse; active drug abuse in the previous year; history of head trauma with loss of consciousness; any other critical medical condition.

B2. Trial-by-trial analysis

The analysis employed linear mixed-effects models to examine changes in VStr BOLD signal over trials in relation to cluster membership and other covariates. The models included only individuals correctly identified as GT or ST. Before entering the linear mixed model analysis, the dependent variables underwent a winsorization procedure to reduce the influence of outliers while retaining all data points. Specifically, values below the first percentile and above the 99th percentile were winsorized, a process known as 2% winsorization. Sex, age, and quadratic age effects were included as covariates to account for potential non-linear age-related differences in brain activity.

B3. BOLD activity in the anticipation and outcome ROIS – Pharmacological Challenge cohort

In the Pharmacological Challenge cohort, the BOLD signal extracted from each ROI was averaged within each phase to obtain a single estimate per participant. In particular, to assess the effect of the drug on BOLD activity, preliminary analyses were conducted to determine whether the BOLD signal extracted from anticipation and outcome ROIs could be combined to generate a single estimate for each reward phase. Two approaches were employed for this purpose. The first approach involved computing the mean BOLD signal separately for anticipation and outcome, yielding two distinct estimates corresponding to the respective reward phases. The second approach standardized and linearly combined the BOLD signal from anticipation and outcome ROIs into a single principal component (PC) for each phase. Given the high correlation between the mean-based and PC-based estimates (Anticipation: $r = .999$; Outcome: $r = .987$), subsequent analyses in this cohort were conducted using the mean-based estimates.

B4. Machine Learning pipeline

Feature selection and feature extraction

For each cohort, we selected our imaging features by first creating a total of 22 meta-ROIs (Figure III.9, Panel A, Step 1) based on a published meta-analysis (Chen et al., 2022) on the MID task—which identified a set of regions exhibiting enhanced BOLD activity during both reward anticipation (Reward vs. No Reward) and reward outcome (Reward vs. No Reward). Specifically, for all four cohorts, we created 12 ROIs for the anticipation phase (including bilateral VStr, thalamus, SMA, bilateral motor cortex, bilateral anterior insula, bilateral lateral occipital cortex, and bilateral middle frontal gyrus) and 10 ROIs for the outcome phase (including bilateral VStr, vmPFC, bilateral lateral occipital cortex,

posterior cingulate cortex, bilateral superior parietal lobule, left middle frontal gyrus, and right inferior frontal gyrus), showed in Figure III.10.

We then extracted the individual signal from the anticipation and outcome ROIs in each individual activation map grouped by the task phase (Figure III.9, Panel A, Step 2) using an uncorrected threshold of $p<.001$ through MarsBar (Brett et al., 2002) implemented in SPM12 (http://www.fil.ion.ucl.ac.uk/spm). This procedure was applied consistently across all cohorts.

Cluster analysis

Clustering was first performed in the Discovery cohort to examine whether participants could be grouped based on reward-related activity (Figure III.9, panel B). Then, to assess the validity of the algorithm, we implemented a leave-site-out cross validation in the Discovery cohort. Finally, we used the Discovery algorithm in the Replication cohort to obtain similar ST-GT imaging profiles, compared with clusters obtained from an independent algorithm on the Replication data to further support the generalizability of the clustering.

Implementation in the Discovery cohort

To investigate whether individuals could be grouped based on their brain activity during a rewarding task, we performed a hierarchical K-means clustering (Peterson et al., 2018) using Euclidean distance and Ward's linkage – implemented in the *factoextra* R package – on the full dataset including as features the BOLD signal extracted from the meta-ROIs on the individual task-related activity during i) the anticipation and ii) the outcome phases during the MID task (Figure III.9, step 3.1).

The features were initially checked for the potential presence of outliers using the Rosner's generalized extreme Studentized deviate test (Rosner, 1975) from the EnvStats R package ($p<0.05$). It is indicated for up to k potential outliers in a large dataset, assuming the data without any outliers come from a normal (Gaussian) distribution. It removes one potential outlier at a time, recalculates

test statistics, and continues until the specified number of potential outliers is evaluated, making it useful for datasets where multiple anomalies may be present.

Then, values were scaled to a [-1,1] range to remove the differences between features, and we adjusted the data for the effect of the site with the function *adjust* from the datawizard R package, which returned residuals of the regression models.

The number of optimal clusters (k=3) was defined through the hypothesis from animal studies Flagel et al., 2009; Flagel et al., 2017; Saunders and Robinson, 2012; Schad et al., 2020; for a review see Felix and Flagel, 2024) suggesting there are three distinct phenotypes, i.e., ST, INT, and GT, supported by the results from the *NbClust()* function [in NbClust R package] (Charrad et al., 2014), which provides the best clustering scheme by varying all combinations of number of clusters, distance measures, and clustering methods with 30 indices (Figure III.9, Step 3.1).

In our clustering, we hypothesized finding similar groups based on brain activity during the two separate phases of the task. Specifically, ST with elevated brain activity during the anticipation phase, GT with elevated brain activity during the outcome phase, and INT with mixed brain activity during both phases.

To test the potential effect of the sex in the clustering, we performed a hierarchical k-means clustering (k=3) separated for males (N=434) and females (N=454), replicating the results obtained with the clustering applied on the full dataset. Results are shown in Figure III.S1.

Leave-Site-Out Validation

Discovery cohort

To validate the clustering performed on the full dataset (N=890) described above, we performed the hierarchical k-means clustering (Peterson et al., 2018) within a leave-site-out validation framework iteratively for each site (Figure III.9, panel B, step 3.2). To this end, the Discovery dataset was initially split into a training set (seven sites) and a testing set (one site). Scaling (-1 to 1 range) was applied

separately in both sets, and residuals of regression models for the effect of the site were implemented only in the training set. The hierarchical k-means clustering (Peterson et al., 2018) was performed on the training set with k=3 using Euclidean distance and Ward's linkage (Johnson and Wichern, 2002). To validate this algorithm on the testing set, we assigned the cluster membership by computing the Euclidean distance between the testing set and the two training cluster centers – corresponding to the sign- and goal-trackers, due to the small sample size of the testing set. Then, we computed the clustering (k=2) on the testing set to label the testing set data. We calculated a confusion matrix (Kuhn, 2008) from *caret* R package to evaluate the cross-tabulation of the cluster assignments with associated statistics, i.e., sensitivity, specificity, balanced and overall accuracy. For two class problems, the sensitivity, specificity, positive predictive value and negative predictive value is calculated using the *positive* argument (if there are only two factor levels, the first level will be used as the "positive" result). The overall accuracy rate is computed along with a 95 percent confidence interval for this rate and a one-sided test to see if the accuracy is better than the "no information rate", which is taken to be the largest class percentage in the data. Balanced accuracy is computed in relation to sensitivity and specificity and, in this study, is used as a measure of replicability. To overcome the issue related to the incorrect order, we previously realigned the levels with *solve_LSAP()* function from the *clue* R package (Papadimitriou and Steiglitz, 1998).

Generalizability of the Discovery clustering algorithm

To assess the generalizability of our neurophysiological-based clustering, we applied the Discovery clustering algorithm to the independent Replication cohort.

As in the Discovery cohort, the BOLD signal was scaled (-1 to 1 range) to remove differences between features across samples. To assess the goodness of the Discovery algorithm in identifying the two imaging profiles in an independent dataset, everyone in the Replication dataset was assigned to a cluster based on its proximity to the Discovery cluster centers computed on the full dataset. In line

with both animal (Fitzpatrick & Morrow, 2020) and human studies (Colom, 2023; Schad et al., 2020; Dinu et al., 2024), which exclude the INT group from further analyses, we projected only the two centroids – ST and GT – to obtain predicted clusters (Figure III.9, panel B, step 4.1) and then we performed a hierarchical k-means clustering (Peterson et al., 2018) on the independent Replication dataset (k=2) with Euclidean distance and Ward's linkage to obtain the actual data-driven cluster memberships (Figure III.9, panel B, step 4.2).

The generalizability of the Discovery clustering algorithm was evaluated in the Replication cohort by implementing a confusion matrix (Kuhn, 2008), which compared the predicted clusters, based on the projection of the two Discovery centroids, i.e., ST/GT, and the two actual data-driven classes obtained from the independent hierarchical k-means algorithm on Replication data (Figure III.9, panel B, Step 4.3). The confusion matrix output distinguished the correctly identified from misidentified subjects (false positives and false negatives) for each class, i.e., ST/GT. Thus, we obtained overall accuracy and statistical measures of the sensitivity (true positive rates), specificity (true negative rates), and balanced accuracy for the two classes (Figure III.9, panel B, step 4.3).

The Discovery model was then applied to the Pharmacological Challenge, and Clinical cohorts (Figure III.9, panel B, Steps 4.1–4.3). In each case, data were scaled to the same range, clustering was repeated, and predicted assignments were compared with independent clustering results via confusion matrices. Sensitivity, specificity, and balanced accuracy were computed.

Finally, misclassified individuals resulting from the contingency table were excluded from subsequent individual-level analyses for all the four cohorts (Figure III.9 panel C, Steps 5.1, 5.2, 5.3), which instead included the predicted cluster membership based on the Discovery clustering algorithm.

B5. Placebo vs. drug effect using whole reward networks

Similarly to what was observed in the VStr, in the estimates obtained from the reward anticipation and outcome networks, we found a significant three-way interaction between phase, drug, and cluster ($F(1,96)=4.172$, $p=0.0438$). We conducted follow-up analyses separately for the anticipation and the outcome phases (Figure III.15A). During anticipation (Figure III.15A, left), a Drug × Cluster interaction ($F(1,32)=9.10$, $p=0.005$) showed that ST under placebo had greater responses than GT ($\beta=-0.721$, $SE=0.123$, $t(44.4)=-5.85$, $p < 0.001$). Drug administration significantly reduced ST BOLD response ($\beta=-0.4125$, $SE=0.0949$, $t(32)=-4.35$, $p < 0.001$) but had no effect on GT ($\beta=-0.0644$, $SE=0.0656$, $t(32)=-0.981$, $p=0.334$). Despite this reduction, ST still exhibited larger BOLD changes than GT under drug conditions ($\beta=-0.373$, $SE=0.123$, $t(44.4)=-3.03$, $p=0.0041$). Levene's tests confirmed no significant variance differences between groups under placebo ($F(1,32)=0.54$, $p=0.468$) or drug conditions ($F(1,32)=1.42$, $p=0.242$), ruling out variance differences as a confound in the observed drug effects. In the reward outcome phase (Figure III.15A, right), the Drug × Cluster interaction was not significant ($F(1,32)=2.62$, $p=0.115$).

B6. Dose-response effect in the reward anticipation phase using whole reward network

In the whole-network estimates, dose-response effects tested in the anticipation phase yielded during anticipation, for ST, a main effect of Drug ($\chi^2(2)=6.827$, $p=0.033$) with a significant linear trend ($\chi^2(1)=6.814$, $p=0.009$), indicating a dose-dependent response decrease (Figure III.15B). The quadratic model was not significant ($\chi^2(1)=0.013$, $p=0.910$), confirming a linear effect. High-dose risperidone significantly reduced responses compared to placebo and low-dose ($\chi^2(1)=4.578$, $p=0.032$). For GT, no significant drug effects were found ($\chi^2(2)=4.617$, $p=0.099$), with non-significant linear ($\chi^2(1)=3.523$, $p=0.060$) and quadratic ($p=0.296$) trends. Post-hoc comparisons showed no differences

between placebo and low-dose (β=-0.177, SE=0.097, t(16)=-1.826, pFDR=0.130), placebo and high-dose (β=-0.186, SE=0.097, t(16)=-1.917, pFDR=0.130), or low-dose and high-dose (β=-0.0089, SE=0.097, t(16)=-0.092, pFDR=0.928). Bayesian analysis strongly supported the null hypothesis, showing no risperidone effect on anticipatory responses in GT (BF_{01}=6.608, 5.883, 3.153).

B.7 Supplementary Figures

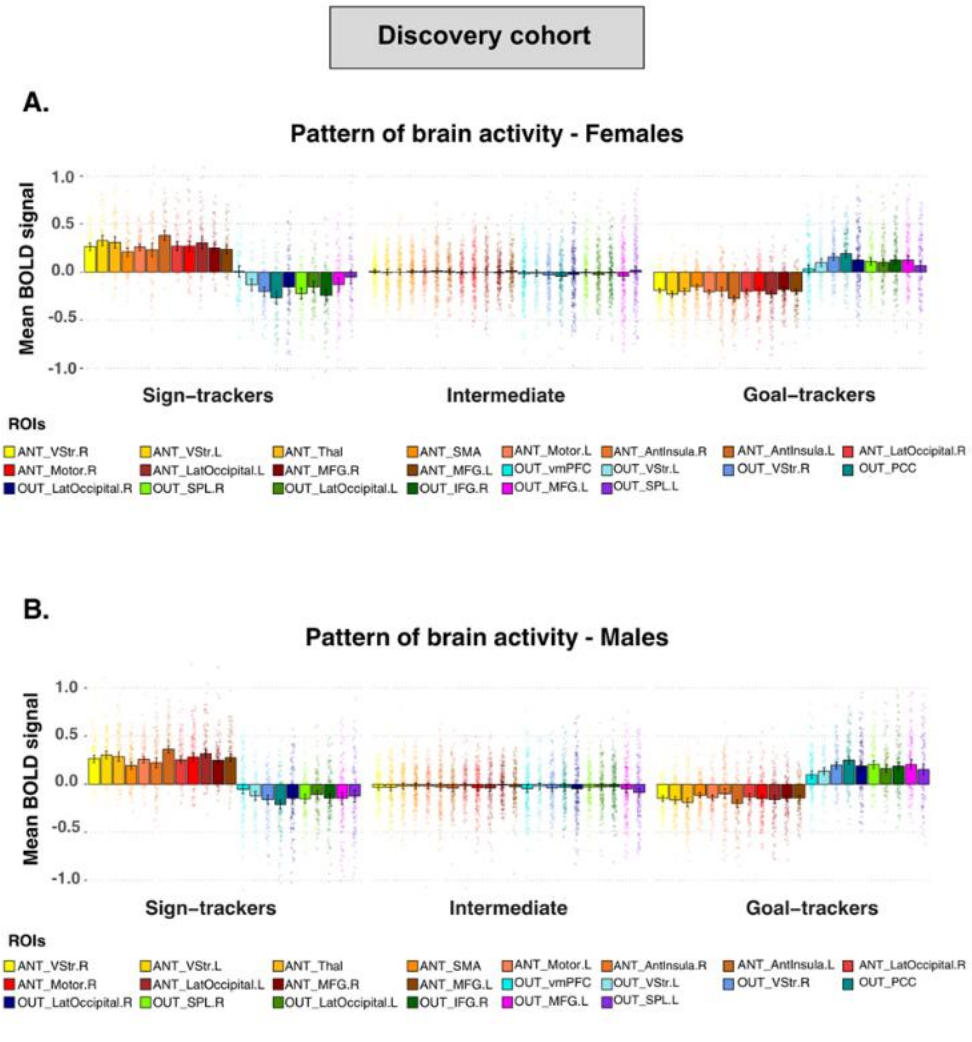

Figure III.S 1 Pattern of brain activity grouped by sex in the Discovery cohort. Barplots showing BOLD activity in each ROI for each cluster, i.e., sign-trackers, intermediate and goal-trackers, in females (**A**) and males (**B**) in the Discovery cohort. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

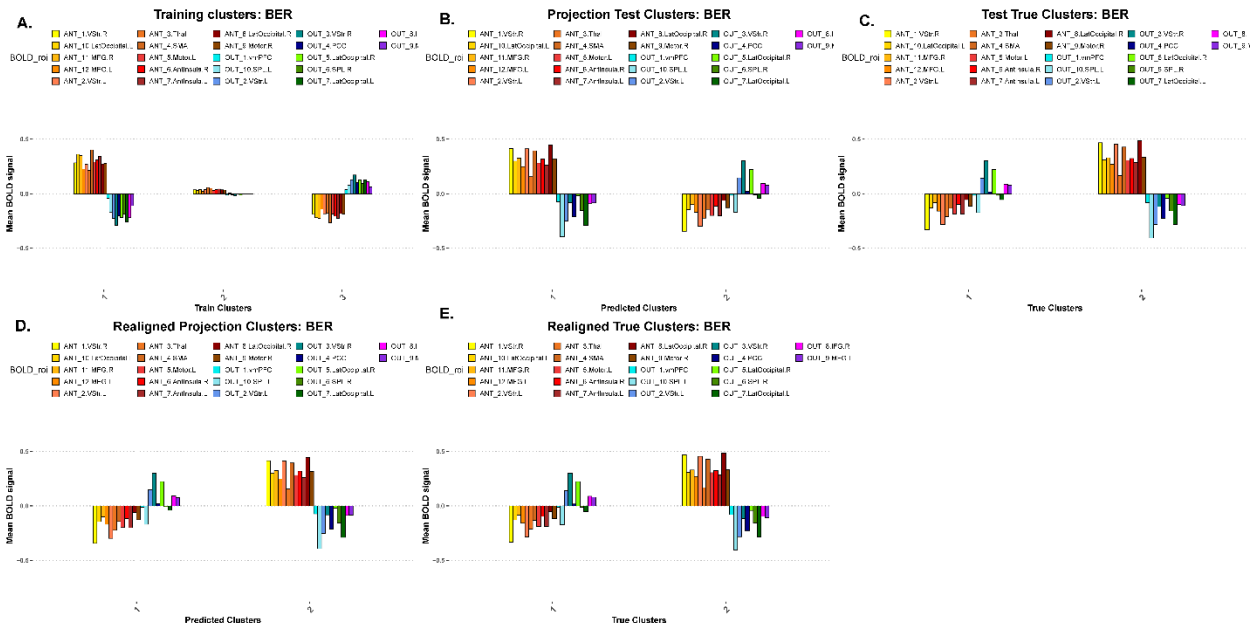

Figure III.S 2 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: BERLIN. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Berlin site), derived from the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

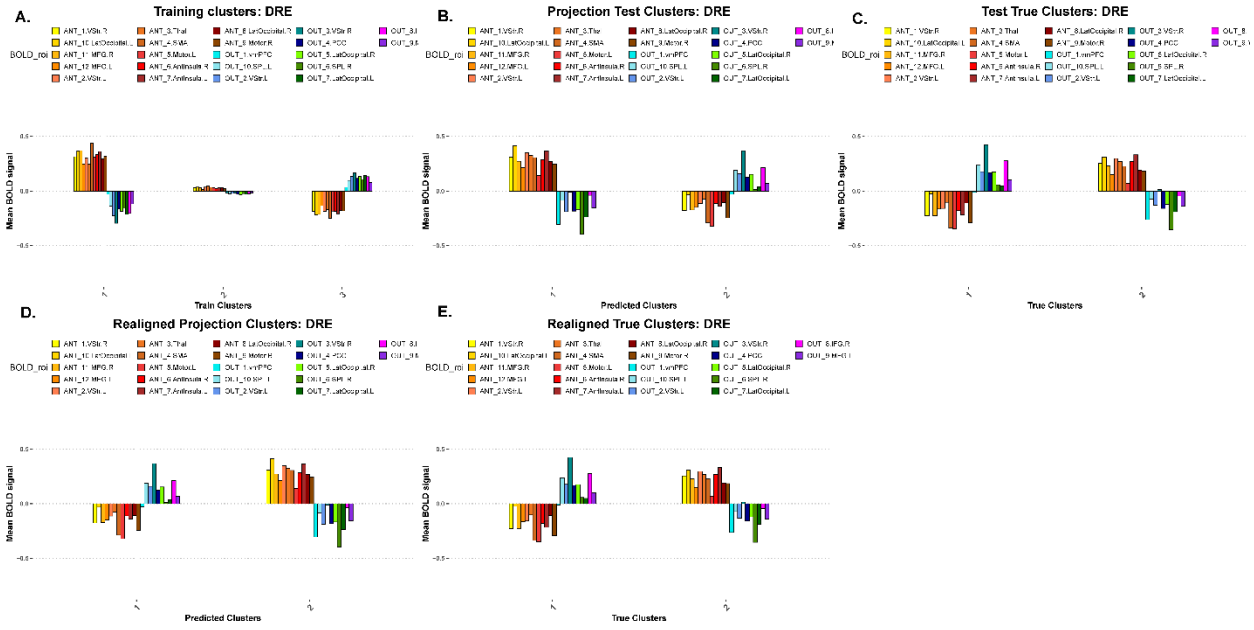

Figure III.S 3 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: DRESDEN. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Desden site), derived from

the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

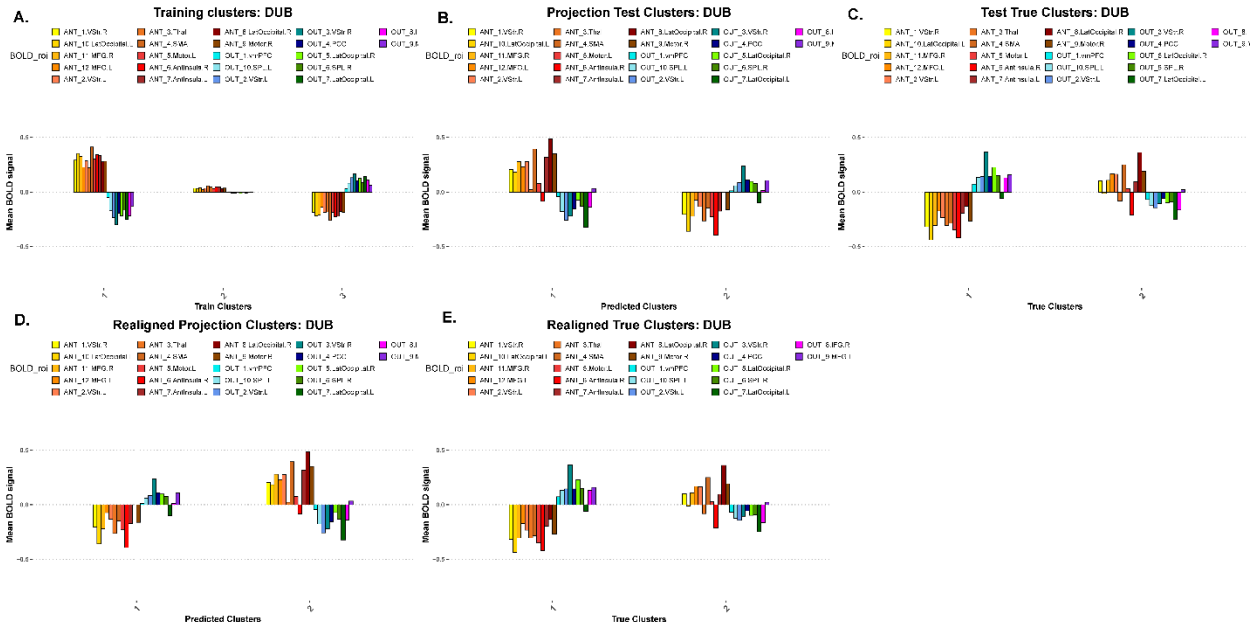

Figure III.S 4 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: DUBLIN. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Dublin site), derived from the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

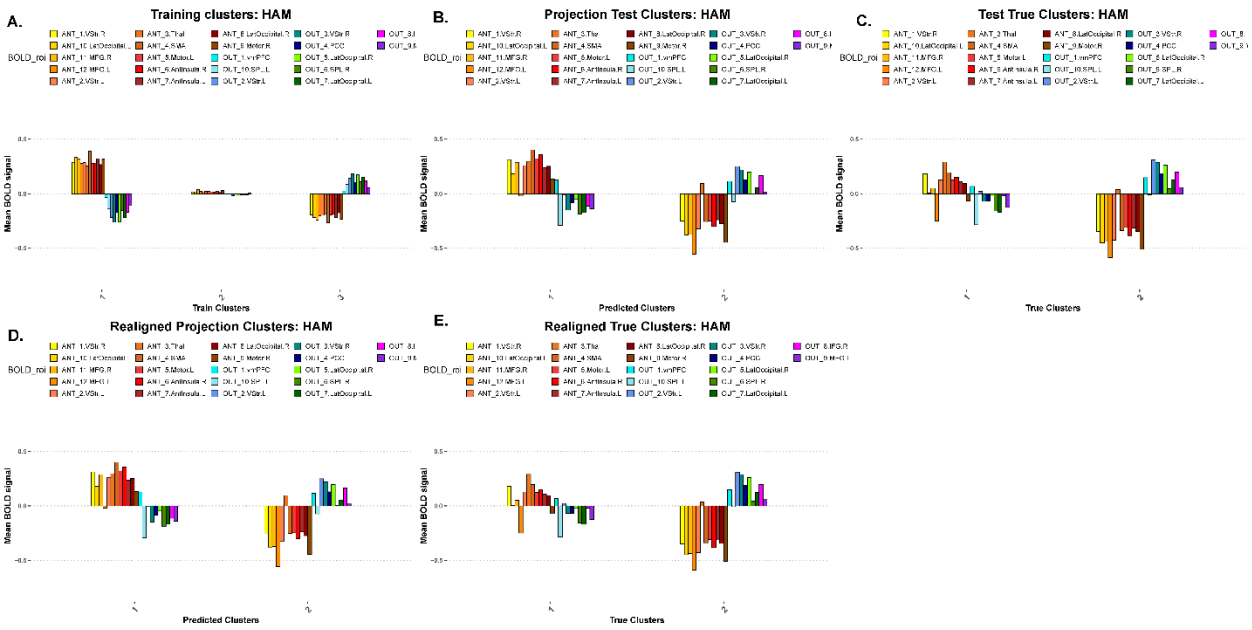

Figure III.S 5 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: HAMBURG. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Hamburg site), derived from the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

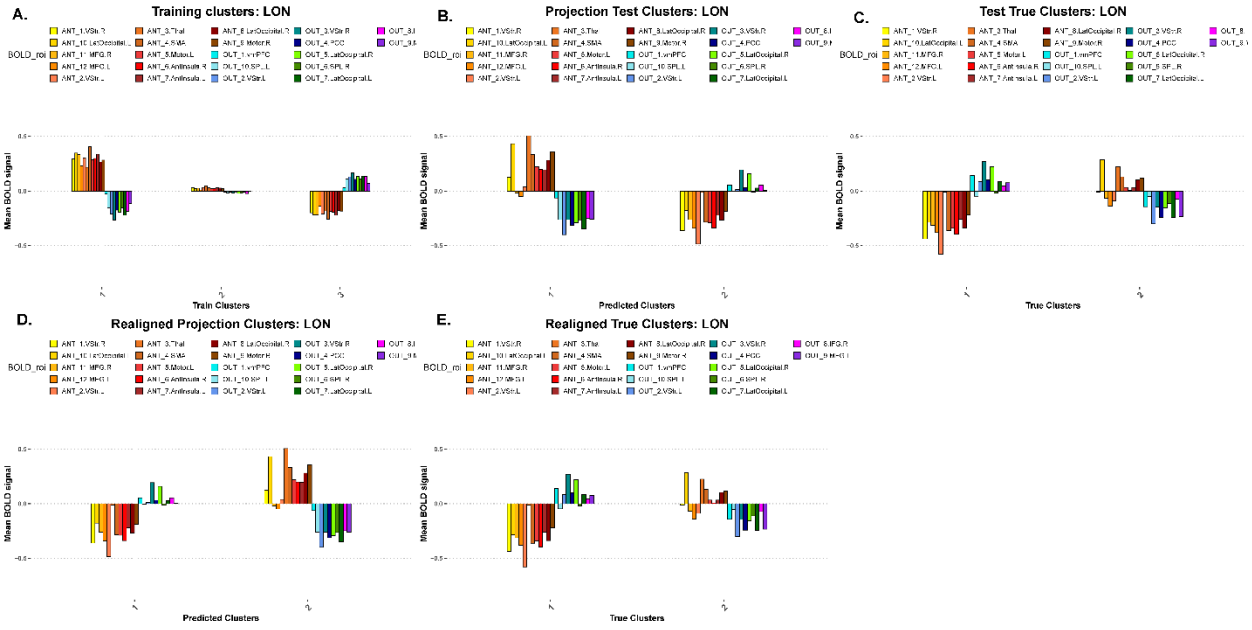

Figure III.S 6 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: LONDON. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (London site), derived from

the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

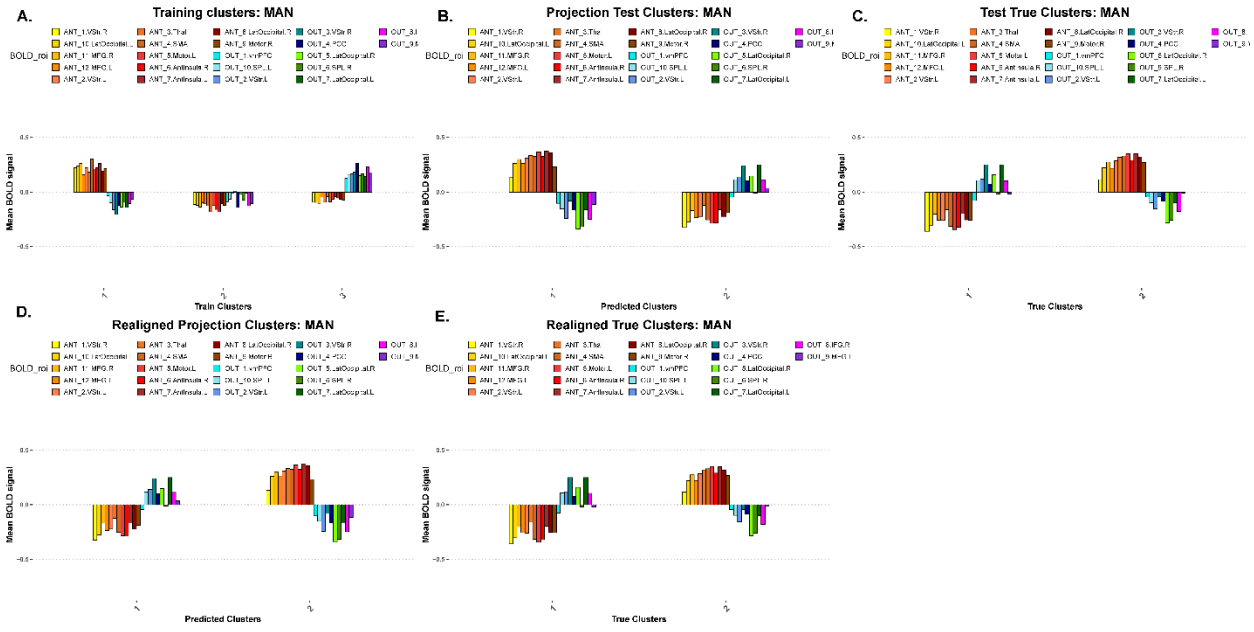

Figure III.S 7 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: MANNHEIM. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Mannehim site), derived from the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

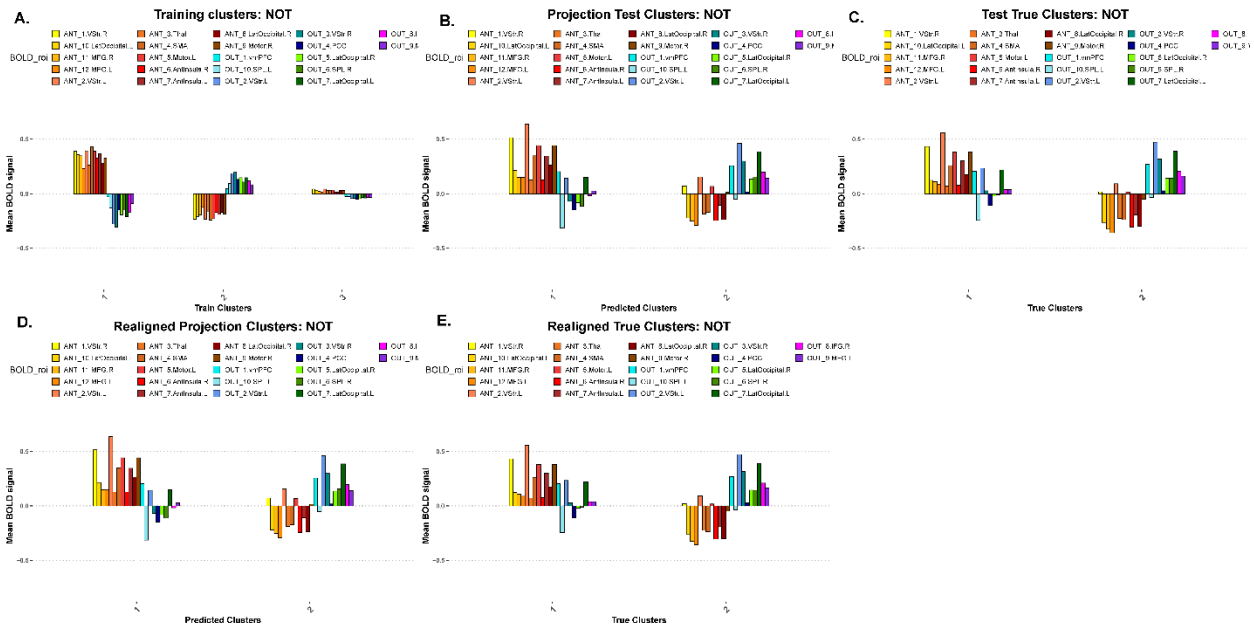

Figure III.S 8 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: NOTTINGHAM. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Nottingham site), derived from the projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

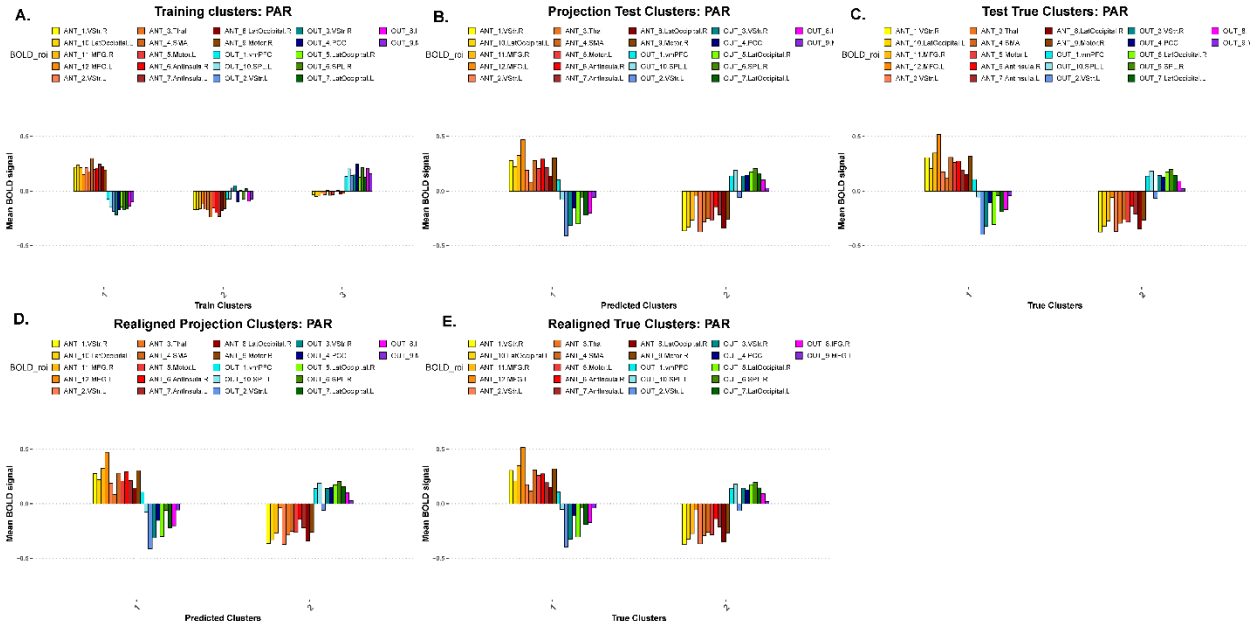

Figure III.S 9 Pattern of brain activity in the leave-site-out validation framework of the Discovery cohort. Testing site: PARIS. Barplots showing BOLD activity in each ROI for each cluster in the (**A**) training set (made up of seven sites) derived with a k=3 solution, (**B-D**) testing set (Paris site), derived from the

projection of the two centroids corresponding to ST and GT, and in the (**C-E**), testing set, derived from the implementation of a hierarchical k-means algorithm with k=2. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

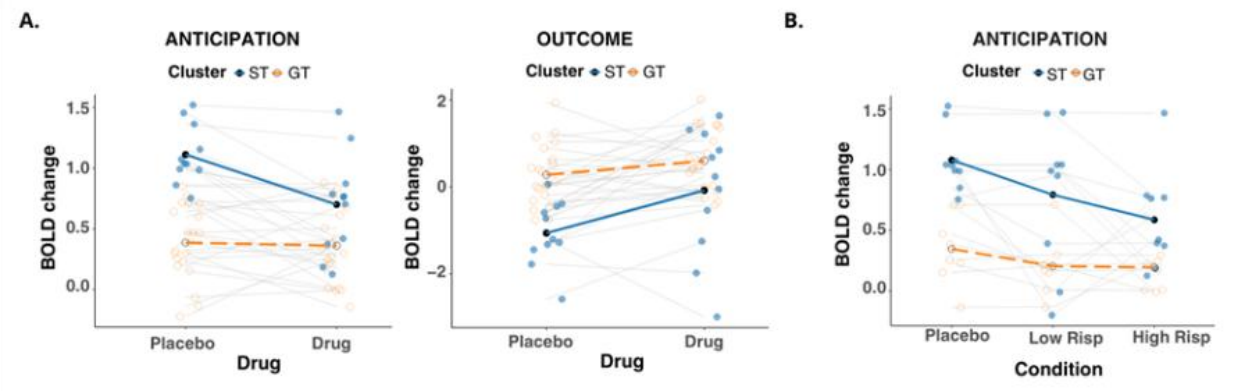

Figure III.S 10 Whole-network analyses in the Pharmacological challenge cohort. Plots representing the Placebo vs. Drug effect (A) and the Dose-Response effect (B) across the whole reward network. *Abbreviations*: ST=sign-trackers; GT=goal-trackers; BOLD=blood-oxygen-level-dependent.

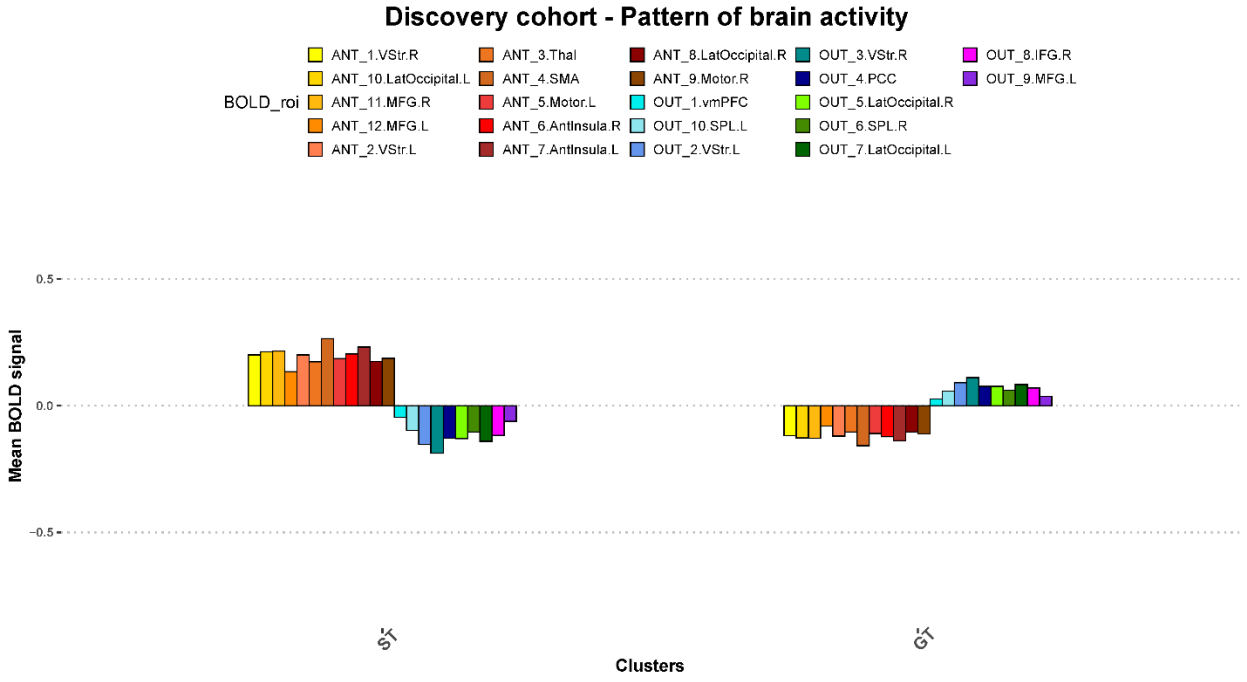

Figure III.S 11 Pattern of brain activity in the Discovery cohort with k=2. Barplots showing BOLD activity in each ROI for each cluster, i.e., sign-trackers and goal-trackers, in the Discovery cohort. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

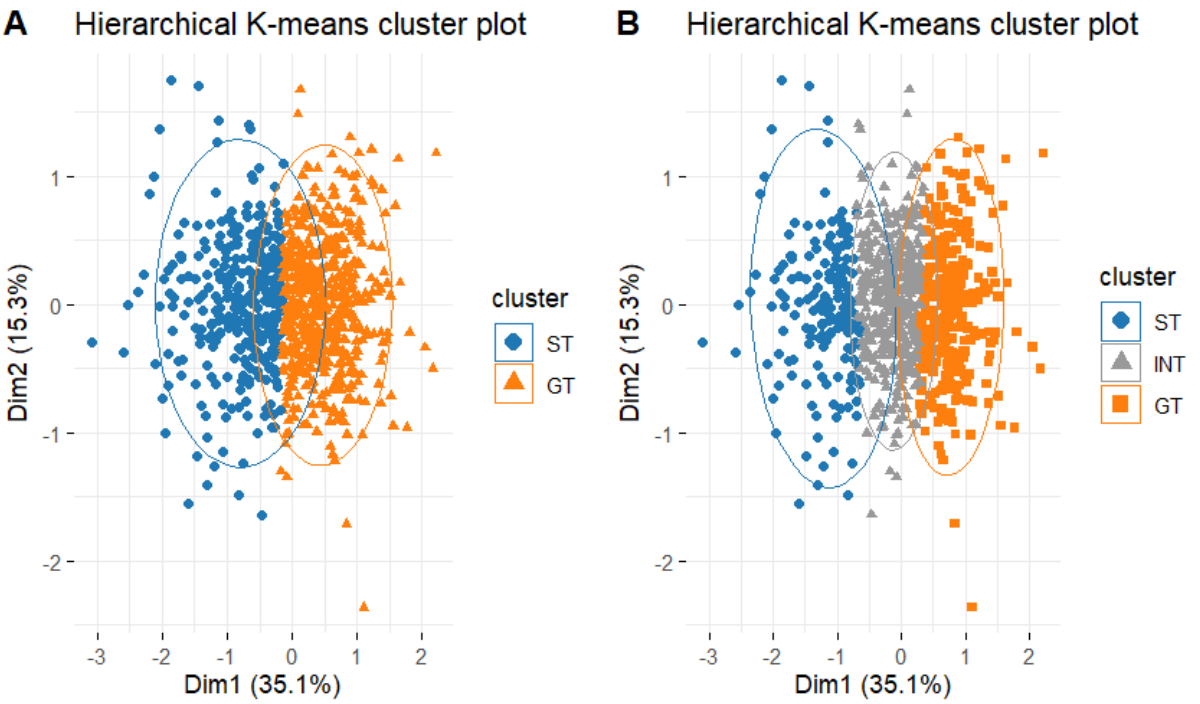

Figure III.S 12 Hierarchical k-means clustering solution visualized on the first two principal components. (**A**) Two-cluster solution (k = 2) derived from hierarchical k-means clustering, identifying Sign-Trackers (ST) and Goal-Trackers (GT). (**B**) Three-cluster solution (k = 3), revealing an intermediate group (INT) positioned between the ST and GT clusters. Each point represents an individual participant projected onto the first two principal components (Dim1 and Dim2), explaining 35.1% and 15.3% of the variance, respectively. Ellipses indicate the dispersion of observations within each cluster. Colors denote cluster membership.

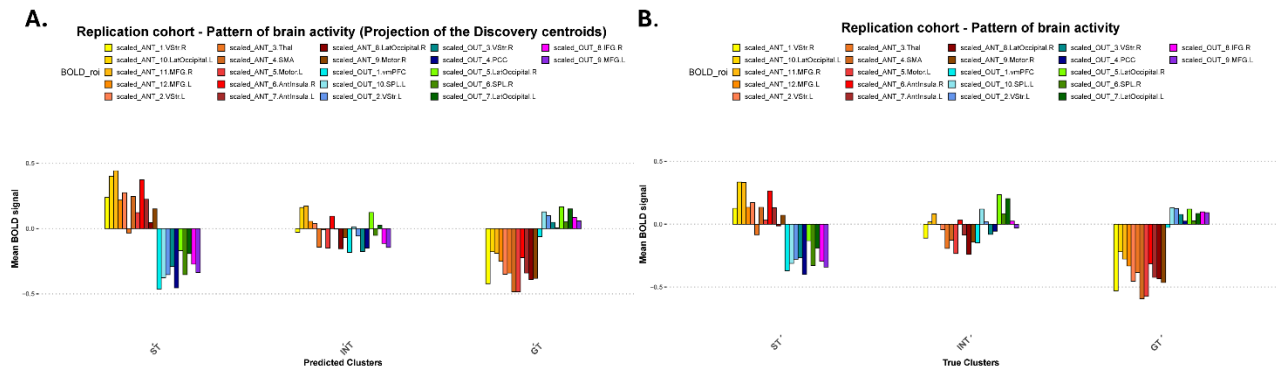

Figure III.S 13 Pattern of brain activity in the Replication cohort. Barplots showing BOLD activity in each ROI for each cluster, i.e., sign-trackers, intermediate and goal-trackers, predicted through the projection of all three centroids from the Discovery algorithm (**A**), and derived from an independent hierarchical k-means clustering (**B**) in the Replication cohort. *Abbreviations:* ANT = anticipation; OUT = outcome; R = right; L = left; VStr = ventral striatum; SPL = superior parietal lobule; Thal = thalamus; MFG = middle frontal gyrus; SMA = supplementary motor area; IFG = inferior frontal gyrus; vmPFC = ventromedial prefrontal cortex; PCC = posterior cingulate cortex.

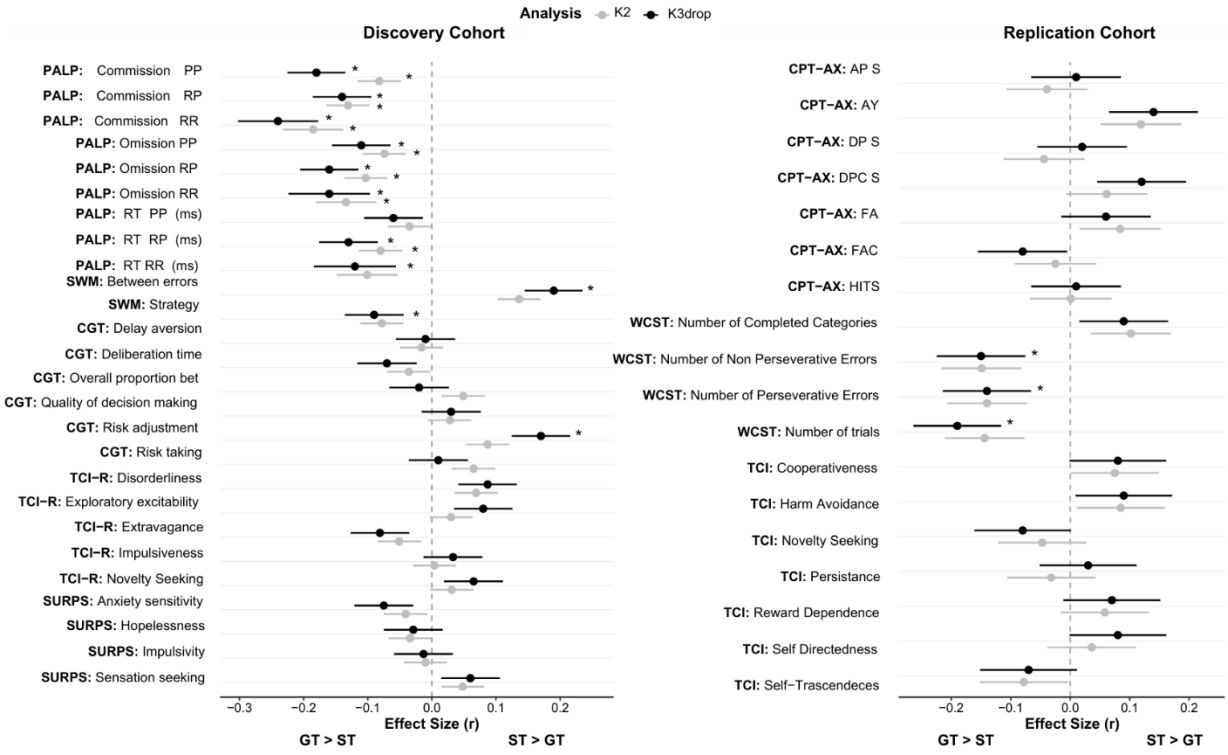

Figure III.S 14 Neuropsychological differences between Sign-Trackers and Goal-Trackers across Discovery and Replication cohorts. Forest plots showing effect sizes (Pearson's *r*) for group differences between Sign-Trackers (ST) and Goal-Trackers (GT) across neuropsychological measures in the Discovery cohort (left) and Replication cohort (right). Effect sizes are displayed for two clustering approaches: the two-cluster solution (K = 2) and the three-cluster solution excluding the INT group (K = 3 drop), retaining only ST and GT individuals. Points represent effect size estimates, and horizontal bars indicate confidence intervals. Positive values indicate ST > GT, whereas negative values indicate GT > ST. Asterisks denote statistically significant differences after multiple-comparison correction. The vertical dashed line marks zero effect.

B.7 Supplementary Tables

Table III.S1

Cohort	Sample Size (N)	M:F	Mean Age ± SD
Discovery	890	436:454	22.07 ± 0.7
Sites			
Berlin	123	55:68	22.1 ± 0.6
Dresden	102	55:47	21.5 ± 0.7
Dublin	103	58:45	22.3 ± 0.7
Hamburg	122	62:60	22.1 ± 0.9
London	104	54:50	22.2 ± 0.6
Mannheim	115	55:60	22 ± 0.5
Nottingham	116	54:62	22.1 ± 0.6
Paris	105	43:62	22.2 ± 0.6
Replication	245	104:141	26 ± 6
Pharmacological Challenge	34	34:0	26.9 ± 6.8
Clinical	35	24:11	24.7 ± 5.6

Table III.S 1 Demographic characteristics across the four study cohorts. The table presents demographic data for all four cohorts, with the Discovery cohort further stratified by site. For each cohort and subgroup, the sample size, sex distribution (number of males and females), and mean age with standard deviation (SD) are reported. *Abbreviations:* M=males; F=females.

Table III.S2

Phase: Contrast	Group	MNI coordinates X Y Z	Regions
Discovery Cohort			
Anticipation: Reward > No Reward	*Sign-trackers*	-28 -78 -18	Lingual Gyrus Fusiform L
		-20 -78 -18	BA18
		-36 -64 -20	Cerebellum L

		10 -80 -14	Lingual R BA18
		12 -84 8	Cuneus R
		-30 -58 -22	Cuneus R
		4 -86 6	Calcarine R BA17
		30 -76 -20	BA19 Fusiform R
		12 -2 16	Caudate Body R
		-12 0 14	Caudate L
Anticipation: ***Reward > No Reward***	*Goal-trackers*	10 -86 0	Calcarine R
		-16 -90 -12	Lingual L
		38 -22 58	BA4 Precentral R
		12 -16 8	Thalamus R
		-12 -18 8	Thalamus L
		-16 -24 12	Pulvinar L
		14 16 -2	Caudate R
		-10 8 -2	Pallidum L
		30 -54 -18	Fusiform R
Outcome: ***Reward > No Reward***	*Sign-trackers*	-6 2 24	Corpus Callosum
		0 24 -6	Anterior Cingulum L
		-16 -26 -16	ParaHippocampal L
		-6 -32 -58	Midbrain
		32 14 -18	Insula R
Outcome: ***Reward > No Reward***	*Goal-trackers*	2 50 2	Frontal Superior Medial R
		4 46 -6	Frontal Medial Orbital R
		2 46 20	Cingulum Anterior R
		0 60 4	Frontal Superior Medial L
Replication Cohort			

Anticipation: ***Reward > No Reward***	*Sign-trackers*	4 0 58	Supplementary Motor Area R
		-4 6 56	Supplementary Motor Area L
		28 16 -44	Temporal Pole R
		-42 14 54	Primary Motor Cortex L
Anticipation: ***Reward > No Reward***	*Goal-trackers*		
Outcome: ***Reward > No Reward***	*Sign-trackers*	-14 -88 -18	Secondary Visual Cortex L
		2 -84 -10	Secondary Visual Cortex R
		-52 18 -10	Superior Temporal Pole L
		-40 18 -10	Inferior Orbital Frontal Gyrus L
Outcome: ***Reward > No Reward***	*Goal-trackers*	-16 -90 -6	Middle Secondary Visual Cortex L
		14 -90 -2	Primary Visual Cortex R
		14 -84 -12	Secondary Visual Cortex R
Clinical cohort			
Anticipation: ***Reward > No Reward***	*Sign-trackers*	-42 -24 58	Supplementary Motor Area L
		-38 -28 64	Primary Motor Cortex L
		-32 -40 58	Primary Sensory Cortex L
		4 -14 44	Ventral Anterior Cingulum R
		-54 -22 34	Supramarginal Gyrus L
		-62 -18 18	Primary Sensory L
		36 -16 18	Insula R
		-24 12 -8	Putamen L
		-14 -66 10	Primary Visual Cortex R
		46 -12 54	Primary Motor Cortex R
		58 -2 44	Supplementary Motor Area R
Anticipation:	*Goal-trackers*		

Reward > No Reward			
Outcome: ***Reward > No Reward***	*Sign-trackers*		
Outcome: ***Reward > No Reward***	*Goal-trackers*	54 26 34	Dorsolateral Prefrontal Cortex R
		12 10 8	Caudate R
		-28 -96 12	Secondary Visual Cortex L
		16 -62 8	Primary Visual Cortex R
		50 -5 -18	Superior Temporal Gyrus R
		24 -88 -8	Secondary Visual Cortex R
		-42 34 24	Dorsolateral Prefrontal Cortex L
		22 -2 -20	Amygdala R
		-40 -12 12	Insula L
		-36 -30 16	Primary Auditory L
		8 16 -12	Subgenual R
		-42 -12 18	Primary Motor Cortex L
		-50 -12 16	Primary Sensory Cortex L

Table III.S 2 Whole-brain activation during the MID task in the Discovery, Replication, and Clinical cohorts. Brain regions' activations were detected through one-sample t-tests on the Reward > No-Reward contrast during anticipation and outcome in the Discovery, Replication, and Clinical cohorts. All regions are shown at $p_{TFCE\text{-}FWE}<0.05$; cluster extent=20 voxels. *Abbreviations*: MNI=Montreal Neurological Institute; L=left; R= right; BA=Broadmann Area.

Table III.S3

Sample	Questionnaire/Task	Score	Z	p(unc.)	p_{FDR}	effect size r	Cluster	
							ST	GT
							Mean ± SD (N)	Mean ± SD (N)
Discovery	**Temperament and Character Inventory-Revised (TCI-R)**	*Disorderliness*	-1.902	0.057	0.135	0.087	21.65 ± 4.03 (160)	21 ± 4.12 (317)

		Exploratory excitability	-1.745	0.081	0.135	0.08	34.63 ± 4.24 (160)	33.96 ± 4.24 (317)
		Extravagance	-1.777	0.076	0.135	0.081	26.21 ± 3.69 (160)	26.78 ± 4.03 (317)
		Impulsiveness	-0.71	0.478	0.478	0.033	25.27 ± 3.96 (160)	24.96 ± 4.15 (317)
		Novelty Seeking	-1.411	0.158	0.198	0.065	107.78 ± 9.18 (160)	106.70 ± 10.46 (317)
	Substance Use Risk Profile Scale (SURPS)	*Anxiety sensitivity*	-1.635	0.102	0.379	0.075	11.31 ± 2.51 (160)	11.75 ± 2.51(317)
		Hopelessness	-0.624	0.533	0.710	0.029	12.63 ± 3.5 (160)	12.89 ± 3.77 (317)
		Impulsivity	-0.289	0.773	0.773	0.013	10.67 ± 2.11 (160)	10.7 ± 2.27 (317)
		Sensation seeking	-1.313	0.189	0.379	0.06	16.84 ± 3.37 (160)	16.47 ± 3.38 (317)
	Cambridge guessing task (CGT)	*Delay aversion*	-0.173	0.863	0.863	0.01	0.165 ± 0.13 (156)	0.17 ± 0.15 (314)
		Deliberation time	-1.502	0.133	0.399	0.07	1510.02 ± 443.43 (156)	1569.68± 478.33 (314)
		Overall proportion bet	-0.322	0.748	0.863	0.02	0.51 ± 0.13 (156)	0.51± 0.13 (314)
		Quality of decision making	-0.662	0.508	0.863	0.03	0.97 ± 0.05 (156)	0.97 ± 0.05 (314)

		Risk adjustment	-3.588	0.000	0.002	0.17	2.25 ± 0.93 (156)	1.93 ± 1.03 (314)
		Risk taking	-0.299	0.765	0.863	0.01	0.57 ± 0.12 (156)	0.57 ± 0.13 (314)
	Passive Avoidance Learning Paradigm (PALP)	*Commission RP*	-3.040	0.000	0.001	0.14	0.078 ± 0.07 (158)	0.11 ± 0.09 (315)
		Commission PP	-3.945	0.002	0.005	0.18	0.08 ± 0.08 (160)	0.11 ± 0.1 (315)
		Commission RR	-3.713	0.000	0.001	0.24	0.06 ± 0.08 (81)	0.1 ± 0.09 (162)
		Omission PP	-2.449	0.014	0.018	0.11	0.03 ± 0.04 (160)	0.03 ± 0.05 (315)
		Omission RP	-3.394	0.001	0.002	0.16	0.02 ± 0.05 (158)	0.05 ± 0.05 (315)
		Omission RR	-2.530	0.011	0.017	0.16	0.02 ± 0.04 (81)	0.03 ± 0.04 (162)
		RT PP (ms)	-1.277	0.202	0.202	0.06	876.58 ± 206.03 (160)	905.66 ± 221.4 (315)
		RT RP (ms)	-2.743	0.006	0.011	0.13	820.57 ± 173.99 (158)	876.11 ± 192.67 (315)
		RT RR (ms)	-1.825	0.068	0.077	0.12	857.91 ± 213.33 (81)	918.65 ± 255.19 (162)

Replication	**Temperament and Character Inventory (TCI)**	*Cooperativeness*	-0.962	0.336	0.456	0.08	19.73 ± 3.6 (42)	19.08 ± 4.01 (110)
		Harm Avoidance	-1.05	0.294	0.456	0.09	8.96 ± 4.11 (42)	8.4 ± 4.06 (110)
		Novelty Seeking	-0.93	0.352	0.456	0.08	7.89 ± 4.52 (42)	8.5 ± 4.43 (110)
		Persistance	-0.336	0.737	0.737	0.03	3.05 ± 1.56 (42)	2.97 ± 1.91 (110)
		Reward Dependence	-0.859	0.390	0.456	0.07	9.7 ± 2.43 (42)	9.4 ± 2.47 (110)
		Self Directedness	-1.03	0.303	0.456	0.08	19.626 ± ± 6.01 (42)	18.62 ± 6.14 (110)
		Self-Trascendeces	-0.903	0.366	0.456	0.07	5.3 ± 6.25 (42)	6.01 ± 6.36 (110)
	Wisconsin Card Sorting Test (WCST)	*Number of Completed Categories*	-1.144	0.126	0.126	0.09	5.86 ± 0.53 (51)	5.67 ± 1.06 (129)
		Number of Non-Perseverative Errors	-2.019	0.022	0.042	0.15	7.06 ± 7.03 (51)	9.58 ± 9.35 (127)
		Number of Perseverative Errors	-1.861	0.031	0.042	0.14	7.06 ± 3.75 (51)	9.95 ± 8.46 (128)
		Number of trials	-2.484	0.007	0.026	0.19	80.54 ± 15.84 (51)	88.33 ± 20.42 (128)
	Continuous Performance Test (CPT-AX)	*HIT*	-0.143	0.886	0.430	0.01	0.97 ± 0.04 (50)	0.98 ± 0.09 (128)

		FA	-0.801	0.423	0.430	0.06	0.06 ± 0.06 (50)	0.06 ± 0.07 (128)
		FAC	-1.072	0.284	0.662	0.08	0.02± 0.03 (50)	0.04 ± 0.06 (128)
		AP S	-0.115	0.909	0.740	0.01	0.97 ± 0.03 (50)	0.98 ± 0.03 (128)
		AY	-1.826	0.068	0.909	0.14	0.09 ± 0.11 (50)	0.08 ± 0.1 (128)
		DP S	-0.199	0.842	0.909	0.02	4.19 ± 0.99 (50)	4.18 ± 1.04 (128)
		DPC S	-1.543	0.123	0.909	0.12	4.91 ± 0.97 (50)	4.62 ± 1.11 (128)

Table III.S 3 Descriptive statistics by imaging-derived clusters and Wilcoxon rank-sum test results in the Discovery and Replication cohorts. This table reports descriptive and inferential statistics for each imaging-derived cluster across both the Discovery and Replication cohorts. Analyses were conducted on standardized data (scaled 0–1). Effect sizes (r equivalent) were computed using the *rstatix* R package (Tomczak and Tomczak, 2014), by dividing the Wilcoxon Z-transformed values by the square root of the number of observations (Rosenthal and Rubin, 2003). For each variable, the table includes Wilcoxon Z-values, effect sizes, uncorrected p-values, and p-values corrected for multiple comparisons using the FDR correction ($p_{FDR} < 0.05$). Means and standard deviations (SD) are reported based on raw (non-standardized) data to facilitate interpretability. Cluster sample sizes (N) are also indicated. *Abbreviations:* FDR=False Discovery rate; ST=sign-trackers; GT=goal-trackers; SD=standard deviation; RT=reaction times, PP=punishment-punishment, RR=reward-reward, RP=reward-punishment.

References

Abi-Dargham, A., Rodenhiser, J., Printz, D., Zea-Ponce, Y., Gil, R., Kegeles, L. S., ... & Laruelle, M. (2000). Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proceedings of the National Academy of Sciences*, *97*(14), 8104-8109.

Abi-Dargham, A., van de Giessen, E., Slifstein, M., Kegeles, L. S., & Laruelle, M. (2009). Baseline and amphetamine-stimulated dopamine activity are related in drug-naive schizophrenic subjects. *Biological psychiatry*, *65*(12), 1091-1093.

Abler, B., Erk, S., & Walter, H. (2007). Human reward system activation is modulated by a single dose of olanzapine in healthy subjects in an event-related, double-blind, placebo-controlled fMRI study. *Psychopharmacology*, *191*(3), 823-833.

Acsády, L. (2022). Organization of thalamic inputs. *The thalamus*, 27-44.

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507-517.

Adhikari, B. M., Jahanshad, N., Shukla, D., Glahn, D. C., Blangero, J., Fox, P. T., ... & Kochunov, P. (2018). Comparison of heritability estimates on resting state fMRI connectivity phenotypes using the ENIGMA analysis pipeline. *Human brain mapping*, *39*(12), 4893-4902.

Admon, R., Kaiser, R. H., Dillon, D. G., Beltzer, M., Goer, F., Olson, D. P., ... & Pizzagalli, D. A. (2017). Dopaminergic enhancement of striatal response to reward in major depression. *American Journal of Psychiatry*, *174*(4), 378-386.

Aggleton, J. P. (2012). Multiple anatomical systems embedded within the primate medial temporal lobe: implications for hippocampal function. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1579-1596.

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and brain sciences*, *22*(3), 425-444.

Aggleton, J. P., & Mishkin, M. (1984). Projections of the amygdala to the thalamus in the cynomolgus monkey. *Journal of comparative neurology*, *222*(1), 56-68.

Aggleton, J. P., & O'Mara, S. M. (2022). The anterior thalamic nuclei: core components of a tripartite episodic memory system. *Nature Reviews Neuroscience*, *23*(8), 505-516.

Aggleton, J. P., Dumont, J. R., & Warburton, E. C. (2011). Unraveling the contributions of the diencephalon to recognition memory: a review. *Learning & memory*, *18*(6), 384-400.

Aggleton, J. P., O'Mara, S. M., Vann, S. D., Wright, N. F., Tsanov, M., & Erichsen, J. T. (2010). Hippocampal–anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *European Journal of Neuroscience*, *31*(12), 2292-2307.

Albertella, L., Le Pelley, M. E., Chamberlain, S. R., Westbrook, F., Fontenelle, L. F., Segrave, R., ... & Yücel, M. (2019a). Reward-related attentional capture is associated with severity of addictive and obsessive–compulsive behaviors. *Psychology of Addictive Behaviors*, *33*(5), 495.

Albertella, L., Watson, P., Yücel, M., & Le Pelley, M. E. (2019b). Persistence of value-modulated attentional capture is associated with risky alcohol use. *Addictive behaviors reports*, *10*, 100195.

Alcauter, S., Lin, W., Smith, J. K., Short, S. J., Goldman, B. D., Reznick, J. S., ... & Gao, W. (2014). Development of thalamocortical connectivity during infancy and its cognitive correlations. *Journal of Neuroscience*, *34*(27), 9067-9075.

Alkam, Tursun, Andrew H. Van Benschoten, and Ebrahim Tarshizi. "Reinforcement learning in artificial intelligence and neurobiology." Neuroscience Informatics (2025): 100220.

Alkemade, A., Bazin, P. L., Balesar, R., Pine, K., Kirilina, E., Möller, H. E., ... & Forstmann, B. U. (2022). A unified 3D map of microscopic architecture and MRI of the human brain. *Science advances*, *8*(17), eabj7892.

Allen, E. A., Erhardt, E. B., Damaraju, E., Gruner, W., Segall, J. M., Silva, R. F., ... & Calhoun, V. D. (2011). A baseline for the multivariate comparison of resting-state networks. *Frontiers in systems neuroscience*, *5*, 2.

Amaral, D. G., & Cowan, W. M. (1980). Subcortical afferents to the hippocampal formation in the monkey. *Journal of Comparative Neurology*, *189*(4), 573-591.

Ambrus, G. G. (2024). Shared neural codes of recognition memory. *Scientific Reports*, *14*(1), 15846.

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in cognitive sciences*, *17*(8), 379-390.

Anderson, J. R. (2000). *Learning and memory: An integrated approach*. John Wiley & Sons Inc.

Anderson, J. R. (2013). *The architecture of cognition*. Psychology Press.

Andreasen, N. C., O'Leary, D. S., Cizadlo, T., Arndt, S., Rezai, K., Ponto, L. L., ... & Hichwa, R. D. (1996). Schizophrenia and cognitive dysmetria: a positron-emission tomography study of dysfunctional prefrontal-thalamic-cerebellar circuitry. *Proceedings of the National Academy of Sciences*, *93*(18), 9985-9990.

Anselme, P., & Robinson, M. J. (2020). From sign-tracking to attentional bias: Implications for gambling and substance use disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *99*, 109861.

Anticevic, A., Haut, K., Murray, J. D., Repovs, G., Yang, G. J., Diehl, C., ... & Cannon, T. D. (2015). Association of thalamic dysconnectivity and conversion to psychosis in youth and young adults at elevated clinical risk. *JAMA psychiatry*, *72*(9), 882-891.

Antonucci, L. A., Di Carlo, P., Passiatore, R., Papalino, M., Monda, A., Amoroso, N., ... & Blasi, G. (2019). Thalamic connectivity measured with fMRI is associated with a polygenic index predicting thalamo-prefrontal gene co-expression. *Brain Structure and Function*, *224*(3), 1331-1344.

Antonucci, L. A., Nettis, M. A., Juckel, G., Roser, P., Pergola, G., & Thoma, P. (2020). Selective recall deficits for heterogeneous associations in detoxified individuals with alcohol use disorder. *Behavioural brain research*, *390*, 112688.

Antonucci, L. A., Penzel, N., Pigoni, A., Dominke, C., Kambeitz, J., & Pergola, G. (2021). Flexible and specific contributions of thalamic subdivisions to human cognition. *Neuroscience & Biobehavioral Reviews*, *124*, 35-53.

Antonucci, L. A., Pergola, G., Pigoni, A., Dwyer, D., Kambeitz-Ilankovic, L., Penzel, N., ... & Bertolino, A. (2020). A pattern of cognitive deficits stratified for genetic and environmental risk reliably classifies patients with schizophrenia from healthy control subjects. *Biological psychiatry*, *87*(8), 697-707.

Antonucci, L. A., Taurisano, P., Fazio, L., Gelao, B., Romano, R., Quarto, T., ... & Blasi, G. (2016). Association of familial risk for schizophrenia with thalamic and medial prefrontal functional connectivity during attentional control. *Schizophrenia research*, *173*(1-2), 23-29.

Arbabshirani, M. R., & Calhoun, V. D. (2011, August). Functional network connectivity during rest and task: comparison of healthy controls and schizophrenic patients. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4418-4421). IEEE.

Arbabshirani, M. R., Plis, S., Sui, J., & Calhoun, V. D. (2017). Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage*, *145*, 137-165.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection.

Arnett, P. A., & Newman, J. P. (2000). Gray's three-arousal model: An empirical investigation. *Personality and Individual Differences*, *28*(6), 1171-1189.

Asensio, S., Romero, M. J., Palau, C., Sanchez, A., Senabre, I., Morales, J. L., ... & Romero, F. J. (2010). Altered neural response of the appetitive emotional system in cocaine addiction: an fMRI Study. *Addiction biology*, *15*(4), 504-516.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, *38*(1), 95-113.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *neuroimage*, *26*(3), 839-851.

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, *54*(3), 2033-2044.

Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTS). *Insight j*, *2*(365), 1-35.

Bachevalier, J., Meunier, M., Lu, M. X., & Ungerleider, L. G. (1997). Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys. *Experimental brain research*, *115*(3), 430-444.

Bandettini, P. A. (2012). The BOLD plot thickens: sign-and layer-dependent hemodynamic changes with activation. *Neuron*, *76*(3), 468-469.

Baranger, D. A., Lindenmuth, M., Nance, M., Guyer, A. E., Keenan, K., Hipwell, A. E., ... & Forbes, E. E. (2021). The longitudinal stability of fMRI activation during reward processing in adolescents and young adults. *NeuroImage*, *232*, 117872.

Barbas, H., García-Cabezas, M. Á., & Zikopoulos, B. (2013). Frontal-thalamic circuits associated with language. *Brain and language*, *126*(1), 49-61.

Barbas, H., Henion, T. H., & Dermon, C. R. (1991). Diverse thalamic projections to the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, *313*(1), 65-94.

Barber, A. D., Hegarty, C. E., Lindquist, M., & Karlsgodt, K. H. (2021). Heritability of functional connectivity in resting state: assessment of the dynamic mean, dynamic variance, and static connectivity across networks. *Cerebral Cortex*, *31*(6), 2834-2844.

Barker, E. D., Ing, A., Biondo, F., Jia, T., Pingault, J. B., Du Rietz, E., ... & IMAGEN Consortium. (2021). Do ADHD-impulsivity and BMI have shared polygenic and neural correlates?. *Molecular psychiatry*, *26*(3), 1019-1028.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. Trends in cognitive sciences, 17(10), 502-509.

Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1457), 1001-1013.

Beckmann, C. F., Mackay, C. E., Filippini, N., & Smith, S. M. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. *Neuroimage*, *47*(Suppl 1), S148.

Belin, D., Belin-Rauscent, A., Everitt, B. J., & Dalley, J. W. (2016). In search of predictive endophenotypes in addiction: insights from preclinical research. *Genes, Brain and Behavior*, *15*(1), 74-88.

Bell, M. D., Corbera, S., Johannesen, J. K., Fiszdon, J. M., & Wexler, B. E. (2013). Social cognitive impairments and negative symptoms in schizophrenia: are there subtypes with distinct functional correlates?. *Schizophrenia bulletin*, *39*(1), 186-196.

Bellec, P., Rosa-Neto, P., Lyttelton, O. C., Benali, H., & Evans, A. C. (2010). Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *Neuroimage*, *51*(3), 1126-1139.

Belotti, F., Deb, P., Manning, W. G., & Norton, E. C. (2015). twopm: Two-part models. *The Stata Journal*, *15*(1), 3-20.

Benjamin, K. J., Chen, Q., Jaffe, A. E., Stolz, J. M., Collado-Torres, L., Huuki-Myers, L. A., ... & Paquola, A. C. (2022). Analysis of the caudate nucleus transcriptome in individuals with schizophrenia highlights effects of antipsychotics and new risk genes. *Nature neuroscience*, *25*(11), 1559-1568.

Benjamini Y, & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. Journal of the Royal statistical society: series B (Methodological) 57(1):289-300.

Bennett, C. M., & Miller, M. B. (2013). fMRI reliability: Influences of task and experimental design. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 690-702.

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of open source software*, *5*(56), 2815.

Bentivoglio, M., Kultas-Ilinsky, K., & Ilinsky, I. (1993). Limbic thalamus: structure, intrinsic organization, and connections. *Neurobiology of cingulate cortex and limbic thalamus: A comprehensive handbook*, 71-122.

Bergamini, G., Sigrist, H., Ferger, B., Singewald, N., Seifritz, E., & Pryce, C. R. (2016). Depletion of nucleus accumbens dopamine leads to impaired reward and aversion processing in mice: relevance to motivation pathologies. *Neuropharmacology*, *109*, 306-319.

Bermingham, M. L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campbell, H., ... & Haley, C. S. (2015). Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Scientific reports*, *5*(1), 10312.

Berridge, K. C. (1996). Food reward: brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews*, *20*(1), 1-25.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391-431.

Berridge, K. C. (2009). 'Liking'and 'wanting'food rewards: brain substrates and roles in eating disorders. *Physiology & behavior*, *97*(5), 537-550.

Berridge, K. C., & Dayan, P. (2021). Liking. *Current Biology*, *31*(24), R1555-R1557.

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron*, *86*(3), 646-664.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?. *Brain research reviews*, *28*(3), 309-369.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in neurosciences*, *26*(9), 507-513.

Berridge, K. C., & Robinson, T. E. (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist*, *71*(8), 670.

Berridge, K. C., Venier, I. L., & Robinson, T. E. (1989). Taste reactivity analysis of 6-hydroxydopamine-induced aphagia: implications for arousal and anhedonia hypotheses of dopamine function. *Behavioral neuroscience*, *103*(1), 36.

Besle, J., Sánchez-Panchuelo, R. M., Bowtell, R., Francis, S., & Schluppeck, D. (2013). Single-subject fMRI mapping at 7 T of the representation of fingertips in S1: a comparison of event-related and phase-encoding designs. *Journal of neurophysiology*, *109*(9), 2293-2305.

Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 61-73.

Bialleck, K. A., Schaal, H. P., Kranz, T. A., Fell, J., Elger, C. E., & Axmacher, N. (2011). Ventromedial prefrontal cortex activation is associated with memory formation for predictable rewards. *PLoS one*, *6*(2), e16695.

Biswal, B. B., & Uddin, L. Q. (2025). The history and future of resting-state functional magnetic resonance imaging. *Nature*, *641*(8065), 1121-1131.

Biswal, B. B., Kylen, J. V., & Hyde, J. S. (1997). Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. *NMR in Biomedicine*, *10*(4-5), 165-170.

Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., ... & Milham, M. P. (2010). Toward discovery science of human brain function. *Proceedings of the national academy of sciences*, *107*(10), 4734-4739.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine*, *34*(4), 537-541.

Blumenfeld, R. S., Parks, C. M., Yonelinas, A. P., & Ranganath, C. (2011). Putting the pieces together: the role of dorsolateral prefrontal cortex in relational memory encoding. *Journal of cognitive neuroscience*, *23*(1), 257-265.

Bocchetta, M., Malpetti, M., Todd, E. G., Rowe, J. B., & Rohrer, J. D. (2021). Looking beneath the surface: the importance of subcortical structures in frontotemporal dementia. *Brain communications*, *3*(3), fcab158.

Bond, A., & Lader, M. (1974). The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*.

Boorman, L., Harris, S., Bruyns-Haylett, M., Kennerley, A., Zheng, Y., Martin, C., ... & Berwick, J. (2015). Long-latency reductions in gamma power predict hemodynamic changes that underlie the negative BOLD signal. *Journal of Neuroscience*, *35*(11), 4641-4656.

Boorman, L., Kennerley, A. J., Johnston, D., Jones, M., Zheng, Y., Redgrave, P., & Berwick, J. (2010). Negative blood oxygen level dependence in the rat: a model for investigating the role of suppression in neurovascular coupling. *Journal of Neuroscience*, *30*(12), 4285-4294.

Bossier H et al. (2020) The empirical replicability of task-based fMRI as a function of sample size. Neuroimage 212:116601.

Braak, H., & Braak, E. (1991). Alzheimer's disease affects limbic nuclei of the thalamus. *Acta neuropathologica*, *81*(3), 261-268.

Bradfield, L. A., Bertran-Gonzalez, J., Chieng, B., & Balleine, B. W. (2013). The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. *Neuron*, *79*(1), 153-166.

Bradfield, L. A., Hart, G., & Balleine, B. W. (2013). The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. *Frontiers in systems neuroscience*, *7*, 51.

Braga, R. M., Van Dijk, K. R., Polimeni, J. R., Eldaief, M. C., & Buckner, R. L. (2019). Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *Journal of neurophysiology*, *121*(4), 1513-1534.

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, *16*(3), 199-231.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*, *16*(2), S497.

Briand, L. A., Flagel, S. B., Seeman, P., & Robinson, T. E. (2008). Cocaine self-administration produces a persistent increase in dopamine D2High receptors. *European Neuropsychopharmacology*, *18*(8), 551-556.

Brodersen, K. H., Deserno, L., Schlagenhauf, F., Lin, Z., Penny, W. D., Buhmann, J. M., & Stephan, K. E. (2014). Dissecting psychiatric spectrum disorders by generative embedding. *NeuroImage: Clinical*, *4*, 98-111.

Brodt, S., Gais, S., Beck, J., Erb, M., Scheffler, K., & Schönauer, M. (2018). Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science*, *362*(6418), 1045-1048.

Brown, P. L., & Jenkins, H. M. (1968). AUTO-SHAPING OF THE PIGEON'S KEY-PECK 1. *Journal of the experimental analysis of behavior*, *11*(1), 1-8.

Brucar, L. R., Feczko, E., Fair, D. A., & Zilverstand, A. (2023). Current approaches in computational psychiatry for the data-driven identification of brain-based subtypes. *Biological psychiatry*, *93*(8), 704-716.

Buch, A. M., Vértes, P. E., Seidlitz, J., Kim, S. H., Grosenick, L., & Liston, C. (2023). Molecular and network-level mechanisms explaining individual differences in autism spectrum disorder. *Nature neuroscience*, *26*(4), 650-663.

Buckner, R. L., Krienen, F. M., & Yeo, B. T. (2013). Opportunities and limitations of intrinsic functional connectivity MRI. *Nature neuroscience*, *16*(7), 832-837.

Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., ... & Mintun, M. A. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *Journal of neuroscience*, *25*(34), 7709-7717.

Bunney, B. S., & Grace, A. A. (1978). Acute and chronic haloperidol treatment: comparison of effects on nigral dopaminergic cell activity. *Life sciences*, *23*(16), 1715-1727.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, *14*(5), 365-376.

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nature neuroscience*, *7*(5), 446-451.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, *68*(3), 362-385.

Buzsáki, G., Kaila, K., & Raichle, M. (2007). Inhibition and brain work. *Neuron*, *56*(5), 771-783.

Bymaster, F. P., Rasmussen, K., Calligaro, D. O., Nelson, D. L., DeLapp, N. W., Wong, D. T., & Moore, N. A. (1997). In vitro and in vivo biochemistry of olanzapine: a novel, atypical antipsychotic drug. *Journal of Clinical Psychiatry*, *58*(10), 28-36.

Byne, W., Hazlett, E. A., Buchsbaum, M. S., & Kemether, E. (2009). The thalamus and schizophrenia: current status of research. *Acta neuropathologica*, *117*(4), 347-368.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage*, *17*(3), 1394-1402.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, *16*(2), 317-330.

Calhoun, V. D., & Adali, T. (2012). Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE reviews in biomedical engineering*, *5*, 60-73.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human brain mapping*, *14*(3), 140-151.

Campus, P., Accoto, A., Maiolati, M., Latagliata, C., & Orsini, C. (2016). Role of prefrontal 5-HT in the strain-dependent variation in sign-tracking behavior of C57BL/6 and DBA/2 mice. *Psychopharmacology*, *233*(7), 1157-1169.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral cortex*, *12*(10), 1048-1056.

Cao, Z., Bennett, M., Orr, C., Icke, I., Banaschewski, T., Barker, G. J., ... & IMAGEN Consortium. (2019). Mapping adolescent reward anticipation, receipt, and prediction error during the monetary incentive delay task. *Human brain mapping*, *40*(1), 262-283.

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, *26*(3), 321-352.

Carl, H., Walsh, E., Eisenlohr-Moul, T., Minkel, J., Crowther, A., Moore, T., ... & Smoski, M. J. (2016). Sustained anterior cingulate cortex activation during reward processing predicts response to psychotherapy in major depressive disorder. *Journal of affective disorders*, *203*, 204-212.

Carlesimo, G. A., Lombardi, M. G., & Caltagirone, C. (2011). Vascular thalamic amnesia: a reappraisal. *Neuropsychologia*, *49*(5), 777-789.

Carlesimo, G. A., Lombardi, M. G., Caltagirone, C., & Barban, F. (2015). Recollection and familiarity in the human thalamus. *Neuroscience & Biobehavioral Reviews*, *54*, 18-28.

Carrera, E., & Bogousslavsky, J. (2006). The thalamus and behavior: effects of anatomically distinct strokes. *Neurology*, *66*(12), 1817-1823.

Carroll, L. S., & Owen, M. J. (2009). Genetic overlap between autism, schizophrenia and bipolar disorder. *Genome medicine*, *1*(10), 102.

Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental transfer: a review. *Neuroscience & Biobehavioral Reviews*, *71*, 829-848.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological psychology*, *54*(1-3), 241-257.

Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & cognition*, *36*(2), 429-437.

Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suárez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cerebral cortex*, *10*(3), 220-242.

Cetin, M. S., Christensen, F., Abbott, C. C., Stephen, J. M., Mayer, A. R., Cañive, J. M., ... & Calhoun, V. D. (2014). Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. *Neuroimage*, *97*, 117-126.

Chand, G. B., Dwyer, D. B., Erus, G., Sotiras, A., Varol, E., Srinivasan, D., ... & Davatzikos, C. (2020). Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain*, *143*(3), 1027-1038.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. Journal of statistical software, 61, 1-36.

Chen, G., Pine, D. S., Brotman, M. A., Smith, A. R., Cox, R. W., & Haller, S. P. (2021). Trial and error: A hierarchical modeling approach to test-retest reliability. *NeuroImage*, *245*, 118647.

Chen, Y., Chaudhary, S., & Li, C. S. R. (2022). Shared and distinct neural activity during anticipation and outcome of win and loss: A meta-analysis of the monetary incentive delay task. *Neuroimage*, *264*, 119764.

Cheng, B., Zhang, D., Chen, S., Kaufer, D. I., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2013). Semi-supervised multimodal relevance vector regression improves cognitive performance estimation from imaging and biological biomarkers. *Neuroinformatics*, *11*(3), 339-353.

Child, N. D., & Benarroch, E. E. (2013). Anterior nucleus of the thalamus: functional organization and clinical implications. *Neurology*, *81*(21), 1869-1876.

Chiodo, L. A., & Bunney, B. S. (1983). Typical and atypical neuroleptics: differential effects of chronic administration on the activity of A9 and A10 midbrain dopaminergic neurons. *Journal of Neuroscience*, *3*(8), 1607-1619.

Cho, Y. T., Moujaes, F., Schleifer, C. H., Starc, M., Ji, J. L., Santamauro, N., ... & Anticevic, A. (2022). Reward and loss incentives improve spatial working memory by shaping trial-by-trial posterior frontoparietal signals. *NeuroImage*, *254*, 119139.

Choi, U. S., Sung, Y. W., & Ogawa, S. (2022). Effects of Physiological Signal Removal on Resting-State Functional MRI Metrics. *Brain Sciences*, *13*(1), 8.

Chong, M., Bhushan, C., Joshi, A. A., Choi, S., Haldar, J. P., Shattuck, D. W., ... & Leahy, R. M. (2017). Individual parcellation of resting fMRI with a group functional connectivity prior. *NeuroImage*, *156*, 87-100.

Chow, J. J., Nickell, J. R., Darna, M., & Beckmann, J. S. (2016). Toward isolating the role of dopamine in the acquisition of incentive salience attribution. *Neuropharmacology*, *109*, 320-331.

Cipolotti, L., Husain, M., Crinion, J., Bird, C. M., Khan, S. S., Losseff, N., ... & Leff, A. P. (2008). The role of the thalamus in amnesia: a tractography, high-resolution MRI and neuropsychological study. *Neuropsychologia*, *46*(11), 2745-2758.

Clark, J. J., Collins, A. L., Sanford, C. A., & Phillips, P. E. (2013). Dopamine encoding of Pavlovian incentive stimuli diminishes with extended training. *Journal of Neuroscience*, *33*(8), 3526-3532.

Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D., ... & Tamminga, C. A. (2016). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry*, *173*(4), 373-384.

Cloninger, C. R., & Svrakic, D. M. (1997). Integrative psychobiological approach to psychiatric assessment and treatment. *Psychiatry*, *60*(2), 120-141.

Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of general psychiatry*, *50*(12), 975-990.

Cohen, A. O., Phaneuf, C. V., Rosenbaum, G. M., Glover, M. M., Avallone, K. N., Shen, X., & Hartley, C. A. (2022). Reward-motivated memories influence new learning across development. *Learning & Memory*, *29*(11), 421-429.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*(4466), 207-210.

Colaizzi, J. M., Flagel, S. B., Gearhardt, A. N., Borowitz, M. A., Kuplicki, R., Zotev, V., ... & Paulus, M. P. (2023). The propensity to sign-track is associated with externalizing behavior and distinct patterns of reward-related brain activation in youth. *Scientific Reports*, *13*(1), 4402.

Colaizzi, J. M., Flagel, S. B., Joyner, M. A., Gearhardt, A. N., Stewart, J. L., & Paulus, M. P. (2020). Mapping sign-tracking and goal-tracking onto human behaviors. *Neuroscience & Biobehavioral Reviews*, *111*, 84-94.

Colas, J. T., Dundon, N. M., Gerraty, R. T., Saragosa-Harris, N. M., Szymula, K. P., Tanwisuth, K., ... & O'Doherty, J. P. (2022). Reinforcement learning with associative or discriminative generalization across states and actions: fMRI at 3 T and 7 T. *Human Brain Mapping*, *43*(15), 4750-4790.

Cole, M. W., Anticevic, A., Repovs, G., & Barch, D. (2011). Variable global dysconnectivity and individual differences in schizophrenia. *Biological psychiatry*, *70*(1), 43-50.

Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature neuroscience*, *19*(12), 1718-1726.

Colom, M. (2023). *Neurobiological and behavioural studies of individual variation in cue-evoked motivation across rodents and humans*. Open University (United Kingdom).

Cope, L. M., Gheidi, A., Martz, M. E., Duval, E. R., Khalil, H., Allerton, T., & Morrow, J. D. (2023). A mechanical task for measuring sign-and goal-tracking in humans: a proof-of-concept study. *Behavioural brain research*, *436*, 114112.

Corbit, L. H., & Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *Journal of Neuroscience*, *31*(33), 11786-11794.

Cordes, D., Haughton, V. M., Arfanakis, K., Wendt, G. J., Turski, P. A., Moritz, C. H., ... & Meyerand, M. E. (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. *American journal of neuroradiology*, *21*(9), 1636-1644.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia*, *6*(3), 255-265.

Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R., Shanks, D. R., Robbins, T. W., ... & Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain*, *130*(9), 2387-2400.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, *29*(3), 162-173.

Cremers, H. R., Wager, T. D., & Yarkoni, T. (2017). The relation between statistical power and inference in fMRI. *PloS one*, *12*(11), e0184923.

Crossley, N. A., Mechelli, A., Vértes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., ... & Bullmore, E. T. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, *110*(28), 11583-11588.

Cui, Z., Pines, A. R., Larsen, B., Sydnor, V. J., Li, H., Adebimpe, A., ... & Satterthwaite, T. D. (2022). Linking individual differences in personalized functional network topography to psychopathology in youth. *Biological psychiatry*, *92*(12), 973-983.

da Silva Alves, F., Bakker, G., Schmitz, N., Abeling, N., Hasler, G., van der Meer, J., ... & van Amelsvoort, T. (2013). Dopaminergic modulation of the reward system in schizophrenia: a placebo-controlled dopamine depletion fMRI study. *European neuropsychopharmacology*, *23*(11), 1577-1586.

da Silva Alves, F., Schmitz, N., Figee, M., Abeling, N., Hasler, G., Van der Meer, J., ... & van Amelsvoort, T. (2011). Dopaminergic modulation of the human reward system: a placebo-controlled dopamine depletion fMRI study. *Journal of psychopharmacology*, *25*(4), 538-549.

Dalley, J. W., & Robbins, T. W. (2017). Fractionating impulsivity: neuropsychiatric implications. *Nature Reviews Neuroscience*, *18*(3), 158-171.

Damoiseaux JS, Beckmann C, Arigita ES, Barkhof F, Scheltens P, Stam C, Smith S, Rombouts S (2008) Reduced resting-state brain activity in the "default network" in normal aging. Cerebral cortex 18:1856-1864.

Damoiseaux, J. S., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, *103*(37), 13848-13853.

Danet, L., Barbeau, E. J., Eustache, P., Planton, M., Raposo, N., Sibon, I., ... & Pariente, J. (2015). Thalamic amnesia after infarct: the role of the mammillothalamic tract and mediodorsal nucleus. *Neurology*, *85*(24), 2107-2115.

Darvas, M., Wunsch, A. M., Gibbs, J. T., & Palmiter, R. D. (2014). Dopamine dependency for acquisition and performance of Pavlovian conditioned response. *Proceedings of the National Academy of Sciences*, *111*(7), 2764-2769.

Daum, I., & Ackermann, H. (1994). Frontal-type memory impairment associated with thalamic damage. *International Journal of Neuroscience*, *77*(3-4), 187-198.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current opinion in neurobiology*, *16*(6), 693-700.

Davis, J., Eyre, H., Jacka, F. N., Dodd, S., Dean, O., McEwen, S., ... & Berk, M. (2016). A review of vulnerability and risks for schizophrenia: Beyond the two hit hypothesis. *Neuroscience & Biobehavioral Reviews*, *65*, 185-194.

De Amorim, L. B., Cavalcanti, G. D., & Cruz, R. M. (2023). The choice of scaling technique matters for classification performance. *Applied Soft Computing*, *133*, 109924.

de la Rosa, N., Ress, D., Taylor, A. J., & Kim, J. H. (2021). Retinotopic variations of the negative blood-oxygen-level dependent hemodynamic response function in human primary visual cortex. *Journal of Neurophysiology*, *125*(4), 1045-1057.

De Wit, H. (2009). Impulsivity as a determinant and consequence of drug use: a review of underlying processes. *Addiction biology*, *14*(1), 22-31.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, *1104*(1), 70-88.

Demidenko, M. I., Mumford, J. A., & Poldrack, R. A. (2024). Impact of analytic decisions on test–retest reliability of individual and group estimates in functional magnetic resonance imaging: A multiverse analysis using the monetary incentive delay task. *Imaging Neuroscience*, *2*, 1-26.

Devi, R., Lepsien, J., Lorenz, K., Schlumm, T., Mildner, T., & Möller, H. E. (2022). Multi-echo investigations of positive and negative CBF and concomitant BOLD changes. *NeuroImage*, *263*, 119661.

Devor, A., Hillman, E. M., Tian, P., Waeber, C., Teng, I. C., Ruvinskaya, L., ... & Boas, D. A. (2008). Stimulus-induced changes in blood flow and 2-deoxyglucose uptake dissociate in ipsilateral somatosensory cortex. *Journal of Neuroscience*, *28*(53), 14347-14357.

Di Carlo, P., Pergola, G., Antonucci, L. A., Bonvino, A., Mancini, M., Quarto, T., ... & Blasi, G. (2020). Multivariate patterns of gray matter volume in thalamic nuclei are associated with positive schizotypy in healthy individuals. *Psychological medicine*, *50*(9), 1501-1509.

Dias, T. G. C., Iyer, S. P., Carpenter, S. D., Cary, R. P., Wilson, V. B., Mitchell, S. H., ... & Fair, D. A. (2015). Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. *Developmental cognitive neuroscience*, *11*, 155-174.

Dickinson, D., Pratt, D. N., Giangrande, E. J., Grunnagle, M., Orel, J., Weinberger, D. R., ... & Berman, K. F. (2018). Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. Schizophrenia bulletin, 44(1), 101-113.

Diederen, K. M., Ziauddeen, H., Vestergaard, M. D., Spencer, T., Schultz, W., & Fletcher, P. C. (2017). Dopamine modulates adaptive prediction error coding in the human midbrain and striatum. *Journal of Neuroscience*, *37*(7), 1708-1720.

DiFeliceantonio, A. G., & Berridge, K. C. (2012). Which cue to 'want'? Opioid stimulation of central amygdala makes goal-trackers show stronger goal-tracking, just as sign-trackers show stronger sign-tracking. *Behavioural brain research*, *230*(2), 399-408.

Dinu, L. M., Georgescu, A. L., Singh, S. N., Byrom, N. C., Overton, P. G., Singer, B. F., & Dommett, E. J. (2024). Sign-tracking and goal-tracking in humans: Utilising eye-tracking in clinical and non-clinical populations. *Behavioural brain research*, *461*, 114846.

Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion–cognition interactions: A review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, *23*(6), 669-694.

Dong, G., Li, H., Wang, Y., & Potenza, M. N. (2018). Individual differences in self-reported reward-approach tendencies relate to resting-state and reward-task-based fMRI measures. *International Journal of Psychophysiology*, *128*, 31-39.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, *104*(26), 11073-11078.

Doubovikov, E. D., & Aksenov, D. P. (2025). THE DIAGNOSTIC POTENTIAL OF RESTING STATE FUNCTIONAL MRI: STATISTICAL CONCERNS. *NeuroImage*, 121334.

Drew, P. J., Mateo, C., Turner, K. L., Yu, X., & Kleinfeld, D. (2020). Ultra-slow oscillations in fMRI and resting-state connectivity: neuronal and vascular contributions and technical confounds. *Neuron*, *107*(5), 782-804.

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., ... & Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature medicine*, *23*(1), 28-38.

Du, J., DiNicola, L. M., Angeli, P. A., Saadon-Grosman, N., Sun, W., Kaiser, S., ... & Buckner, R. L. (2024). Organization of the human cerebral cortex estimated within individuals: networks, global topography, and function. *Journal of neurophysiology*, *131*(6), 1014-1082.

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., ... & Alzheimer's Disease Neuroimaging Initiative. (2020). NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders. *NeuroImage: Clinical*, *28*, 102375.

Dubois, J., & Adolphs, R. (2016). Building a science of individual differences from fMRI. *Trends in cognitive sciences*, *20*(6), 425-443.

Duckworth, J. J., Wright, H., Christiansen, P., Rose, A. K., & Fallon, N. (2022). Sign-tracking modulates reward-related neural activation to reward cues, but not reward feedback. *European Journal of Neuroscience*, *56*(7), 5000-5013.

Dudai, Y., & Eisenberg, M. (2004). Rites of passage of the engram: reconsolidation and the lingering consolidation hypothesis. *Neuron*, *44*(1), 93-100.

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron*, *88*(1), 20-32.

Düzel, E., Bunzeck, N., Guitart-Masip, M., & Düzel, S. (2010). NOvelty-related motivation of anticipation and exploration by dopamine (NOMAD): implications for healthy aging. *Neuroscience & Biobehavioral Reviews*, *34*(5), 660-669.

Dworetsky, A., Seitzman, B. A., Adeyemo, B., Neta, M., Coalson, R. S., Petersen, S. E., & Gratton, C. (2021). Probabilistic mapping of human functional brain networks identifies regions of high group consensus. *Neuroimage*, *237*, 118164.

Dwyer, D. B., Cabral, C., Kambeitz-Ilankovic, L., Sanfelici, R., Kambeitz, J., Calhoun, V., ... & Koutsouleris, N. (2018). Brain subtyping enhances the neuroanatomical discrimination of schizophrenia. Schizophrenia bulletin, 44(5), 1060-1069.

Eichenbaum, H. (2004). Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron*, *44*(1), 109-120.

Eichenbaum, H. (2007). Comparative cognition, hippocampal function, and recollection. *Comparative Cognition & Behavior Reviews*, *2*.

Eichenbaum, H. (2010). Memory systems. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 478-490.

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, *18*(9), 547-558.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.*, *30*(1), 123-152.

Eichenbaum, Howard, and Neal J. Cohen, 'Hippocampal Function in Humans: *Insights From Amnesia and Functional Brain Imaging*', *From Conditioning to Conscious Recollection: Memory systems of the brain*, Oxford Psychology Series (New York, 2004; online edn, Oxford Academic, 1 Jan. 2008), https://doi.org/10.1093/acprof:oso/9780195178043.003.0005, accessed 3 Oct. 2025.

Eisenberger, N. I., Berkman, E. T., Inagaki, T. K., Rameson, L. T., Mashal, N. M., & Irwin, M. R. (2010). Inflammation-induced anhedonia: endotoxin reduces ventral striatum responses to reward. *Biological psychiatry*, *68*(8), 748-754.

Elfman, K. W., Parks, C. M., & Yonelinas, A. P. (2008). Testing a neurocomputational model of recollection, familiarity, and source recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 752.

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., ... & Hariri, A. R. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological science*, *31*(7), 792-806.

Enkel, T., Bartsch, D., & Bähner, F. (2019). Sign-and goal-tracking rats show differences in various executive functions. *Behavioural Brain Research*, *371*, 111979.

Ernst, M. O. N. I. Q. U. E., & Spear, L. P. (2009). Reward systems. *Handbook of developmental social neuroscience*, 324-341.

Esslinger, C., Englisch, S., Inta, D., Rausch, F., Schirmbeck, F., Mier, D., ... & Zink, M. (2012). Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophrenia research*, *140*(1-3), 114-121.

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS one*, *12*(9), e0184661.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature methods*, *16*(1), 111-116.

Fair, D., Bathula, D., Mills, K. L., Costa Dias, T. G., Blythe, M. S., Zhang, D., ... & Nagel, B. J. (2010). Maturing thalamocortical functional connectivity across development. *Frontiers in systems neuroscience*, *4*, 1422.

Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences*, *109*(17), 6769-6774.

Fama, R., & Sullivan, E. V. (2015). Thalamic structures and associated cognitive functions: Relations with age and aging. *Neuroscience and biobehavioral reviews*, *54*, 29.

Farmer, R. F., & Goldberg, L. R. (2008). A psychometric evaluation of the revised Temperament and Character Inventory (TCI-R) and the TCI-140. *Psychological assessment*, *20*(3), 281.

Fedorenko, E. (2021). The early origins and the growing popularity of the individual-subject analytic approach in human neuroscience. *Current Opinion in Behavioral Sciences*, *40*, 105-112.

Fedorenko, E., & Blank, I. A. (2020). Broca's area is not a natural kind. *Trends in cognitive sciences*, *24*(4), 270-284.

Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of neurophysiology*, *104*(2), 1177-1194.

Feilong, M., Nastase, S. A., Guntupalli, J. S., & Haxby, J. V. (2018). Reliable individual differences in fine-grained cortical functional architecture. *NeuroImage*, *183*, 375-386.

Felix, P. C., & Flagel, S. B. (2024). Leveraging individual differences in cue–reward learning to investigate the psychological and neural basis of shared psychiatric symptomatology: The sign-tracker/goal-tracker model. *Behavioral Neuroscience*, *138*(4), 260.

Ferbinteanu, J. (2019). Memory systems 2018–Towards a new paradigm. *Neurobiology of Learning and Memory*, *157*, 61-78.

Ferrucci, R., Brunoni, A. R., Parazzini, M., Vergari, M., Rossi, E., Fumagalli, M., ... & Priori, A. (2013). Modulating human procedural learning by cerebellar transcranial direct current stimulation. *The Cerebellum*, *12*(4), 485-492.

Field, M., & Cox, W. M. (2008). Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug and alcohol dependence*, *97*(1-2), 1-20.

Finn, E. S., & Todd Constable, R. (2016). Individual variation in functional brain connectivity: implications for personalized approaches to psychiatric disease. *Dialogues in clinical neuroscience*, *18*(3), 277-287.

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... & Constable, R. T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience*, *18*(11), 1664-1671.

Fitzpatrick, C. J., & Morrow, J. D. (2020). Individual variation in the attribution of incentive salience to social cues. *Scientific Reports*, *10*(1), 2583.

Flagel, S. B., & Robinson, T. E. (2017). Neurobiological basis of individual variation in stimulus-reward learning. *Current Opinion in Behavioral Sciences*, *13*, 178-185.

Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, *56*, 139-148.

Flagel, S. B., Cameron, C. M., Pickup, K. N., Watson, S. J., Akil, H., & Robinson, T. E. (2011). A food predictive cue must be attributed with incentive salience for it to induce c-fos mRNA expression in cortico-striatal-thalamic brain regions. *Neuroscience*, *196*, 80-96.

Flagel, S. B., Chaudhury, S., Waselus, M., Kelly, R., Sewani, S., Clinton, S. M., ... & Akil, H. (2016). Genetic background and epigenetic modifications in the core of the nucleus accumbens predict addiction-like behavior in a rat model. *Proceedings of the National Academy of Sciences*, *113*(20), E2861-E2870.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., ... & Akil, H. (2011). A selective role for dopamine in stimulus–reward learning. *Nature*, *469*(7328), 53-57.

Flagel, S. B., Robinson, T. E., Clark, J. J., Clinton, S. M., Watson, S. J., Seeman, P., ... & Akil, H. (2010). An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. *Neuropsychopharmacology*, *35*(2), 388-400.

Flagel, S. B., Watson, S. J., Akil, H., & Robinson, T. E. (2008). Individual differences in the attribution of incentive salience to a reward-related cue: influence on cocaine sensitization. *Behavioural brain research*, *186*(1), 48-56.

Flagel, S. B., Watson, S. J., Robinson, T. E., & Akil, H. (2007). Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology*, *191*(3), 599-607.

Flournoy, J. C., Bryce, N. V., Dennison, M. J., Rodman, A. M., McNeilly, E. A., Lurie, L. A., ... & McLaughlin, K. A. (2024). A precision neuroscience approach to estimating reliability of neural responses during emotion processing: Implications for task-fMRI. *NeuroImage*, *285*, 120503.

Fornito, A., Yoon, J., Zalesky, A., Bullmore, E. T., & Carter, C. S. (2011). General and specific functional connectivity disturbances in first-episode schizophrenia during cognitive control performance. *Biological psychiatry*, *70*(1), 64-72.

Foroni, F., Pergola, G., Argiris, G., & Rumiati, R. I. (2013). The FoodCast research image database (FRIDa). *Frontiers in human neuroscience*, *7*, 51.

Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in systems neuroscience*, *4*, 1443.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, *8*(9), 700-711.

Fox, M. D., Buckner, R. L., Liu, H., Chakravarty, M. M., Lozano, A. M., & Pascual-Leone, A. (2014). Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. *Proceedings of the National Academy of Sciences*, *111*(41), E4367-E4375.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046-10051.

Fox, M. D., Halko, M. A., Eldaief, M. C., & Pascual-Leone, A. (2012). Measuring and manipulating brain connectivity with resting state functional connectivity magnetic resonance imaging (fcMRI) and transcranial magnetic stimulation (TMS). *Neuroimage*, *62*(4), 2232-2243.

Fox, M. D., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2007). Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron*, *56*(1), 171-184.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673-9678.

Franken, I. H., Booij, J., & van den Brink, W. (2005). The role of dopamine in human addiction: from reward to motivated attention. *European journal of pharmacology*, *526*(1-3), 199-206.

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature reviews neuroscience*, *6*(2), 119-130.

Franklin, T. R., Wetherill, R. R., Jagannathan, K., Hager, N., O'Brien, C. P., & Childress, A. R. (2015). Limitations of the use of the MP-RAGE to identify neural changes in the brain: recent cigarette smoking alters gray matter indices in the striatum. *Frontiers in human neuroscience*, *8*, 1052.

Fraser, K. M., Haight, J. L., Gardner, E. L., & Flagel, S. B. (2016). Examining the role of dopamine D2 and D3 receptors in Pavlovian conditioned approach behaviors. *Behavioural brain research*, *305*, 87-99.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325-1352.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, *2*(1-2), 56-78.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. (1995). Spatial registration and normalization of images. *Human brain mapping*, *3*(3), 165-189.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow & Metabolism*, *13*(1), 5-14.

Fritsch, V., Da Mota, B., Loth, E., Varoquaux, G., Banaschewski, T., Barker, G. J., ... & IMAGEN consortium. (2015). Robust regression for large-scale neuroimaging studies. *Neuroimage*, *111*, 431-441.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of experimental psychology: General*, *141*(1), 2.

Funahashi, S. (2013). Thalamic mediodorsal nucleus and its participation in spatial working memory processes: comparison with the prefrontal cortex. *Frontiers in systems neuroscience*, *7*, 36.

Fusar-Poli, P., & Meyer-Lindenberg, A. (2013). Striatal presynaptic dopamine in schizophrenia, part II: meta-analysis of [18F/11C]-DOPA PET studies. *Schizophrenia bulletin*, *39*(1), 33-42.

Gao, S., Calhoun, V. D., & Sui, J. (2018). Machine learning in major depression: From classification to treatment outcome prediction. *CNS neuroscience & therapeutics*, *24*(11), 1037-1052.

Garbusow, M., Ebrahimi, C., Riemerschmid, C., Daldrup, L., Rothkirch, M., Chen, K., ... & Rapp, M. A. (2022). Pavlovian-to-instrumental transfer across mental disorders: a review. *Neuropsychobiology*, *81*(5), 418-437.

Garbusow, M., Schad, D. J., Sommer, C., Jünger, E., Sebold, M., Friedel, E., ... & Rapp, M. A. (2014). Pavlovian-to-instrumental transfer in alcohol dependence: a pilot study. *Neuropsychobiology*, *70*(2), 111-121.

Gardner, D. M., Murphy, A. L., O'Donnell, H., Centorrino, F., & Baldessarini, R. J. (2010). International consensus study of antipsychotic dosing. *American Journal of Psychiatry*, *167*(6), 686-693.

Garofalo, S., & di Pellegrino, G. (2015). Individual differences in the influence of task-irrelevant Pavlovian cues on human behavior. *Frontiers in behavioral neuroscience*, *9*, 163.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. Neuroscience & Biobehavioral Reviews, 37(7), 1297-1310.

Gates, K. M., Molenaar, P. C., Iyer, S. P., Nigg, J. T., & Fair, D. A. (2014). Organizing heterogeneous samples using community detection of GIMME-derived resting state functional networks. *PloS one*, *9*(3), e91322.

Ge, T., Holmes, A. J., Buckner, R. L., Smoller, J. W., & Sabuncu, M. R. (2017). Heritability analysis with repeat measurements and its application to resting-state functional connectivity. *Proceedings of the National Academy of Sciences*, *114*(21), 5521-5526.

Geerligs L, Henson RN (2016) Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. NeuroImage 135:16-31.

Geier, K. T. (2023). *Role of the anterior and mediodorsal thalamic nuclei in cognitive processes*. University of Toronto (Canada).

Geier, K. T., Buchsbaum, B. R., Parimoo, S., & Olsen, R. K. (2020). The role of anterior and medial dorsal thalamus in associative memory encoding and retrieval. *Neuropsychologia*, *148*, 107623.

Genon, S., Eickhoff, S. B., & Kharabian, S. (2022). Linking interindividual variability in brain structure to behaviour. *Nature Reviews Neuroscience*, *23*(5), 307-318.

Genon, S., Reid, A., Langner, R., Amunts, K., & Eickhoff, S. B. (2018). How to characterize the function of a brain region. *Trends in cognitive sciences*, *22*(4), 350-364.

Georgescu, I. A., Popa, D., & Zagrean, L. (2020). The anatomical and functional heterogeneity of the mediodorsal thalamus. *Brain Sciences*, *10*(9), 624.

Ghafarimoghadam, M., Mashayekh, R., Gholami, M., Fereydani, P., Shelley-Tremblay, J., Kandezi, N., ... & Motaghinejad, M. (2022). A review of behavioral methods for the evaluation of cognitive performance in animal models: current techniques and links to human cognition. *Physiology & behavior*, *244*, 113652.

Ghika-Schmid, F., & Bogousslavsky, J. (2000). The acute behavioral syndrome of anterior thalamic infarction: a prospective study of 12 cases. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *48*(2), 220-227.

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104-114.

Gillis, Z. S., & Morrison, S. E. (2019). Sign tracking and goal tracking are characterized by distinct patterns of nucleus accumbens activity. *ENeuro*, *6*(2).

Giraldo-Chica, M., Rogers, B. P., Damon, S. M., Landman, B. A., & Woodward, N. D. (2018). Prefrontal-thalamic anatomical connectivity and executive cognitive function in schizophrenia. *Biological psychiatry*, *83*(6), 509-517.

Glahn, D. C., Winkler, A. M., Kochunov, P., Almasy, L., Duggirala, R., Carless, M. A., ... & Blangero, J. (2010). Genetic control over the resting brain. *Proceedings of the National Academy of Sciences*, *107*(3), 1223-1228.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171-178.

Goense, J., Merkle, H., & Logothetis, N. K. (2012). High-resolution fMRI reveals laminar differences in neurovascular coupling between positive and negative BOLD responses. *Neuron*, *76*(3), 629-639.

Golby, A. J., Poldrack, R. A., Brewer, J. B., Spencer, D., Desmond, J. E., Aron, A. P., & Gabrieli, J. D. (2001). Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain*, *124*(9), 1841-1854.

Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: a deficit in the representation of value. *Schizophrenia bulletin*, *34*(5), 835-847.

Goldman-Rakic, P. S., & Porrino, L. J. (1985). The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *Journal of Comparative Neurology*, *242*(4), 535-560.

Gordon, E. M., & Nelson, S. M. (2021). Three types of individual variation in brain networks revealed by single-subject functional connectivity analyses. *Current Opinion in Behavioral Sciences*, *40*, 79-86.

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., ... & Dosenbach, N. U. (2017). Precision functional mapping of individual human brains. *Neuron*, *95*(4), 791-807.

Gordon, E. M., Laumann, T. O., Marek, S., Raut, R. V., Gratton, C., Newbold, D. J., ... & Nelson, S. M. (2020). Default-mode network streams for coupling to language and control systems. *Proceedings of the National Academy of Sciences*, *117*(29), 17308-17319.

Grace, A. A., Bunney, B. S., Moore, H., & Todd, C. L. (1997). Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. *Trends in neurosciences*, *20*(1), 31-37.

Gradin, V. B., Waiter, G., O'Connor, A., Romaniuk, L., Stickle, C., Matthews, K., ... & Steele, J. D. (2013). Salience network-midbrain dysconnectivity and blunted reward signals in schizophrenia. *Psychiatry Research: Neuroimaging*, *211*(2), 104-111.

Gräff J, Joseph NF, Horn ME, Samiei A, Meng J, Seo J, Rei D, Bero AW, Phan TX, Wagner F (2014) Epigenetic priming of memory updating during reconsolidation to attenuate remote fear memories. Cell 156:261-276.

Gratton, C., Kraus, B. T., Greene, D. J., Gordon, E. M., Laumann, T. O., Nelson, S. M., ... & Petersen, S. E. (2020). Defining individual-specific functional neuroanatomy for precision psychiatry. *Biological psychiatry*, *88*(1), 28-39.

Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ... & Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*, *98*(2), 439-452.

Greene, A. S., Gao, S., Scheinost, D., & Constable, R. T. (2018). Task-induced brain state manipulation improves prediction of individual traits. *Nature communications*, *9*(1), 2807.

Greene, D. J., Marek, S., Gordon, E. M., Siegel, J. S., Gratton, C., Laumann, T. O., ... & Dosenbach, N. U. (2020). Integrative and network-specific connectivity of the basal ganglia and thalamus defined in individuals. *Neuron*, *105*(4), 742-758.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the national academy of sciences*, *100*(1), 253-258.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proceedings of the National Academy of Sciences*, *101*(13), 4637-4642.

Grimm, O., Heinz, A., Walter, H., Kirsch, P., Erk, S., Haddad, L., ... & Meyer-Lindenberg, A. (2014). Striatal response to reward anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. *JAMA psychiatry*, *71*(5), 531-539.

Guadagno, A., Belliveau, C., Mechawar, N., & Walker, C. D. (2021). Effects of early life stress on the developing basolateral amygdala-prefrontal cortex circuit: the emerging role of local inhibition and perineuronal nets. *Frontiers in human neuroscience*, *15*, 669120.

Guntupalli, J. S., Hanke, M., Halchenko, Y. O., Connolly, A. C., Ramadge, P. J., & Haxby, J. V. (2016). A model of representational spaces in human cortex. *Cerebral cortex*, *26*(6), 2919-2934.

Gunzler, D., Chen, T., Wu, P., & Zhang, H. (2013). Introduction to mediation analysis with structural equation modeling. *Shanghai archives of psychiatry*, *25*(6), 390.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(7), 4259-4264.

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, *35*(1), 4-26.

Hacker, C. D., Laumann, T. O., Szrama, N. P., Baldassarre, A., Snyder, A. Z., Leuthardt, E. C., & Corbetta, M. (2013). Resting state network estimation in individual subjects. *Neuroimage*, *82*, 616-633.

Haight, J. L., Fuller, Z. L., Fraser, K. M., & Flagel, S. B. (2017). A food-predictive cue attributed with incentive salience engages subcortical afferents and efferents of the paraventricular nucleus of the thalamus. *Neuroscience*, *340*, 135-152.

Halassa, M. M., & Kastner, S. (2017). Thalamic functions in distributed cognitive control. *Nature neuroscience*, *20*(12), 1669-1679.

Halassa, M. M., & Sherman, S. M. (2019). Thalamocortical circuit motifs: a general framework. *Neuron*, *103*(5), 762-770.

Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Brain connectivity related to working memory performance. *Journal of Neuroscience*, *26*(51), 13338-13343.

Harel, N., Lee, S. P., Nagaoka, T., Kim, D. S., & Kim, S. G. (2002). Origin of negative blood oxygenation level—dependent fMRI signals. *Journal of cerebral blood flow & metabolism*, *22*(8), 908-917.

Harrewijn, A., Abend, R., Linke, J., Brotman, M. A., Fox, N. A., Leibenluft, E., ... & Pine, D. S. (2020). Combining fMRI during resting state and an attention bias task in children. *NeuroImage*, *205*, 116301.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, *28*(1), 100-108.

Hawkins, P. C., Wood, T. C., Vernon, A. C., Bertolino, A., Sambataro, F., Dukart, J., ... & Mehta, M. A. (2018). An investigation of regional cerebral blood flow and tissue structure changes after acute administration of antipsychotics in healthy male volunteers. *Human Brain Mapping*, *39*(1), 319-331.

Hawkins, P. C., Zelaya, F. O., O'Daly, O., Holiga, S., Dukart, J., Umbricht, D., & Mehta, M. A. (2021). The effect of risperidone on reward-related brain activity is robust to drug-induced vascular changes. *Human brain mapping*, *42*(9), 2766-2777.

Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., ... & Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, *72*(2), 404-416.

Heaton, R. K., & Pendleton, M. G. (1981). Use of Neuropsychological tests to predict adult patients' everyday functioning. *Journal of consulting and clinical psychology*, *49*(6), 807.

Heck, M., Durieux, N., Anselme, P., & Quertemont, E. (2025). Implementations of sign-and goal-tracking behavior in humans: A scoping review. *Cognitive, Affective, & Behavioral Neuroscience*, *25*(2), 263-290.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, *50*(3), 1166-1186.

Heinz, A. (2002). Dopaminergic dysfunction in alcoholism and schizophrenia–psychopathological and behavioral correlates. *European Psychiatry*, *17*(1), 9-16.

Heinz, A., & Schlagenhauf, F. (2010). Dopaminergic dysfunction in schizophrenia: salience attribution revisited. *Schizophrenia bulletin*, *36*(3), 472-485.

Hélie, S., Shamloo, F., Novak, K., & Foti, D. (2017). The roles of valuation and reward processing in cognitive function and psychiatric disorders. *Annals of the New York Academy of Sciences*, *1395*(1), 33-48.

Henderson, J. M., Carpenter, K., Cartwright, H., & Halliday, G. M. (2000). Loss of thalamic intralaminar nuclei in progressive supranuclear palsy and Parkinson's disease: clinical and therapeutic implications. *Brain*, *123*(7), 1410-1421.

Henke, K., Buck, A., Weber, B., & Wieser, H. G. (1997). Human hippocampus establishes associations in memory. *Hippocampus*, *7*(3), 249-256.

Herbert, M. I. C. H. A. E. L., Johns, M. W., & Doré, C. (1976). Factor analysis of analogue scales measuring subjective feelings before and after sleep. *Br J Med Psychol*, *49*(4), 373-379.

Hermes, D., Miller, K. J., Vansteensel, M. J., Aarnoutse, E. J., Leijten, F. S., & Ramsey, N. F. (2012). Neurophysiologic correlates of fMRI in human motor cortex. *Human brain mapping*, *33*(7), 1689-1699.

Hermosillo, R. J., Moore, L. A., Feczko, E., Miranda-Domínguez, Ó., Pines, A., Dworetsky, A., ... & Fair, D. A. (2024). A precision functional atlas of personalized network topography and probabilities. *Nature neuroscience*, *27*(5), 1000-1013.

Herrero, M. T., Barcia, C., & Navarro, J. (2002). Functional anatomy of thalamus and basal ganglia. *Child's Nervous System*, *18*(8), 386-404.

Holiga, Š., Sambataro, F., Luzy, C., Greig, G., Sarkar, N., Renken, R. J., ... & Dukart, J. (2018). Test-retest reliability of task-based and resting-state blood oxygen level dependence and cerebral blood flow measures. *PloS one*, *13*(11), e0206583.

Hollingshead, A. B., & Frederick, C. (1964). Social class and mental illness: A community study. *American Behavioral Scientist*, *8*(3), 30-30.

Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neuroscience & Biobehavioral Reviews*, *34*(8), 1277-1295.

Howes, O. D., & Murray, R. M. (2014). Schizophrenia: an integrated sociodevelopmental-cognitive model. *The Lancet*, *383*(9929), 1677-1687.

Howes, O. D., Hird, E. J., Adams, R. A., Corlett, P. R., & McGuire, P. (2020). Aberrant salience, information processing, and dopaminergic signaling in people at clinical high risk for psychosis. *Biological psychiatry*, *88*(4), 304-314.

Howes, O. D., Kambeitz, J., Kim, E., Stahl, D., Slifstein, M., Abi-Dargham, A., & Kapur, S. (2012). The nature of dopamine dysfunction in schizophrenia and what this means for treatment: meta-analysis of imaging studies. *Archives of general psychiatry*, *69*(8), 776-786.

Howes, O. D., Williams, M., Ibrahim, K., Leung, G., Egerton, A., McGuire, P. K., & Turkheimer, F. (2013). Midbrain dopamine function in schizophrenia and depression: a post-mortem and positron emission tomographic imaging study. *Brain*, *136*(11), 3242-3251.

Hreib, K. K., Rosene, D. L., & Moss, M. B. (1988). Basal forebrain efferents to the medial dorsal thalamic nucleus in the rhesus monkey. *Journal of Comparative Neurology*, *277*(3), 365-390.

Hsu, W. T., Rosenberg, M. D., Scheinost, D., Constable, R. T., & Chun, M. M. (2018). Resting-state functional connectivity predicts neuroticism and extraversion in novel individuals. Social cognitive and affective neuroscience, 13(2), 224-232.

Hu, D., & Huang, L. (2015). Negative hemodynamic response in the cortex: evidence opposing neuronal deactivation revealed via optical imaging and electrophysiological recording. *Journal of Neurophysiology*, *114*(4), 2152-2161.

Huber, L., Goense, J., Kennerley, A. J., Ivanov, D., Krieger, S. N., Lepsien, J., ... & Möller, H. E. (2014). Investigation of the neurovascular coupling in positive and negative BOLD responses in human brain at 7 T. *Neuroimage*, *97*, 349-362.

Huttunen, J., Heinimaa, M., Svirskis, T., Nyman, M., Kajander, J., Forsback, S., ... & Hietala, J. (2008). Striatal dopamine synthesis in first-degree relatives of patients with schizophrenia. *Biological psychiatry*, *63*(1), 114-117.

Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature neuroscience*, *19*(3), 404-413.

Hwang, K., Bertolero, M. A., Liu, W. B., & D'Esposito, M. (2017). The human thalamus is an integrative hub for functional brain networks. *Journal of Neuroscience*, *37*(23), 5594-5607.

Hwang, K., Shine, J. M., Cole, M. W., & Sorenson, E. (2022). Thalamocortical contributions to cognitive task activity. *Elife*, *11*, e81282.

Iglesias, A. G., Chiu, A. S., Wong, J., Campus, P., Li, F., Bhatti, J. K., ... & Flagel, S. B. (2023). Inhibition of dopamine neurons prevents incentive value encoding of a reward cue: With revelations from deep phenotyping. *Journal of Neuroscience*, *43*(44), 7376-7392.

Ilinsky, I. A., Jouandet, M. L., & Goldman-Rakic, P. S. (1985). Organization of the nigrothalamocortical system in the rhesus monkey. *Journal of Comparative Neurology*, *236*(3), 315-330.

Iraji, A., Calhoun, V. D., Wiseman, N. M., Davoodi-Bojd, E., Avanaki, M. R., Haacke, E. M., & Kou, Z. (2016). The connectivity domain: Analyzing resting state fMRI data using feature-based data-driven and model-based methods. *Neuroimage*, *134*, 494-507.

Iraji, A., Deramus, T. P., Lewis, N., Yaesoubi, M., Stephen, J. M., Erhardt, E., ... & Calhoun, V. D. (2019). The spatial chronnectome reveals a dynamic interplay between functional segregation and integration. *Human brain mapping*, *40*(10), 3058-3077.

Jankowski, M. M., Ronnqvist, K. C., Tsanov, M., Vann, S. D., Wright, N. F., Erichsen, J. T., ... & O'mara, S. M. (2013). The anterior thalamus provides a subcortical circuit supporting memory and spatial navigation. *Frontiers in systems neuroscience*, *7*, 45.

Janssen, R. J., Mourão-Miranda, J., & Schnack, H. G. (2018). Making individual prognoses in psychiatry using neuroimaging and machine learning. Biological psychiatry: Cognitive neuroscience and neuroimaging, 3(9), 798-808.

Jauhar, S., Fortea, L., Solanes, A., Albajes-Eizagirre, A., McKenna, P. J., & Radua, J. (2021). Brain activations associated with anticipation and delivery of monetary reward: A systematic review and meta-analysis of fMRI studies. *PloS one*, *16*(8), e0255292.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, *5*(2), 143-156.

Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, *185*, 35-57.

Johansen-Berg, H., Behrens, T. E., Sillery, E., Ciccarelli, O., Thompson, A. J., Smith, S. M., & Matthews, P. M. (2005). Functional–anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cerebral cortex*, *15*(1), 31-39.

Johnson, R. A., & Wichern, D. W. (2002). Applied multivariate statistical analysis.

Johnston, B. A., Steele, J. D., Tolomeo, S., Christmas, D., & Matthews, K. (2015). Structural MRI-based predictions in patients with treatment-refractory depression (TRD). *PloS one*, *10*(7), e0132958.

Jollans, L., Zhipeng, C., Icke, I., Greene, C., Kelly, C., Banaschewski, T., ... & Whelan, R. (2016). Ventral striatum connectivity during reward anticipation in adolescent smokers. *Developmental neuropsychology*, *41*(1-2), 6-21.

Jones, E. G. (2012). *The thalamus*. Springer Science & Business Media.

Joyner, M. (2019). *Examining Behavioral Phenotypes of Overeating and Obesity: Environmental, Psychological, and Neurobiological Influences on Food Motivation and Palatable Food Consumption* (Doctoral dissertation).

Joyner, M. A., Gearhardt, A. N., & Flagel, S. B. (2017). A translational model to assess sign-tracking and goal-tracking behavior in children. *Neuropsychopharmacology*, *43*(1), 228.

Juckel, G. (2016). Inhibition of the reward system by antipsychotic treatment. *Dialogues in Clinical Neuroscience*, *18*(1), 109-114.

Juckel, G., Schlagenhauf, F., Koslowski, M., Filonov, D., Wüstenberg, T., Villringer, A., ... & Heinz, A. (2006). Dysfunction of ventral striatal reward prediction in schizophrenic patients treated with typical, not atypical, neuroleptics. *Psychopharmacology*, *187*(2), 222-228.

Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., ... & Heinz, A. (2006). Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage*, *29*(2), 409-416.

Kambeitz, J., Kambeitz-Ilankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., ... & Koutsouleris, N. (2015). Detecting neuroimaging biomarkers for schizophrenia: a meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology*, *40*(7), 1742-1751.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*(4), 231-242.

Kannurpatti, S. S., & Biswal, B. B. (2004). Negative functional response to sensory stimulation and its origins. *Journal of Cerebral Blood Flow & Metabolism*, *24*(6), 703-712.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*, *17*(11), 4302-4311.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American journal of Psychiatry*, *160*(1), 13-23.

Karalunas, S. L., Fair, D., Musser, E. D., Aykes, K., Iyer, S. P., & Nigg, J. T. (2014). Subtyping attention-deficit/hyperactivity disorder using temperament dimensions: toward biologically based nosologic criteria. *JAMA psychiatry*, *71*(9), 1015-1024.

Karoly, H. C., Bryan, A. D., Weiland, B. J., Mayer, A., Dodd, A., & Ewing, S. W. F. (2015). Does incentive-elicited nucleus accumbens activation differ by substance of abuse? An examination with adolescents. *Developmental cognitive neuroscience*, *16*, 5-15.

Katthagen, T., Mathys, C., Deserno, L., Walter, H., Kathmann, N., Heinz, A., & Schlagenhauf, F. (2018). Modeling subjective relevance in schizophrenia and its relation to aberrant salience. *PLoS computational biology*, *14*(8), e1006319.

Keller, A. S., Pines, A. R., Shanmugan, S., Sydnor, V. J., Cui, Z., Bertolero, M. A., ... & Satterthwaite, T. D. (2023). Personalized functional brain network topography is associated with individual differences in youth cognition. *Nature communications*, *14*(1), 8411.

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., ... & Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron*, *20*(5), 927-936.

Ketz, N. A., Jensen, O., & O'Reilly, R. C. (2015). Thalamic pathways underlying prefrontal cortex–medial temporal lobe oscillatory interactions. *Trends in neurosciences*, *38*(1), 3-12.

Khandaker, G. M., Barnett, J. H., White, I. R., & Jones, P. B. (2011). A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. *Schizophrenia research*, *132*(2-3), 220-227.

Kherif, F., Josse, G., Seghier, M. L., & Price, C. J. (2009). The main sources of intersubject variability in neuronal activation for reading aloud. *Journal of Cognitive Neuroscience*, *21*(4), 654-668.

Kiar, G., Mumford, J. A., Xu, T., Vogelstein, J. T., Glatard, T., & Milham, M. P. (2024). Why experimental variation in neuroimaging should be embraced. *Nature communications*, *15*(1), 9411.

Klein, J. C., Rushworth, M. F., Behrens, T. E., Mackay, C. E., de Crespigny, A. J., D'Arceuil, H., & Johansen-Berg, H. (2010). Topography of connections between human prefrontal cortex and mediodorsal thalamus studied with diffusion tractography. *Neuroimage*, *51*(2), 555-564.

Klingner, C. M., Hasler, C., Brodoehl, S., & Witte, O. W. (2010). Dependence of the negative BOLD response on somatosensory stimulus intensity. *Neuroimage*, *53*(1), 189-195.

Klöppel, S., Abdulkadir, A., Jack Jr, C. R., Koutsouleris, N., Mourão-Miranda, J., & Vemuri, P. (2012). Diagnostic neuroimaging across diseases. *Neuroimage*, *61*(2), 457-463.

Knowlton, B. J., & Castel, A. D. (2022). Memory and reward-based learning: A value-directed remembering perspective. *Annual review of psychology*, *73*(1), 25-52.

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1511), 3771-3786.

Knutson, B., Bjork, J. M., Fong, G. W., Hommer, D., Mattay, V. S., & Weinberger, D. R. (2004). Amphetamine modulates human incentive processing. *Neuron*, *43*(2), 261-269.

Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, *12*(17), 3683-3687.

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage*, *12*(1), 20-27.

Koch, S., Perry, K. W., & Bymaster, F. P. (2004). Brain region and dose effects of an olanzapine/fluoxetine combination on extracellular monoamine concentrations in the rat. *Neuropharmacology*, *46*(2), 232-242.

Koenen, K. C., Moffitt, T. E., Roberts, A. L., Martin, L. T., Kubzansky, L., Harrington, H., ... & Caspi, A. (2009). Childhood IQ and adult mental disorders: a test of the cognitive reserve hypothesis. *American Journal of Psychiatry, 166*(1), 50-57.

Kong, R., Li, J., Orban, C., Sabuncu, M. R., Liu, H., Schaefer, A., ... & Yeo, B. T. (2019). Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cerebral cortex*, *29*(6), 2533-2551.

Koush, Y., de Graaf, R. A., Kupers, R., Dricot, L., Ptito, M., Behar, K. L., ... & Hyder, F. (2021). Metabolic underpinnings of activated and deactivated cortical areas in human brain. *Journal of Cerebral Blood Flow & Metabolism*, *41*(5), 986-1000.

Koutsouleris, N., Riecher-Rössler, A., Meisenzahl, E. M., Smieskova, R., Studerus, E., Kambeitz-Ilankovic, L., ... & Borgwardt, S. (2015). Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. *Schizophrenia bulletin*, *41*(2), 471-482.

Koutsouleris, N., Wobrock, T., Guse, B., Langguth, B., Landgrebe, M., Eichhammer, P., ... & Hasan, A. (2018). Predicting response to repetitive transcranial magnetic stimulation in patients with schizophrenia using structural magnetic resonance imaging: a multisite machine learning analysis. *Schizophrenia Bulletin*, *44*(5), 1021-1034.

Kraus, B., Zinbarg, R., Braga, R. M., Nusslock, R., Mittal, V. A., & Gratton, C. (2023). Insights from personalized models of brain and behavior for identifying biomarkers in psychiatry. *Neuroscience & Biobehavioral Reviews*, *152*, 105259.

Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., Morel, A., & Székely, G. (2010). A mean three-dimensional atlas of the human thalamus: generation from multiple histological data. *Neuroimage*, *49*(3), 2053-2062.

Kuhn, B. N., Campus, P., Flagel, S. B., Tomie, A., & Morrow, J. (2018). The neurobiological mechanisms underlying sign-tracking behavior. *Sign-tracking and drug addiction*.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of statistical software*, *28*, 1-26.

LaBerge, D. (1997). Attention, awareness, and the triangular circuit. *Consciousness and cognition*, *6*(2-3), 149-181.

Lakatos, P., O'Connell, M. N., & Barczak, A. (2016). Pondering the pulvinar. *Neuron*, *89*(1), 5-7.

Lake, E. M., Finn, E. S., Noble, S. M., Vanderwal, T., Shen, X., Rosenberg, M. D., ... & Constable, R. T. (2019). The functional brain organization of an individual allows prediction of measures of social abilities transdiagnostically in autism and attention-deficit/hyperactivity disorder. *Biological psychiatry*, *86*(4), 315-326.

Lancaster, T. M., Ihssen, N., Brindley, L. M., Tansey, K. E., Mantripragada, K., O'Donovan, M. C., ... & Linden, D. E. (2016). Associations between polygenic risk for schizophrenia and brain function during probabilistic learning in healthy individuals. *Human Brain Mapping*, *37*(2), 491-500.

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M. Y., ... & Petersen, S. E. (2015). Functional system and areal organization of a highly sampled individual human brain. *Neuron*, *87*(3), 657-670.

Laumann, T. O., Ortega, M., Hoyt, C. R., Seider, N. A., Siegel, J. S., Nguyen, A. L., ... & Dosenbach, N. U. (2021). Brain network reorganisation in an adolescent after bilateral perinatal strokes. *The Lancet Neurology*, *20*(4), 255-256.

Laumann, T. O., Snyder, A. Z., Mitra, A., Gordon, E. M., Gratton, C., Adeyemo, B., ... & Petersen, S. E. (2017). On the stability of BOLD fMRI correlations. *Cerebral cortex*, *27*(10), 4719-4732.

Laumann, T. O., Zorumski, C. F., & Dosenbach, N. U. (2023). Precision neuroimaging for localization-related psychiatry. *JAMA psychiatry*, *80*(8), 763-764.

Lee, J., & Park, S. (2006). The role of stimulus salience in CPT-AX performance of schizophrenia patients. *Schizophrenia Research*, *81*(2-3), 191-197.

Lee, L., Boorman, L., Glendenning, E., Christmas, C., Sharp, P., Redgrave, P., ... & Howarth, C. (2020). Key aspects of neurovascular control mediated by specific populations of inhibitory cortical interneurons. *Cerebral Cortex*, *30*(4), 2452-2464.

Lee, M. H., Hacker, C. D., Snyder, A. Z., Corbetta, M., Zhang, D., Leuthardt, E. C., & Shimony, J. S. (2012). Clustering of resting state networks. *PloS one*, *7*(7), e40370.

Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting-state fMRI: a review of methods and clinical applications. *American Journal of neuroradiology*, *34*(10), 1866-1872.

Lehner, R., Balsters, J. H., Bürgler, A., Hare, T. A., & Wenderoth, N. (2017). Food-predicting stimuli differentially influence eye movements and goal-directed behavior in normal-weight, overweight, and obese individuals. *Frontiers in psychiatry*, *8*, 230.

Li, T., Wang, Q., Zhang, J., Rolls, E. T., Yang, W., Palaniyappan, L., ... & Feng, J. (2017). Brain-wide analysis of functional connectivity in first-episode and chronic stages of schizophrenia. *Schizophrenia bulletin*, *43*(2), 436-448.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703-713.

Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in neurosciences*, *34*(10), 536-547.

Liu, C., Yücel, M., Suo, C., Le Pelley, M. E., Tiego, J., Rotaru, K., ... & Albertella, L. (2021). Reward-related attentional capture moderates the association between fear-driven motives and heavy drinking. *European Addiction Research*, *27*(5), 351-361.

Liu, H., Yuan, K. H., & Li, H. (2023). A systematic framework for defining R-squared measures in mediation analysis. *Psychological Methods*.

Liu, J., Liao, X., Xia, M., & He, Y. (2018). Chronnectome fingerprinting: Identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. *Human brain mapping*, *39*(2), 902-915.

Lloyd, B., & Nieuwenhuis, S. (2024). The effect of reward-induced arousal on the success and precision of episodic memory retrieval. *Scientific Reports*, *14*(1), 2105.

Lo, K. Y. H., Hopman, H. J., Chan, S. C., Chau, W. H. S., Cheng, P. W. C., Cheung, K. Y., ... & Chan, S. M. S. (2022). Concurrent anodal transcranial direct current stimulation (tDCS) with cognitive training to improve cognition in schizophrenia.

Logothetis, N. K. (2002). The neural basis of the blood–oxygen–level–dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1003-1037.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol.*, *66*, 735-769.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *nature*, *412*(6843), 150-157.

Lopez, J., Wolff, M., Lecourtier, L., Cosquer, B., Bontempi, B., Dalrymple-Alford, J., & Cassel, J. C. (2009). The intralaminar thalamic nuclei contribute to remote spatial memory. *Journal of Neuroscience*, *29*(10), 3302-3306.

Lovic, V., Saunders, B. T., Yager, L. M., & Robinson, T. E. (2011). Rats prone to attribute incentive salience to reward cues are also prone to impulsive action. *Behavioural brain research*, *223*(2), 255-261.

Lowe, M. J., Beall, E. B., Sakaie, K. E., Koenig, K. A., Stone, L., Marrie, R. A., & Phillips, M. D. (2008). Resting state sensorimotor functional connectivity in multiple sclerosis inversely correlates with transcallosal motor pathway transverse diffusivity. *Human brain mapping*, *29*(7), 818-827.

Lowe, M. J., Dzemidzic, M., Lurito, J. T., Mathews, V. P., & Phillips, M. D. (2000). Correlations in low-frequency BOLD fluctuations reflect cortico-cortical connections. *Neuroimage*, *12*(5), 582-587.

Lu, W., Dong, K., Cui, D., Jiao, Q., & Qiu, J. (2019). Quality assurance of human functional magnetic resonance imaging: a literature review. *Quantitative imaging in medicine and surgery*, *9*(6), 1147.

Luijten, M., Schellekens, A. F., Kühn, S., Machielse, M. W., & Sescousse, G. (2017). Disruption of reward processing in addiction: an image-based meta-analysis of functional magnetic resonance imaging studies. *JAMA psychiatry*, *74*(4), 387-398.

Lynch, C. J., Elbau, I. G., Ng, T., Ayaz, A., Zhu, S., Wolk, D., ... & Liston, C. (2024). Frontostriatal salience network expansion in individuals in depression. *Nature*, *633*(8030), 624-633.

Lynch, C. J., Elbau, I., & Liston, C. (2021). Improving precision functional mapping routines with multi-echo fMRI. *Current opinion in behavioral sciences*, *40*, 113-119.

Lynch, C. J., Power, J. D., Scult, M. A., Dubin, M., Gunning, F. M., & Liston, C. (2020). Rapid precision functional mapping of individuals using multi-echo fMRI. *Cell reports*, *33*(12).

Ma, I., van Holstein, M., Mies, G. W., Mennes, M., Buitelaar, J., Cools, R., ... & Scheres, A. (2016). Ventral striatal hyperconnectivity during rewarded interference control in adolescents with ADHD. *Cortex*, *82*, 225-236.

Ma, X., Feng, N., Palaniyappan, L., Cao, L., Gu, Z., Kang, J., ... & Luo, Q. (2025). Neuroimaging stratification reveals the striatal vulnerability to stress as a risk for schizophrenia. *Translational Psychiatry*, *15*(1), 18.

Ma, Z., Cao, P., Sun, P., Lu, Z., Li, L., Chen, Y., & Chai, X. (2017). Negative hemodynamic response without neuronal inhibition investigated by combining optical imaging and electrophysiological recording. *Neuroscience letters*, *637*, 161-167.

MacQueen, J. (1967). Multivariate observations. In *Proceedings ofthe 5th Berkeley Symposium on Mathematical Statisticsand Probability* (Vol. 1, pp. 281-297).

Malek-Ahmadi, M., Ginsberg, S. D., Alldred, M. J., Counts, S. E., Ikonomovic, M. D., Abrahamson, E. E., ... & Mufson, E. J. (2024). Application of robust regression in translational neuroscience studies with non-Gaussian outcome data. *Frontiers in Aging Neuroscience*, *15*, 1299451.

Marchand, A., Faugère, A., Coutureau, E., & Wolff, M. (2014). A role for anterior thalamic nuclei in contextual fear memory. *Brain Structure and Function*, *219*(5), 1575-1586.

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., ... & Dosenbach, N. U. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654-660.

Margulies, D. S., Kelly, A. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *Neuroimage*, *37*(2), 579-588.

Marini, F., Scott, J., Aron, A. R., & Ester, E. F. (2017). Task-irrelevant distractors in the delay period interfere selectively with visual short-term memory for spatial locations. *Attention, Perception, & Psychophysics*, *79*(5), 1384-1392.

Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biological psychiatry: cognitive neuroscience and neuroimaging*, *1*(5), 433-447.

Mascarell Maričić, L., Walter, H., Rosenthal, A., Ripke, S., Quinlan, E. B., Banaschewski, T., ... & Heinz, A. (2020). The IMAGEN study: a decade of imaging genetics in adolescents. *Molecular psychiatry*, *25*(11), 2648-2671.

Mason, A., Farrell, S., Howard-Jones, P., & Ludwig, C. J. (2017). The role of reward and reward uncertainty in episodic memory. *Journal of memory and language*, *96*, 62-77.

Matsuo, K., Nicoletti, M., Nemoto, K., Hatch, J. P., Peluso, M. A., Nery, F. G., & Soares, J. C. (2009). A voxel-based morphometry study of frontal gray matter correlates of impulsivity. *Human brain mapping*, *30*(4), 1188-1195.

Mattoni, M., Fisher, A. J., Gates, K. M., Chein, J., & Olino, T. M. (2025). Group-to-individual generalizability and individual-level inferences in cognitive neuroscience. *Neuroscience & Biobehavioral Reviews*, 106024.

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in cognitive sciences*, *11*(3), 126-135.

Mayhew, S. D., Mullinger, K. J., Bagshaw, A. P., Bowtell, R., & Francis, S. T. (2014). Investigating intrinsic connectivity networks using simultaneous BOLD and CBF measurements. *Neuroimage*, *99*, 111-121.

McCutcheon, R. A., Harrison, P. J., Howes, O. D., McGuire, P. K., Taylor, D. M., & Pillinger, T. (2023). Data-driven taxonomy for antipsychotic medication: a new classification system. *Biological psychiatry*, *94*(7), 561-568.

McCutcheon, R. A., Keefe, R. S., & McGuire, P. K. (2023). Cognitive impairment in schizophrenia: aetiology, pathophysiology, and treatment. *Molecular psychiatry*, *28*(5), 1902-1918.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature neuroscience*, *11*(1), 103-107.

Mendelson, A. F., Zuluaga, M. A., Lorenzi, M., Hutton, B. F., Ourselin, S., & Alzheimer's Disease Neuroimaging Initiative. (2017). Selection bias in the reported performances of AD classification pipelines. *NeuroImage: Clinical*, *14*, 400-416.

Meng, X., Jiang, R., Lin, D., Bustillo, J., Jones, T., Chen, J., ... & Calhoun, V. D. (2017). Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage*, *145*, 218-229.

Menon, M., Jensen, J., Vitcu, I., Graff-Guerrero, A., Crawley, A., Smith, M. A., & Kapur, S. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biological psychiatry*, *62*(7), 765-772.

Meyer, M., Di Scala, G., Edde, M., Dilharreguy, B., Radat, F., Allard, M., & Chanraud, S. (2017). Brain structural investigation and hippocampal tractography in medication overuse headache: a native space analysis. *Behavioral and Brain Functions*, *13*(1), 6.

Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E. (2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PloS one*, *7*(6), e38987.

Miller, M. B., Donovan, C. L., Bennett, C. M., Aminoff, E. M., & Mayer, R. E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage*, *59*(1), 83-93.

Miller, M. B., Van Horn, J. D., Wolford, G. L., Handy, T. C., Valsangkar-Smyth, M., Inati, S., ... & Gazzaniga, M. S. (2002). Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *Journal of Cognitive Neuroscience*, *14*(8), 1200-1214.

Mishra, A. M., Ellens, D. J., Schridde, U., Motelow, J. E., Purcaro, M. J., DeSalvo, M. N., ... & Blumenfeld, H. (2011). Where fMRI and electrophysiology agree to disagree: corticothalamic and striatal activity patterns in the WAG/Rij rat. *Journal of Neuroscience*, *31*(42), 15053-15064.

Mitchell, A. S. (2015). The mediodorsal thalamus as a higher order thalamic relay nucleus important for learning and decision-making. *Neuroscience & Biobehavioral Reviews*, *54*, 76-88.

Mitchell, A. S., & Chakraborty, S. (2013). What does the mediodorsal thalamus do?. *Frontiers in systems neuroscience*, *7*, 37.

Mitchell, A. S., & Dalrymple-Alford, J. C. (2005). Dissociable memory effects after medial thalamus lesions in the rat. *European Journal of Neuroscience*, *22*(4), 973-985.

Mizrahi, R., Addington, J., Rusjan, P. M., Suridjan, I., Ng, A., Boileau, I., ... & Wilson, A. A. (2012). Increased stress-induced dopamine release in psychosis. *Biological psychiatry*, *71*(6), 561-567.

Moore, N. A., Leander, J. D., Benvenga, M. J., Gleason, S. D., & Shannon, H. (1997). Behavioral pharmacology of olanzapine: a novel antipsychotic drug. *Journal of Clinical Psychiatry*, *58*(10), 37-44.

Moradi, E., Khundrakpam, B., Lewis, J. D., Evans, A. C., & Tohka, J. (2017). Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data. *Neuroimage*, *144*, 128-141.

Morel, A. (2007). *Stereotactic atlas of the human thalamus and basal ganglia*. CRC Press.

Morrison, S. E., Bamkole, M. A., & Nicola, S. M. (2015). Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Frontiers in neuroscience*, *9*, 468.

Morrow, J. D., Maren, S., & Robinson, T. E. (2011). Individual variation in the propensity to attribute incentive salience to an appetitive cue predicts the propensity to attribute motivational salience to an aversive cue. *Behavioural brain research*, *220*(1), 238-243.

Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., ... & Liu, H. (2013). Individual variability in functional connectivity architecture of the human brain. *Neuron*, *77*(3), 586-595.

Mukai, I., Kim, D., Fukunaga, M., Japee, S., Marrett, S., & Ungerleider, L. G. (2007). Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *Journal of Neuroscience*, *27*(42), 11401-11411.

Mullinger, K. J., Mayhew, S. D., Bagshaw, A. P., Bowtell, R., & Francis, S. T. (2014). Evidence that the negative BOLD response is neuronal in origin: a simultaneous EEG–BOLD–CBF study in humans. *Neuroimage*, *94*, 263-274.

Murray, R. M., O'Callaghan, E., Castle, D. J., & Lewis, S. W. (1992). A neurodevelopmental approach to the classification of schizophrenia. *Schizophrenia Bulletin*, *18*(2), 319-332.

Murty, V. P., LaBar, K. S., Hamilton, D. A., & Adcock, R. A. (2011). Is all motivation good for learning? Dissociable influences of approach and avoidance motivation in declarative memory. *Learning & Memory*, *18*(11), 712-717.

Muzzio, I. A., Kentros, C., & Kandel, E. (2009). What is remembered? Role of attention on the encoding and retrieval of hippocampal representations. *The Journal of physiology*, *587*(12), 2837-2854.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current opinion in neurobiology*, *7*(2), 217-227.

Nagaoka, T., Zhao, F., Wang, P., Harel, N., Kennan, R. P., Ogawa, S., & Kim, S. G. (2006). Increases in oxygen consumption without cerebral blood volume change during visual stimulation under hypotension condition. *Journal of Cerebral Blood Flow & Metabolism*, *26*(8), 1043-1051.

Nakajima, M., & Halassa, M. M. (2017). Thalamic control of functional cortical connectivity. *Current opinion in neurobiology*, *44*, 127-131.

Nakajima, M., Schmitt, L. I., & Halassa, M. M. (2019). Prefrontal cortex regulates sensory filtering through a basal ganglia-to-thalamus pathway. *Neuron*, *103*(3), 445-458.

Nakata, H., Domoto, R., Mizuguchi, N., Sakamoto, K., & Kanosue, K. (2019). Negative BOLD responses during hand and foot movements: An fMRI study. *PLoS One*, *14*(4), e0215736.

Nakuci, J., & Rahnev, D. (2025). A Practical Guide to Identifying Robust Clusters in Neuroimaging Data. Human Brain Mapping, 46(13), e70330.

Nakuci, J., Yeon, J., Haddara, N., Kim, J. H., Kim, S. P., & Rahnev, D. (2025). Multiple brain activation patterns for the same perceptual decision-making task. Nature Communications, 16(1), 1785.

Nelson, J. B., Navarro, A., & del Carmen Sanjuan, M. (2019). Conditioned anticipatory outcome searching in humans. *Behavioural processes*, *164*, 237-251.

Ngo, C. T., Horner, A. J., Newcombe, N. S., & Olson, I. R. (2019). Development of holistic episodic recollection. *Psychological science*, *30*(12), 1696-1706.

Nielsen, M. O., Rostrup, E., Wulff, S., Bak, N., Broberg, B. V., Lublin, H., ... & Glenthoj, B. (2012). Improvement of brain reward abnormalities by antipsychotic monotherapy in schizophrenia. *Archives of general psychiatry*, *69*(12), 1195-1204.

Nielsen, M. Ø., Rostrup, E., Wulff, S., Bak, N., Lublin, H., Kapur, S., & Glenthøj, B. (2012). Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biological psychiatry*, *71*(10), 898-905.

Nucifora Jr, F. C., Woznica, E., Lee, B. J., Cascella, N., & Sawa, A. (2019). Treatment resistant schizophrenia: Clinical, biological, and therapeutic perspectives. *Neurobiology of disease*, *131*, 104257.

Nymberg, C., Jia, T., Lubbe, S., Ruggeri, B., Desrivieres, S., Barker, G., ... & IMAGEN Consortium. (2013). Neural mechanisms of attention-deficit/hyperactivity disorder symptoms are stratified by MAOA genotype. Biological psychiatry, 74(8), 607-614.

O'Daly, O. G., Joyce, D., Tracy, D. K., Azim, A., Stephan, K. E., Murray, R. M., & Shergill, S. S. (2019). Amphetamine Sensitization Alters Reward Processing in the Human Striatum and Amygdala (vol 9, e93955, 2014). *PLOS ONE*, *14*(6).

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329-337.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *science*, *304*(5669), 452-454.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *proceedings of the National Academy of Sciences*, *87*(24), 9868-9872.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97-113.

Oldham, S., Murawski, C., Fornito, A., Youssef, G., Yücel, M., & Lorenzetti, V. (2018). The anticipation and outcome phases of reward and loss processing: A neuroimaging meta-analysis of the monetary incentive delay task. *Human brain mapping*, *39*(8), 3398-3418.

O'muircheartaigh, J., Vollmar, C., Barker, G. J., Kumari, V., Symms, M. R., Thompson, P., ... & Richardson, M. P. (2011). Focal structural changes and cognitive dysfunction in juvenile myoclonic epilepsy. *Neurology*, *76*(1), 34-40.

Opitz, B. (2014). Memory function and the hippocampus. *Front Neurol Neurosci*, *34*, 51-59.

Orru, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1140-1152.

Osuch, E., Gao, S., Wammes, M., Théberge, J., Williamson, P., Neufeld, R. J., ... & Calhoun, V. (2018). Complexity in mood disorder diagnosis: fMRI connectivity networks predicted medication-class of response in complex patients. *Acta Psychiatrica Scandinavica*, *138*(5), 472-482.

Osugo, M., Wall, M. B., Selvaggi, P., Zahid, U., Finelli, V., Chapman, G. E., ... & Howes, O. D. (2025). Striatal dopamine D2/D3 receptor regulation of human reward processing and behaviour. *Nature Communications*, *16*(1), 1852.

Owen, M. J. (2014). New approaches to psychiatric diagnostic classification. *Neuron*, *84*(3), 564-571.

Paolone, G., Angelakos, C. C., Meyer, P. J., Robinson, T. E., & Sarter, M. (2013). Cholinergic control over attention in rats prone to attribute incentive salience to reward cues. *Journal of Neuroscience*, *33*(19), 8321-8335.

Papadimitriou, C. H., & Steiglitz, K. (1998). *Combinatorial optimization: algorithms and complexity*. Courier Corporation.

Parnaudeau, S., Bolkan, S. S., & Kellendonk, C. (2018). The mediodorsal thalamus: an essential partner of the prefrontal cortex for cognition. *Biological psychiatry*, *83*(8), 648-656.

Pasley, B. N., Inglis, B. A., & Freeman, R. D. (2007). Analysis of oxygen metabolism implies a neural origin for the negative BOLD response in human visual cortex. *Neuroimage*, *36*(2), 269-276.

Passiatore, R., Antonucci, L. A., Bierstedt, S., Saranathan, M., Bertolino, A., Suchan, B., & Pergola, G. (2021). How recent learning shapes the brain: Memory-dependent functional reconfiguration of brain circuits. *NeuroImage*, *245*, 118636.

Passiatore, R., Antonucci, L. A., DeRamus, T. P., Fazio, L., Stolfa, G., Sportelli, L., ... & Pergola, G. (2023). Changes in patterns of age-related network connectivity are associated with risk for schizophrenia. *Proceedings of the National Academy of Sciences*, *120*(32), e2221533120.

Pattyn, T., Van Den Eede, F., Lamers, F., Veltman, D., Sabbe, B. G., & Penninx, B. W. (2015). Identifying panic disorder subtypes using factor mixture modeling. *Depression and anxiety*, *32*(7), 509-517.

Pavlov, I. P. (1927). *Les réflexes conditionnels: étude objective de l'activité nerveuse supérieure des animaux*. F. Alcan.

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of neurosciences*, *17*(3), 136.

Perakyla J, Sun L, Lehtimaki K, Peltola J, Ohman J, Mottonen T, Ogawa KH, Hartikainen KM (2017) Causal Evidence from Humans for the Role of Mediodorsal Nucleus of the Thalamus in Working Memory. J Cogn Neurosci 29:2090-2102.

Peräkylä, J., Sun, L., Lehtimäki, K., Peltola, J., Öhman, J., Möttönen, T., ... & Hartikainen, K. M. (2017). Causal evidence from humans for the role of mediodorsal nucleus of the thalamus in working memory. *Journal of cognitive neuroscience*, *29*(12), 2090-2102.

Pergola, G., & Suchan, B. (2013). Associative learning beyond the medial temporal lobe: many actors on the memory stage. *Frontiers in behavioral neuroscience*, *7*, 162.

Pergola, G., Bellebaum, C., Gehlhaar, B., Koch, B., Schwarz, M., Daum, I., & Suchan, B. (2013). The involvement of the thalamus in semantic retrieval: a clinical group study. *Journal of Cognitive Neuroscience*, *25*(6), 872-886.

Pergola, G., Danet, L., Pitel, A. L., Carlesimo, G. A., Segobin, S., Pariente, J., ... & Barbeau, E. J. (2018). The regulatory role of the human mediodorsal thalamus. *Trends in cognitive sciences*, *22*(11), 1011-1025.

Pergola, G., Güntürkün, O., Koch, B., Schwarz, M., Daum, I., & Suchan, B. (2012). Recall deficits in stroke patients with thalamic lesions covary with damage to the parvocellular mediodorsal nucleus of the thalamus. *Neuropsychologia*, *50*(10), 2477-2491.

Pergola, G., Ranft, A., Mathias, K., & Suchan, B. (2013). The role of the thalamic nuclei in recognition memory accompanied by recall during encoding and retrieval: an fMRI study. *Neuroimage*, *74*, 195-208.

Pergola, G., Selvaggi, P., Trizio, S., Bertolino, A., & Blasi, G. (2015). The role of the thalamus in schizophrenia from a neuroimaging perspective. *Neuroscience & Biobehavioral Reviews*, *54*, 57-75.

Pergola, G., Suchan, B., Koch, B., Schwarz, M., Daum, I., & Güntürkün, O. (2013). Quantitative assessment of chronic thalamic stroke. *American Journal of Neuroradiology*, *34*(5), E51-E55.

Pergola, G., Trizio, S., Di Carlo, P., Taurisano, P., Mancini, M., Amoroso, N., ... & Blasi, G. (2017). Grey matter volume patterns in thalamic nuclei are associated with familial risk for schizophrenia. *Schizophrenia research*, *180*, 13-20.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042-1045.

Pessoa, L., Gutierrez, E., Bandettini, P. A., & Ungerleider, L. G. (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron*, *35*(5), 975-987.

Petersen, D., Raudales, R., Silva, A. K., Kellendonk, C., & Canetta, S. (2024). Adolescent Thalamoprefrontal inhibition leads to changes in intrinsic prefrontal network connectivity. *eneuro*, *11*(8).

Peterson, A. D., Ghosh, A. P., & Maitra, R. (2018). Merging K-means with hierarchical clustering for identifying general-shaped groups. *Stat*, *7*(1), e172.

Phillips, J. M., Kambi, N. A., Redinbaugh, M. J., Mohanta, S., & Saalmann, Y. B. (2021). Disentangling the influences of multiple thalamic nuclei on prefrontal cortex and cognitive control. *Neuroscience & Biobehavioral Reviews*, *128*, 487-510.

Pitchers, K. K., Sarter, M., & Robinson, T. E. (2018). The hot 'n'cold of cue-induced drug relapse. *Learning & Memory*, *25*(9), 474-480.

Plichta, M. M., Schwarz, A. J., Grimm, O., Morgen, K., Mier, D., Haddad, L., ... & Meyer-Lindenberg, A. (2012). Test–retest reliability of evoked BOLD signals from a cognitive–emotive fMRI test battery. Neuroimage, 60(3), 1746-1758.

Poil, S. S., Hardstone, R., Mansvelder, H. D., & Linkenkaer-Hansen, K. (2012). Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks. *Journal of Neuroscience*, *32*(29), 9817-9823.

Poldrack, R. A., & Farah, M. J. (2015). Progress and challenges in probing the human brain. *Nature*, *526*(7573), 371-379.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., ... & Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature reviews neuroscience*, *18*(2), 115-126.

Polosecki, P., Castro, E., Rish, I., Pustina, D., Warner, J. H., Wood, A., ... & Cecchi, G. A. (2020). Resting-state connectivity stratifies premanifest Huntington's disease by longitudinal cognitive decline rate. *Scientific reports*, *10*(1), 1252.

Polosecki, P., Castro, E., Wood, A., Warner, J. H., Rish, I., & Cecchi, G. A. (2017). Computational psychiatry: Advancing predictive modeling of neurodegeneration with neuroimaging of Huntington's disease. *IBM Journal of Research and Development*, *61*(2/3), 4-1.

Pool, E. R., Tord, D. M., Delplanque, S., Stussi, Y., Cereghetti, D., Vuilleumier, P., & Sander, D. (2022). Differential contributions of ventral striatum subregions to the motivational and hedonic components of the affective processing of reward. *Journal of Neuroscience*, *42*(13), 2716-2728.

Popken, G. J., Bunney Jr, W. E., Potkin, S. G., & Jones, E. G. (2000). Subnucleus-specific loss of neurons in medial thalamus of schizophrenics. *Proceedings of the National Academy of Sciences*, *97*(16), 9276-9280.

Powell, E. W. (1973). Limbic projections to the thalamus. *Experimental Brain Research*, *17*(4), 394-401.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage, 59(3), 2142-2154.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, *72*(4), 665-678.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *neuroimage*, *84*, 320-341.

Preuss, T. M., & Goldman-Rakic, P. S. (1987). Crossed corticothalamic and thalamocortical connections of macaque prefrontal cortex. *Journal of Comparative Neurology*, *257*(2), 269-281.

Puckett, A. M., Mathis, J. R., & DeYoe, E. A. (2014). An investigation of positive and inverted hemodynamic response functions across multiple visual areas. *Human brain mapping*, *35*(11), 5550-5564.

Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus*, *17*(3), 192-200.

Rabini, G., Ubaldi, S., & Fairhall, S. L. (2023). Task-based activation and resting-state connectivity predict individual differences in semantic capacity for complex semantic knowledge. *Communications Biology*, *6*(1), 1020.

Radua, J., Schmidt, A., Borgwardt, S., Heinz, A., Schlagenhauf, F., McGuire, P., & Fusar-Poli, P. (2015). Ventral striatal activation during reward processing in psychosis: a neurofunctional meta-analysis. *JAMA psychiatry*, *72*(12), 1243-1251.

Ragland, J. D., Ranganath, C., Barch, D. M., Gold, J. M., Haley, B., MacDonald III, A. W., ... & Carter, C. S. (2012). Relational and Item-Specific Encoding (RISE): task development and psychometric characteristics. *Schizophrenia bulletin*, *38*(1), 114-124.

Ragland, J. D., Ranganath, C., Harms, M. P., Barch, D. M., Gold, J. M., Layher, E., ... & Carter, C. S. (2015). Functional and neuroanatomic specificity of episodic memory dysfunction in schizophrenia: a functional magnetic resonance imaging study of the relational and item-specific encoding task. *JAMA psychiatry*, *72*(9), 909-916.

Raichle, M. E. (2015). The brain's default mode network. *Annual review of neuroscience*, *38*(1), 433-447.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the national academy of sciences*, *98*(2), 676-682.

Raimondo, L., Heij, J., Priovoulos, N., Kundu, P., Leoni, R. F., & van der Zwaag, W. (2021). Advances in resting state fMRI acquisitions for functional connectomics. Neuroimage, 243, 118503.

Raio, A., Pergola, G., Rampino, A., Russo, M., D'Ambrosio, E., Selvaggi, P., ... & Apulian Network on Risk for Psychosis Manzari Anna 6 Carofiglio Angela 7 Barrasso Giuseppe 5 Bellomo Antonello 4 Leccisotti Ivana 4 Di Fino Melania 4 Andriola Ileana 1 2 Pennacchio Teresa Claudia 1. (2023). Similarities and differences between multivariate patterns of cognitive and socio-cognitive deficits in schizophrenia, bipolar disorder and related risk. *Schizophrenia*, *9*(1), 11.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature reviews neuroscience*, *13*(10), 713-726.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia*, *41*(3), 378-389.

Rasero, J., Sentis, A. I., Yeh, F. C., & Verstynen, T. (2021). Integrating across neuroimaging modalities boosts prediction accuracy of cognitive ability. *PLoS computational biology*, *17*(3), e1008347.

Rashid, B., Arbabshirani, M. R., Damaraju, E., Cetin, M. S., Miller, R., Pearlson, G. D., & Calhoun, V. D. (2016). Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity. *Neuroimage*, *134*, 645-657.

Ray, J. P., & Price, J. L. (1993). The organization of projections from the mediodorsal nucleus of the thalamus to orbital and medial prefrontal cortex in macaque monkeys. *Journal of Comparative Neurology*, *337*(1), 1-31.

Reineberg, A. E., Andrews-Hanna, J. R., Depue, B. E., Friedman, N. P., & Banich, M. T. (2015). Resting-state networks predict individual differences in common and specific aspects of executive function. *Neuroimage*, *104*, 69-78.

Rikhye, R. V., Gilra, A., & Halassa, M. M. (2018). Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nature neuroscience*, *21*(12), 1753-1763.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. Neuroimage, 23(2), 752-763.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain research reviews*, *18*(3), 247-291.

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biological psychiatry*, *65*(10), 869-873.

Robison, A. J., Thakkar, K. N., & Diwadkar, V. A. (2020). Cognition and reward circuits in schizophrenia: synergistic, not separate. *Biological psychiatry*, *87*(3), 204-214.

Romaniuk, L., Honey, G. D., King, J. R., Whalley, H. C., McIntosh, A. M., Levita, L., ... & Hall, J. (2010). Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Archives of general psychiatry*, *67*(12), 1246-1254.

Rosazza, C., & Minati, L. (2011). Resting-state brain networks: literature review and clinical applications. *Neurological sciences*, *32*(5), 773-785.

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature neuroscience*, *19*(1), 165-171.

Rosenthal, R., & Rubin, D. B. (2003). requivalent: A simple effect size indicator. Psychological methods, 8(4), 492.

Rosenthal, R., Cooper, H., & Hedges, L. (1994). Parametric measures of effect size. *The handbook of research synthesis*, *621*(2), 231-244.

Rosner, B. (1975). On the detection of many outliers. *Technometrics*, *17*(2), 221-227.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of statistical software, 48(1), 1-36.

Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1430.

Roy, D. S., Zhang, Y., Halassa, M. M., & Feng, G. (2022). Thalamic subnetworks as units of function. *Nature neuroscience*, *25*(2), 140-153.

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current opinion in neurobiology*, *23*(2), 255-260.

Russchen, F. T., Amaral, D. G., & Price, J. L. (1987). The afferent input to the magnocellular division of the mediodorsal thalamic nucleus in the monkey, Macaca fascicularis. *Journal of Comparative Neurology*, *256*(2), 175-210.

Saalmann, Y. B. (2014). Intralaminar and medial thalamic influence on cortical synchrony, information transmission and cognition. *Frontiers in systems neuroscience*, *8*, 83.

Saalmann, Y. B., & Kastner, S. (2011). Cognitive and perceptual functions of the visual thalamus. *Neuron*, *71*(2), 209-223.

Saalmann, Y. B., Pinsk, M. A., Wang, L., Li, X., & Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. *science*, *337*(6095), 753-756.

Sacco, R., Lenti, C., Saccani, M., Curatolo, P., Manzi, B., Bravaccio, C., & Persico, A. M. (2012). Cluster analysis of autistic patients based on principal pathogenetic components. *Autism Research*, *5*(2), 137-147.

Sadikot, A. F., Parent, A., & Francois, C. (1992). Efferent connections of the centromedian and parafascicular thalamic nuclei in the squirrel monkey: a PHA-L study of subcortical projections. *Journal of Comparative Neurology*, *315*(2), 137-159.

Sambuco, N. (2022). fMRI replicability during emotional scene viewing: Functional regions and sample size. *Psychophysiology*, *59*(6), e14000.

Sambuco, N., Bradley, M. M., Herring, D. R., & Lang, P. J. (2020). Common circuit or paradigm shift? The functional brain in emotional scene perception and emotional imagery. *Psychophysiology*, *57*(4), e13522.

Sanfratello, L., Caprihan, A., Stephen, J. M., Knoefel, J. E., Adair, J. C., Qualls, C., ... & Aine, C. J. (2014). Same task, different strategies: How brain networks can be influenced by memory strategy. *Human brain mapping*, *35*(10), 5127-5140.

Sarakam, B. C. (2025). Reinforcement Learning in Neuroimaging. *International Journal of Innovative Science and Research Technology*, *10*(6), 123-132.

Sarpal, D. K., Robinson, D. G., Lencz, T., Argyelan, M., Ikuta, T., Karlsgodt, K., ... & Malhotra, A. K. (2015). Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. *JAMA psychiatry*, *72*(1), 5-13.

Sarter, M., & Phillips, K. B. (2018). The neuroscience of cognitive-motivational styles: Sign-and goal-trackers as animal models. *Behavioral neuroscience*, *132*(1), 1.

Saunders, B. T., & Robinson, T. E. (2012). The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *European Journal of Neuroscience*, *36*(4), 2521-2532.

Saunders, B. T., & Robinson, T. E. (2013). Individual variation in resisting temptation: implications for addiction. *Neuroscience & Biobehavioral Reviews*, *37*(9), 1955-1975.

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, *35*(2), 319-334.

Schad, D. J., Rapp, M. A., Garbusow, M., Nebe, S., Sebold, M., Obst, E., ... & Huys, Q. J. (2020). Dissociating neural learning signals in human sign-and goal-trackers. *Nature human behaviour*, *4*(2), 201-214.

Schäfer, K., Blankenburg, F., Kupers, R., Grüner, J. M., Law, I., Lauritzen, M., & Larsson, H. B. (2012). Negative BOLD signal changes in ipsilateral primary somatosensory cortex are associated with perfusion decreases and behavioral evidence for functional inhibition. *Neuroimage*, *59*(4), 3119-3127.

Scheinost, D., Noble, S., Horien, C., Greene, A. S., Lake, E. M., Salehi, M., ... & Constable, R. T. (2019). Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage*, *193*, 35-45.

Schembre, S., Hoffman, K., Versace, F., Petrosini, J., & Daniel-MacDougall, C. (2016). Gut microbiome profiles vary by the propensity to attribute incentive salience to palatable food cues in humans. *Psychoneuroendocrinology*, (71), 69.

Scheres, A., Milham, M. P., Knutson, B., & Castellanos, F. X. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivity disorder. *Biological psychiatry*, *61*(5), 720-724.

Schettino, M., Mauti, M., Parrillo, C., Ceccarelli, I., Giove, F., Napolitano, A., ... & Orsini, C. (2024). Resting-state brain activation patterns and network topology distinguish human sign and goal trackers. *Translational Psychiatry*, *14*(1), 446.

Schlagenhauf, F., Sterzer, P., Schmack, K., Ballmaier, M., Rapp, M., Wrase, J., ... & Heinz, A. (2009). Reward feedback alterations in unmedicated schizophrenia patients: relevance for delusions. *Biological psychiatry*, *65*(12), 1032-1039.

Schmahmann, J. D. (2003). Vascular syndromes of the thalamus. *Stroke*, *34*(9), 2264-2278.

Schmidt, K., Nolte-Zenker, B., Patzer, J., Bauer, M., Schmidt, L. G., & Heinz, A. (2001). Psychopathological correlates of reduced dopamine receptor sensitivity in depression, schizophrenia, and opiate and alcohol dependence. *Pharmacopsychiatry*, *34*(02), 66-72.

Schmitt, L. I., Wimmer, R. D., Nakajima, M., Happ, M., Mofakham, S., & Halassa, M. M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature*, *545*(7653), 219-223.

Schnack, H. G. (2019). Improving individual predictions: machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophrenia research*, *214*, 34-42.

Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., ... & Bauer, A. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *Journal of Neuroscience*, *28*(52), 14311-14319.

Schridde, U., Khubchandani, M., Motelow, J. E., Sanganahalli, B. G., Hyder, F., & Blumenfeld, H. (2008). Negative BOLD with large increases in neuronal activity. *Cerebral cortex*, *18*(8), 1814-1827.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*.

Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. *Physiological reviews*, *95*(3), 853-951.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual review of neuroscience*, *23*(1), 473-500.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. Science, 275(5306), 1593-1599.

Schumann, G., Binder, E. B., Holte, A., de Kloet, E. R., Oedegaard, K. J., Robbins, T. W., ... & Wittchen, H. U. (2014). Stratified medicine for mental disorders. *European Neuropsychopharmacology*, *24*(1), 5-50.

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., ... & Struve, M. (2010). The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Molecular psychiatry*, *15*(12), 1128-1139.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry*, *20*(1), 11.

Sebold, M., Garbusow, M., Cerci, D., Chen, K., Sommer, C., Huys, Q. J., ... & Friedel, E. (2021). Association of the OPRM1 A118G polymorphism and Pavlovian-to-instrumental transfer: Clinical relevance for alcohol dependence. *Journal of Psychopharmacology*, *35*(5), 566-578.

Seghier, M. L., & Price, C. J. (2009). Dissociating functional brain networks by decoding the between-subject variability. *Neuroimage*, *45*(2), 349-359.

Seghier, M. L., & Price, C. J. (2018). Interpreting and utilising intersubject variability in brain function. *Trends in cognitive sciences*, *22*(6), 517-530.

Segobin, S., & Pitel, A. L. (2021). The specificity of thalamic alterations in Korsakoff's syndrome: Implications for the study of amnesia. *Neuroscience & Biobehavioral Reviews*, *130*, 292-300.

Seibert, T. M., Majid, D. A., Aron, A. R., Corey-Bloom, J., & Brewer, J. B. (2012). Stability of resting fMRI interregional correlations analyzed in subject-native space: a one-year longitudinal study in healthy adults and premanifest Huntington's disease. *Neuroimage*, *59*(3), 2452-2463.

Seitzman, B. A., Gratton, C., Laumann, T. O., Gordon, E. M., Adeyemo, B., Dworetsky, A., ... & Petersen, S. E. (2019). Trait-like variants in human functional brain networks. *Proceedings of the National Academy of Sciences*, *116*(45), 22851-22861.

Selemon, L. D., & Goldman-Rakic, P. S. (1988). Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. *Journal of Neuroscience*, *8*(11), 4049-4068.

Serrano-Barroso, A., Vargas, J. P., Diaz, E., Gómez-González, I. M., Ruiz, G., & López, J. C. (2022). A videogame as a tool for clinical screening of possible vulnerability to impulsivity and attention disturbances in children. *Children*, *9*(11), 1652.

Sescousse, G., Caldú, X., Segura, B., & Dreher, J. C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(4), 681-696.

Shackman, A. J., Smith, J. F., Orth, R. D., Savage, C. L., Didier, P. R., McCarthy, J. M., ... & Blanchard, J. J. (2025). Blunted Ventral Striatal Reactivity to Social Reward Is Associated with More Severe Motivation and Pleasure Deficits in Psychosis. *Schizophrenia Bulletin*, sbae221.

Shao, T., Wang, W., Hei, G., Yang, Y., Long, Y., Wang, X., ... & Wu, R. (2022). Identifying and revealing different brain neural activities of cognitive subtypes in early course schizophrenia. *Frontiers in Molecular Neuroscience*, *15*, 983995.

Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. nature protocols, 12(3), 506-518.

Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature neuroscience*, *19*(4), 533-541.

Sherman, S. M., & Guillery, R. W. (2011). Distinct functions for direct and transthalamic corticocortical connections. *Journal of neurophysiology*, *106*(3), 1068-1077.

Shettleworth, S. J. (2009). *Cognition, evolution, and behavior*. Oxford university press.

Shibata, H., & Yukie, M. (2003). Differential thalamic connections of the posteroventral and dorsal posterior cingulate gyrus in the monkey. *European Journal of Neuroscience*, *18*(6), 1615-1626.

Shigemune, Y., Tsukiura, T., Kambara, T., & Kawashima, R. (2014). Remembering with gains and losses: effects of monetary reward and punishment on successful encoding activation of source memories. *Cerebral cortex*, *24*(5), 1319-1331.

Shih, Y. Y. I., Chen, C. C. V., Shyu, B. C., Lin, Z. J., Chiang, Y. C., Jaw, F. S., ... & Chang, C. (2009). A new scenario for negative functional magnetic resonance imaging signals: endogenous neurotransmission. *Journal of Neuroscience*, *29*(10), 3036-3044.

Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. K. (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature neuroscience*, *9*(4), 569-577.

Shmuel, A., Yacoub, E., Pfeuffer, J., Van de Moortele, P. F., Adriany, G., Hu, X., & Ugurbil, K. (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron*, *36*(6), 1195-1210.

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in cognitive sciences*, *14*(10), 464-472.

Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human brain mapping*, *35*(5), 1981-1996.

Simon, J. J., Biller, A., Walther, S., Roesch-Ely, D., Stippich, C., Weisbrod, M., & Kaiser, S. (2010). Neural correlates of reward processing in schizophrenia—relationship to apathy and depression. *Schizophrenia research*, *118*(1-3), 154-161.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature reviews neuroscience*, *4*(8), 637-648.

Singer, B. F., Guptaroy, B., Austin, C. J., Wohl, I., Lovic, V., Seiler, J. L., ... & Aragona, B. J. (2016). Individual variation in incentive salience attribution and accumbens dopamine transporter expression and function. *European Journal of Neuroscience*, *43*(5), 662-670.

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., ... & Sulzer, J. (2017). Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, *18*(2), 86-100.

Smedley, E. B., & Smith, K. S. (2018). Evidence of structure and persistence in motivational attraction to serial Pavlovian cues. *Learning & Memory*, *25*(2), 78-89.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage, 44(1), 83-98.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the national academy of sciences*, *106*(31), 13040-13045.

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., ... & Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature neuroscience*, *18*(11), 1565-1567.

Song, Z., Tustison, N., Avants, B., & Gee, J. C. (2006, October). Integrated graph cuts for brain MRI segmentation. In International conference on medical image computing and computer-assisted intervention (pp. 831-838). Berlin, Heidelberg: Springer Berlin Heidelberg.

Speer, N. K., Jacoby, L. L., & Braver, T. S. (2003). Strategy-dependent changes in memory: Effects on behavior and brain activity. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(3), 155-167.

Spitzer, M. (1997). A cognitive neuroscience view of schizophrenic thought disorder. *Schizophrenia Bulletin*, *23*(1), 29-50.

Sportelli, L., Eisenberg, D. P., Passiatore, R., D'Ambrosio, E., Antonucci, L. A., Bettina, J. S., ... & Pergola, G. (2024). Dopamine signaling enriched striatal gene set predicts striatal dopamine synthesis and physiological activity in vivo. *Nature communications*, *15*(1), 3342.

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of cognitive neuroscience*, *4*(3), 232-243.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory*, *82*(3), 171-177.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Current opinion in neurobiology*, *5*(2), 169-177.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, *93*(24), 13515-13522.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor perspectives in biology*, *7*(8), a021766.

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci.*, *27*(1), 279-306.

Stark, C. E., & Squire, L. R. (2001). When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the national Academy of Sciences*, *98*(22), 12760-12766.

Stefanovic, B., Warnking, J. M., & Pike, G. B. (2004). Hemodynamic and metabolic responses to neuronal inhibition. *Neuroimage*, *22*(2), 771-778.

Sten, S., Lundengård, K., Witt, S. T., Cedersund, G., Elinder, F., & Engström, M. (2017). Neural inhibition can explain negative BOLD responses: A mechanistic modelling and fMRI study. *Neuroimage*, *158*, 219-231.

Stepien, M., Manoliu, A., Kubli, R., Schneider, K., Tobler, P. N., Seifritz, E., ... & Kirschner, M. (2018). Investigating the association of ventral and dorsal striatal dysfunction during reward anticipation with negative symptoms in patients with schizophrenia and healthy individuals. *PLoS One*, *13*(6), e0198215.

Štillová, K., Jurák, P., Chládek, J., Chrastina, J., Halámek, J., Bočková, M., ... & Rektor, I. (2015). The role of anterior nuclei of the thalamus: a subcortical gate in memory processing: an intracerebral recording study. *PloS one*, *10*(11), e0140778.

Strauss, G. P., & Gold, J. M. (2012). A new perspective on anhedonia in schizophrenia. *American Journal of Psychiatry*, *169*(4), 364-373.

Strauss, G. P., Waltz, J. A., & Gold, J. M. (2014). A review of reward processing and motivational impairment in schizophrenia. *Schizophrenia bulletin*, *40*(Suppl_2), S107-S116.

Su, J. H., Thomas, F. T., Kasoff, W. S., Tourdias, T., Choi, E. Y., Rutt, B. K., & Saranathan, M. (2019). Thalamus Optimized Multi Atlas Segmentation (THOMAS): fast, fully automated segmentation of thalamic nuclei from structural MRI. *Neuroimage*, *194*, 272-282.

Sui, J., Jiang, R., Bustillo, J., & Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. *Biological psychiatry*, *88*(11), 818-828.

Sweeney-Reed, C. M., Zaehle, T., Voges, J., Schmitt, F. C., Buentjen, L., Kopitzki, K., ... & Richardson-Klavehn, A. (2014). Corticothalamic phase synchrony and cross-frequency coupling predict human memory formation. *elife*, *3*, e05352.

Sweeney-Reed, C. M., Zaehle, T., Voges, J., Schmitt, F. C., Buentjen, L., Kopitzki, K., ... & Richardson-Klavehn, A. (2015). Thalamic theta phase alignment predicts human memory formation and anterior thalamic cross-frequency coupling. *Elife*, *4*, e07578.

Sylvester, C. Y. C., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, *41*(3), 357-370.

Taber, K. H., Wen, C., Khan, A., & Hurley, R. A. (2004). The limbic thalamus. *The Journal of neuropsychiatry and clinical neurosciences*, *16*(2), 127-132.

Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian–instrumental transfer. *Journal of Neuroscience*, *28*(2), 360-368.

Tambini, A., Ketz, N., & Davachi, L. (2010). Enhanced brain correlations during rest are related to memory for recent experiences. *Neuron*, *65*(2), 280-290.

Tan, J. B., Orlando, I. F., Whyte, C., Bryant, A. G., Munn, B. R., Baracchini, G., ... & Shine, J. M. (2025). Cerebellar and subcortical contributions to working memory manipulation. *Communications Biology*, *8*(1), 1028.

Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 216-220.

Taylor, P. A., Reynolds, R. C., Calhoun, V., Gonzalez-Castillo, J., Handwerker, D. A., Bandettini, P. A., ... & Chen, G. (2023). Highlight results, don't hide them: Enhance interpretation, reduce biases and improve reproducibility. NeuroImage, 274, 120138.

Test, W. C. S. (1993). Wisconsin Card Sorting Test. *Odessa, FL: Psychological Assessment Resources*.

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J. B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage*, *35*(1), 105-120.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515-518.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size.

Tomie, A., Grimes, K. L., & Pohorecky, L. A. (2008). Behavioral characteristics and neurobiological substrates shared by Pavlovian sign-tracking and drug abuse. *Brain research reviews*, *58*(1), 121-135.

Tschernegg, M., Pletzer, B., Schwartenbeck, P., Ludersdorfer, P., Hoffmann, U., & Kronbichler, M. (2015). Impulsivity relates to striatal gray matter volumes in humans: evidence from a delay discounting paradigm. *Frontiers in Human Neuroscience*, *9*, 384.

Tseng, H. H., Watts, J. J., Kiang, M., Suridjan, I., Wilson, A. A., Houle, S., ... & Mizrahi, R. (2018). Nigral stress-induced dopamine release in clinical high risk and antipsychotic-naïve schizophrenia. *Schizophrenia bulletin*, *44*(3), 542-551.

Tubridy, S., & Davachi, L. (2011). Medial temporal lobe contributions to episodic sequence encoding. *Cerebral cortex*, *21*(2), 272-280.

Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, *1*(381-403), 1.

Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications biology*, *1*(1), 62.

Van Buuren, M., Wagner, I. C., & Fernández, G. (2019). Functional network interactions at rest underlie individual differences in memory ability. *Learning & Memory*, *26*(1), 9-19.

Van Den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Specific somatotopic organization of functional connections of the primary motor network during resting state. *Human brain mapping*, *31*(4), 631-644.

Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European neuropsychopharmacology*, *20*(8), 519-534.

Van Den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Pol, H. E. H. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, *29*(23), 7619-7624.

Van Der Werf, Y. D., Jolles, J., Witter, M. P., & Uylings, H. B. (2003). Contributions of thalamic nuclei to declarative memory functioning. *Cortex*, *39*(4-5), 1047-1062.

Van der Werf, Y. D., Scheltens, P., Lindeboom, J., Witter, M. P., Uylings, H. B., & Jolles, J. (2003). Deficits of memory, executive functioning and attention following infarction in the thalamus; a study of 22 cases with localised lesions. *Neuropsychologia*, *41*(10), 1330-1344.

Van der Werf, Y. D., Witter, M. P., Uylings, H. B., & Jolles, J. (2000). Neuropsychology of infarctions in the thalamus: a review. *Neuropsychologia*, *38*(5), 613-627.

Van Loo, H. M., De Jonge, P., Romeijn, J. W., Kessler, R. C., & Schoevers, R. A. (2012). Data-driven subtypes of major depressive disorder: a systematic review. *BMC medicine*, *10*(1), 156.

Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. *Neuroimage*, *180*, 68-77.

Vaskinn, A., Haatveit, B., Melle, I., Andreassen, O. A., Ueland, T., & Sundet, K. (2020). Cognitive heterogeneity across schizophrenia and bipolar disorder: a cluster analysis of intellectual trajectories. *Journal of the International Neuropsychological Society*, *26*(9), 860-872.

Vatansever, D., Smallwood, J., & Jefferies, E. (2021). Varying demands for cognitive control reveals shared neural processes supporting semantic and episodic memory retrieval. *Nature communications*, *12*(1), 2134.

Veazey, R. B., Amaral, D. G., & Cowan, W. M. (1982). The morphology and connections of the posterior hypothalamus in the cynomolgus monkey (Macaca fascicularis). II. Efferent connections. *Journal of Comparative Neurology*, *207*(2), 135-156.

Versace, F., Frank, D. W., Stevens, E. M., Deweese, M. M., Guindani, M., & Schembre, S. M. (2019). The reality of "food porn": Larger brain responses to food-related cues than to erotic images predict cue-induced eating. *Psychophysiology*, *56*(4), e13309.

Versace, F., Kypriotakis, G., Basen-Engquist, K., & Schembre, S. M. (2016). Heterogeneity in brain reactivity to pleasant and food cues: evidence of sign-tracking in humans. *Social Cognitive and Affective Neuroscience*, *11*(4), 604-611.

Vertes, R. P. (2006). Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neuroscience*, *142*(1), 1-20.

Vertes, R. P., Hoover, W. B., Do Valle, A. C., Sherman, A., & Rodriguez, J. J. (2006). Efferent projections of reuniens and rhomboid nuclei of the thalamus in the rat. *Journal of comparative neurology*, *499*(5), 768-796.

Vidal, J. P., Danet, L., Péran, P., Pariente, J., Bach Cuadra, M., Zahr, N. M., ... & Saranathan, M. (2024). Robust thalamic nuclei segmentation from T1-weighted MRI using polynomial intensity transformation. *Brain Structure and Function*, *229*(5), 1087-1101.

Vogt, B. A., Pandya, D. N., & Rosene, D. L. (1987). Cingulate cortex of the rhesus monkey: I. Cytoarchitecture and thalamic afferents. *Journal of Comparative Neurology*, *262*(2), 256-270.

Vogt, B. A., Rosene, D. L., & Pandya, D. N. (1979). Thalamic and cortical afferents differentiate anterior from posterior cingulate cortex in the monkey. *Science*, *204*(4389), 205-207.

Volkow, N. D., Wang, G. J., Telang, F., Fowler, J. S., Logan, J., Childress, A. R., ... & Wong, C. (2006). Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. *Journal of neuroscience*, *26*(24), 6583-6588.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in cognitive sciences*, *9*(9), 445-453.

Wagner, I. C., van Buuren, M., & Fernández, G. (2019). Thalamo-cortical coupling during encoding and consolidation is linked to durable memory formation. *Neuroimage*, *197*, 80-92.

Wagner, I. C., van Buuren, M., Bovy, L., & Fernández, G. (2016). Parallel engagement of regions associated with encoding and later retrieval forms durable memories. *Journal of Neuroscience*, *36*(30), 7985-7995.

Waller, L., Erk, S., Pozzi, E., Toenders, Y. J., Haswell, C. C., Büttner, M., ... & Veer, I. M. (2022). ENIGMA HALFpipe: interactive, reproducible, and efficient analysis for resting-state and task-based fMRI data. *Human Brain Mapping*, *43*(9), 2727-2742.

Waltz, J. A., Schweitzer, J. B., Gold, J. M., Kurup, P. K., Ross, T. J., Jo Salmeron, B., ... & Stein, E. A. (2009). Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology*, *34*(6), 1567-1577.

Waltz, J. A., Schweitzer, J. B., Ross, T. J., Kurup, P. K., Salmeron, B. J., Rose, E. J., ... & Stein, E. A. (2010). Abnormal responses to monetary outcomes in cortex, but not in the basal ganglia, in schizophrenia. *Neuropsychopharmacology*, *35*(12), 2427-2439.

Wang, D., Li, M., Wang, M., Schoeppe, F., Ren, J., Chen, H., ... & Liu, H. (2020). Individual-specific functional connectivity markers track dimensional and categorical features of psychotic illness. *Molecular psychiatry*, *25*(9), 2119-2129.

Wang, K. S., Smith, D. V., & Delgado, M. R. (2016). Using fMRI to study reward processing in humans: past, present, and future. *Journal of neurophysiology*, *115*(3), 1664-1678.

Wang, X. J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, *84*(3), 638-654.

Wang, Y., Tang, S., Zhang, L., Bu, X., Lu, L., Li, H., ... & Huang, X. (2021). Data-driven clustering differentiates subtypes of major depressive disorder with distinct brain connectivity and symptom features. The British Journal of Psychiatry, 219(5), 606-613.

Wardle, M., Flagel, S., Lopez-Gamundi, P., & Amador, C. (2017). Measuring sign-tracking in humans: The promise and pitfalls of translation. *Drug and Alcohol Dependence*, *100*(171), e212-e213.

Warren, T. L., Tubbs, J. D., Lesh, T. A., Corona, M. B., Pakzad, S. S., Albuquerque, M. D., ... & Nord, A. S. (2024). Association of neurotransmitter pathway polygenic risk with specific symptom profiles in psychosis. *Molecular Psychiatry*, *29*(8), 2389-2398.

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature*, *217*(5132), 972-974.

Wassum, K. M., Ostlund, S. B., Loewinger, G. C., & Maidment, N. T. (2013). Phasic mesolimbic dopamine release tracks reward seeking during expression of Pavlovian-to-instrumental transfer. *Biological psychiatry*, *73*(8), 747-755.

Wassum, K. M., Ostlund, S. B., Loewinger, G. C., & Maidment, N. T. (2013). Phasic mesolimbic dopamine release tracks reward seeking during expression of Pavlovian-to-instrumental transfer. *Biological psychiatry*, *73*(8), 747-755.

Wechsler, D. (1981). WAIS-R: Manual: Wechsler adult intelligence scale-revised. *(No Title)*.

Weiland, B. J., Welsh, R. C., Yau, W. Y. W., Zucker, R. A., Zubieta, J. K., & Heitzeg, M. M. (2013). Accumbens functional connectivity during reward mediates sensation-seeking and alcohol use in high-risk youth. *Drug and alcohol dependence*, *128*(1-2), 130-139.

Weinberger, D. R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of general psychiatry*, *44*(7), 660-669.

Weinberger, D. R. (2017). Future of days past: neurodevelopment and schizophrenia. *Schizophrenia bulletin*, *43*(6), 1164-1168.

Wen, Z., Hammoud, M. Z., Siegel, C. E., Laska, E. M., Abu-Amara, D., Etkin, A., ... & Marmar, C. R. (2025). Neuroimaging-based variability in subtyping biomarkers for psychiatric heterogeneity. *Molecular Psychiatry*, *30*(5), 1966-1975.

Whitton, A. E., Treadway, M. T., & Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Current opinion in psychiatry*, *28*(1), 7-12.

Wig, G. S. (2017). Segregated systems of human brain networks. *Trends in cognitive sciences*, *21*(12), 981-996.

Wig, G. S., Grafton, S. T., Demos, K. E., Wolford, G. L., Petersen, S. E., & Kelley, W. M. (2008). Medial temporal lobe BOLD activity at rest predicts individual differences in memory ability in healthy young adults. *Proceedings of the National Academy of Sciences*, *105*(47), 18555-18560.

Wig, G. S., Laumann, T. O., Cohen, A. L., Power, J. D., Nelson, S. M., Glasser, M. F., ... & Petersen, S. E. (2014). Parcellating an individual subject's cortical and subcortical brain structures using snowball sampling of resting-state correlations. *Cerebral cortex*, *24*(8), 2036-2054.

Wilson, R. P., Colizzi, M., Bossong, M. G., Allen, P., Kempton, M., & Bhattacharyya, S. (2018). The neural substrate of reward anticipation in health: a meta-analysis of fMRI findings in the monetary incentive delay task. *Neuropsychology review*, *28*(4), 496-506.

Wilson, R., Mullinger, K. J., Francis, S. T., & Mayhew, S. D. (2019). The relationship between negative BOLD responses and ERS and ERD of alpha/beta oscillations in visual and motor cortex. *Neuroimage*, *199*, 635-650.

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*, *338*(6104), 270-273.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, *17*(5), 766-780.

Wittmann, B. C., Schiltz, K., Boehler, C. N., & Düzel, E. (2008). Mesolimbic interaction of emotional valence and reward improves memory formation. *Neuropsychologia*, *46*(4), 1000-1008.

Woicik, P. A., Stewart, S. H., Pihl, R. O., & Conrod, P. J. (2009). The substance use risk profile scale: a scale measuring traits linked to reinforcement-specific substance use profiles. *Addictive behaviors*, *34*(12), 1042-1055.

Wolfers, T., Buitelaar, J. K., Beckmann, C. F., Franke, B., & Marquand, A. F. (2015). From estimating activation locality to predicting disorder: a review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neuroscience & Biobehavioral Reviews*, *57*, 328-349.

Wolfers, T., Doan, N. T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., ... & Marquand, A. F. (2018). Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA psychiatry*, *75*(11), 1146-1155.

Wolff, M., & Halassa, M. M. (2024). The mediodorsal thalamus in executive control. *Neuron*, *112*(6), 893-908.

Wolff, M., Alcaraz, F., Marchand, A. R., & Coutureau, E. (2015). Functional heterogeneity of the limbic thalamus: from hippocampal to cortical functions. *Neuroscience & Biobehavioral Reviews*, *54*, 120-130.

Wolke, S. A., Mehta, M. A., O'Daly, O., Zelaya, F., Zahreddine, N., Keren, H., ... & Stringaris, A. (2019). Modulation of anterior cingulate cortex reward and penalty signalling in medication-naive young-adult subjects with depressive symptoms following acute dose lurasidone. *Psychological Medicine*, *49*(8), 1365-1377.

Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of cognitive neuroscience*, *24*(7), 1532-1547.

Woo, C. W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: brain models in translational neuroimaging. *Nature neuroscience*, *20*(3), 365-377.

Wood, M. L., & Mark Henkelman, R. (1985). Truncation artifacts in magnetic resonance imaging. *Magnetic resonance in medicine*, *2*(6), 517-526.

Woodward, N. D., & Heckers, S. (2016). Mapping thalamocortical functional connectivity in chronic and early stages of psychotic disorders. *Biological psychiatry*, *79*(12), 1016-1025.

Woodward, N. D., Giraldo-Chica, M., Rogers, B. & Cascio, C. J. T halamocortical dysconnectivity in autism spectrum disorder: an analysis of the Autism Brain Imaging Data Exchange. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 2, 76–84 (2017).

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage*, *14*(6), 1370-1386.

Yadav, N., Noble, C., Niemeyer, J. E., Terceros, A., Victor, J., Liston, C., & Rajasethupathy, P. (2022). Prefrontal feature representations drive memory recall. *Nature*, *608*(7921), 153-160.

Yager, L. M., Pitchers, K. K., Flagel, S. B., & Robinson, T. E. (2015). Individual variation in the motivational and neurobiological effects of an opioid cue. *Neuropsychopharmacology*, *40*(5), 1269-1277.

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., ... & Kawato, M. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature communications*, *7*(1), 11254.

Yakovlev, P. I., Locke, S., Koskoff, D. Y., & Patton, R. A. (1960). Limbic nuclei of thalamus and connections of limbic cortex. *Archives of Neurology*, *3*(6), 620-641.

Yang, Z., Zuo, X. N., McMahon, K. L., Craddock, R. C., Kelly, C., de Zubicaray, G. I., ... & Wright, M. J. (2016). Genetic and environmental contributions to functional connectivity architecture of the human brain. *Cerebral cortex*, *26*(5), 2341-2352.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100-1122.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*.

Yeterian, E. H., & Pandya, D. N. (1988). Corticothalamic connections of paralimbic regions in the rhesus monkey. *Journal of Comparative Neurology*, *269*(1), 130-146.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characterstics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1415.

Yoon, Y. B., & Jung, W. H. (2025). The role of the ventral striatum in the relationship between impulsive decision-making and emotional self-regulation by cognitive reappraisal. *Scientific Reports*, *15*(1), 27830.

Yoshida, K., Shimizu, Y., Yoshimoto, J., Takamura, M., Okada, G., Okamoto, Y., ... & Doya, K. (2017). Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression. *PloS one*, *12*(7), e0179638.

Yoshikawa, A., Masaoka, Y., Yoshida, M., Koiwa, N., Honma, M., Watanabe, K., ... & Izumizaki, M. (2020). Heart rate and respiration affect the functional connectivity of default mode network in resting-state functional magnetic resonance imaging. *Frontiers in Neuroscience*, *14*, 631.

Yu, H., Chapman, B., Di Florio, A., Eischen, E., Gotz, D., Jacob, M., & Blair, R. H. (2019). Bootstrapping estimates of stability for clusters, observations and model selection. *Computational Statistics*, *34*(1), 349-372.

Zald, D. H., & Treadway, M. T. (2017). Reward processing, neuroeconomics, and psychopathology. *Annual review of clinical psychology*, *13*(1), 471-495.

Zeng, J., Yan, J., Cao, H., Su, Y., Song, Y., Luo, Y., & Yang, X. (2022). Neural substrates of reward anticipation and outcome in schizophrenia: a meta-analysis of fMRI findings in the monetary incentive delay task. *Translational Psychiatry*, *12*(1), 448.

Zhang, W., Perry, K. W., Wong, D. T., Potts, B. D., Bao, J., Tollefson, G. D., & Bymaster, F. P. (2000). Synergistic effects of olanzapine and other antipsychotic agents in combination with fluoxetine on norepinephrine and dopamine release in rat prefrontal cortex. *Neuropsychopharmacology*, *23*(3), 250-262.

Zhang, W., Sweeney, J. A., Bishop, J. R., Gong, Q., & Lui, S. (2023). Biological subtyping of psychiatric syndromes as a pathway for advances in drug discovery and personalized medicine. *Nature Mental Health*, *1*(2), 88-99.

Zhou, D., Luo, J., Yi, Z., Li, Y., Yang, S., Verguts, T., & Chen, Q. (2020). The hand-lateralization of spatial associations in working memory and long-term memory. *Quarterly Journal of Experimental Psychology*, *73*(8), 1150-1161.

Zhou, H., Schafer, R. J., & Desimone, R. (2016). Pulvinar-cortex interactions in vision and attention. *Neuron*, *89*(1), 209-220.

Zhu, X., Tang, H. D., Dong, W. Y., Kang, F., Liu, A., Mao, Y., ... & Zhang, Z. (2021). Distinct thalamocortical circuits underlie allodynia induced by tissue injury and by depression-like states. *Nature neuroscience*, *24*(4), 542-553.

Zlotnik, G., & Vansintjan, A. (2019). Memory: An extended definition. *Frontiers in psychology*, *10*, 487439.