

Unsupervised identification of internal perceptual states influencing psychomotor performance

Vardal, Ozan; Karapanagiotidis, Theodoros; Stafford, Tom; Drachen, Anders; Wade, Alex

Published in:
NeuroImage

DOI:
[10.1016/j.neuroimage.2025.121134](https://doi.org/10.1016/j.neuroimage.2025.121134)

Publication date:
2025

Document version:
Final published version

Document license:
CC BY

Citation for pulished version (APA):
Vardal, O., Karapanagiotidis, T., Stafford, T., Drachen, A., & Wade, A. (2025). Unsupervised identification of internal perceptual states influencing psychomotor performance. *NeuroImage*, 310, Article 121134.
<https://doi.org/10.1016/j.neuroimage.2025.121134>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk



Unsupervised identification of internal perceptual states influencing psychomotor performance

Ozan Vardal ^a,*, Theodoros Karapanagiotidis ^b, Tom Stafford ^c, Anders Drachen ^d, Alex Wade ^e

^a Institute of Science and Technology for Brain-inspired Intelligence, Fudan University, No. 825, Zhangheng Road, Zhangjiang High Tech Park, Shanghai, 200120, China

^b School of Psychology, University of Sussex, Pevensey 1 Building, Falmer, BN1 9QH, United Kingdom

^c Department of Psychology, University of Sheffield, ICOS Building, 219 Portobello, Sheffield, S1 4DP, United Kingdom

^d Maersk Mc-Kinney Møller Institute, University of Southern Denmark, Campusvej 55, Odense, DK-5230, Denmark

^e Department of Psychology, University of York, Heslington, York, YO10 5DD, United Kingdom

ARTICLE INFO

Dataset link: <https://openneuro.org>

Keywords:

Internal states
Magnetoencephalography
Hidden Markov models
Endogenous rhythms
Video games

ABSTRACT

When humans perform repetitive tasks over long periods, their performance is not constant. People drift in and out of states that might be loosely categorised as engagement, disengagement or 'flow' and these states will be reflected in aspects of their performance (for example, reaction time, accuracy, criteria shifts and potentially longer-term strategy). Until recently it has been challenging to relate these behavioural states to the underlying neural mechanisms that generate them. Here, we acquired magnetoencephalography recordings and contemporaneous, dense behavioural data from participants performing an engaging task (Tetris) that required rapid, strategic behavioural responses over the period of an entire game. We asked whether it was possible to infer the presence of distinct behavioural states from the behavioural data and, if so, whether these states would have distinct neural correlates. We used hidden Markov Modelling to segment the behavioural time series into states with unique behavioural signatures, finding that we could identify three distinct and robust behavioural states. We then computed occipital alpha power across each state. These within-participant differences in alpha power were statistically significant, suggesting that individuals shift between behaviourally and neurally distinct states during complex performance, and that visuo-spatial attention change across these states.

1. Introduction

Digital games are a promising paradigm for research into human cognition. In recent years, researchers have used telemetry data recorded in commercial games to investigate theories of motor chunking (Thompson et al., 2017, 2019), ageing (Thompson et al., 2014a,b), and sleep consolidation (Stafford and Haasnoot, 2017), among some examples of problem domains. On the other end of the methodological spectrum, games have been tailor-made for laboratory research on neural plasticity (Lee et al., 2012; Voss et al., 2012), skill transfer effects (Anderson et al., 2011; Boot et al., 2010), and have been used to test cognitive architectures that model human learning as a whole (Anderson et al., 2019).

A challenge in game-based research is the reliance on total or end-game scores to assess performance. Games are often complex and engaging by design, presenting players with varied challenges across

interactions. This complexity makes analysis using total scores problematic, as moment-to-moment variability can obscure underlying factors that change across trials and sessions, such as players' responses to novel situations. Relatedly, individuals may alternate between periods of good and bad performance within single sessions despite proficiency in the game.

To address these challenges, studies have proposed examining more complex measures of behaviour, such as patterns of players' control inputs and decisions (Towne et al., 2014; Gobet, 2017; Stafford and Vaci, 2022). Additionally, it has been suggested that performance fluctuations can be explained in terms of changes in latent cognitive factors, such as shifts in attention (van Maanen et al., 2011; Renart and Machens, 2014). Recent applications of unsupervised learning techniques have further advanced this field by modelling performance as underpinned by discrete shifts in internal states (Chen, 2015; Calhoun et al., 2019; Ashwood et al., 2022). Inspired by these approaches, this study aims to identify shifts in latent states as individuals engage in

* Corresponding author.

E-mail addresses: ozanvardal@gmail.com, ozanvardal@fudan.edu.cn (O. Vardal).

<https://doi.org/10.1016/j.neuroimage.2025.121134>

Received 8 February 2024; Received in revised form 25 January 2025; Accepted 7 March 2025

Available online 16 March 2025

1053-8119/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

a complex psychomotor task with high ecological validity. By combining high-dimensional behavioural data from a laboratory version of a commercial video game (Tetris (Lindstedt and Gray, 2015)) with magnetoencephalography (MEG) recordings and unsupervised learning techniques, we show that individuals playing a video game shift between behaviourally and neurally distinct states.

1.1. Identifying latent states using hidden Markov models

Previous studies in this vein have shown that brain states during waking behaviour shape the dynamics of cortical activity, stimulus-response and task performance in different animals and in recent cases have been able to describe behaviour with striking accuracy (Vidaurre et al., 2019; Eyjolfsson et al., 2017; Wiltchko et al., 2015; Calhoun et al., 2019). For instance, researchers investigating the acoustic courtship behaviours of fruit flies were able to precisely predict distinct patterns of song behaviour by statistically inferring latent states from flies' movement data, capturing 84.6% of all remaining song patterning information that previous models lacking a latent state component could not explain Calhoun et al. (2019). Accurate segmentation of the latent state sequence allowed detailed description of the flies' sensorimotor-strategies corresponding to each state and, following an optogenetics component of the study, identification of the neurons responsible for switching between states. Similarly, latent state models of rodent decision-making can accurately predict choice strategies corresponding to states of optimal engagement versus bias (Roy et al., 2021; Ashwood et al., 2022), permitting reliable detection of blocks of trials with heterogeneous error-rates, as opposed to previous models that would assume errors are scattered throughout all trials in a session with equal probability.

While varying in scope and problem domain, common to some of these studies is their use of Hidden Markov Models (HMM), which model observable processes in terms of an underlying sequence of unobservable (i.e., hidden) states that transition with fixed probabilities. This approach first involves specification of the number of states that are assumed to influence the process as well as the probabilities of the model initialising each state, following which the parameters of the model are estimated via maximum likelihood estimation. As described previously, successful validation of HMMs in cognitive task environments allows post-hoc relation of observable behavioural dynamics to underlying brain states, resulting in rich descriptions of moment-to-moment performance and cognition. These can exist at the group level but also the individual level, for instance by analysing how much time individual participants spend in each state and how often they transition between states (Vidaurre et al., 2018).

Depending on the objectives of modelling, the specification of the states can take on different forms. In the examples outlined above, researchers specified a distinct generalised linear model (GLM) for each state that acted as a psychometric function mapping stimulus to sensorimotor response (Calhoun et al., 2019; Ashwood et al., 2022). This approach paired the HMM with a previously tested GLM with proven application in tasks with discrete outputs. A similar approach tested stage-wise models of human skill acquisition by pairing each latent state (i.e., stage of learning) with a different speedup function describing participants' response latencies in a novel arithmetic task (Tenison and Anderson, 2016). Other investigations of latent states in humans have included the identification of brain states during wakeful rest or motor task performance by fitting HMMs to electrophysiological time series (Vidaurre et al., 2018, 2019; Karapanagiotidis et al., 2020). These studies have demonstrated that HMMs provide a flexible and task-agnostic framework for segmenting behavioural or neural time series into meaningful state-dependent epochs.

1.2. Aims and approach

In this study, our aims were to (1) use an HMM approach to identify and characterise behavioural states that occur during an ecologically valid psychomotor task and (2) relate these states to neural markers of attention using MEG. To achieve these aims, we used a laboratory version of a commercial game (Tetris (Lindstedt and Gray, 2015)) that records detailed performance metrics. To our knowledge, this is the first investigation of latent states to be conducted in an ecologically valid context of human psychomotor performance.

We first analysed a purely behavioural experiment (Experiment 1) using secondary experimental data from an independent lab. Data from Experiment 1 were used to decompose game-state and behavioural logs into distinct features describing Tetris performance, and to test the capacity of these features to describe game-play. We then conducted a new experiment (Experiment 2) inside the MEG scanner where we recorded neuronal activity while subjects played the same task as in Experiment 1. We applied dimensionality reduction techniques validated in Experiment 1 to extract performance components from our new behavioural data, used HMMs to identify distinct behavioural states, and related these states to contemporaneous neural activity.

The focus of the neural analysis was on occipital alpha power, as there is a well-established relationship between occipital alpha and visual spatial attention (Foster and Awh, 2019; Peylo et al., 2021). More specifically, direction of the attentional "spotlight" from one location in the visual field to another in the absence of eye movement has been shown to correlate with modulations in the amplitude of alpha rhythm in both the parietal and primary visual cortices (Yamagishi et al., 2003; Sauseng et al., 2005). We hypothesised that occipital alpha power would differ between states, reflecting changes in participant engagement throughout the task.

2. Methods

Here, we provide a brief overview of the overall study design, followed by a detailed description of stimuli, procedures, and analysis for both experiments. We first detail the stimulus, followed by descriptions of samples and procedures for each experiment, noting adjustments made for neuroimaging in Experiment 2. The analysis protocol spanning both data sets, covering feature extraction, data pre-processing, HMM and MEG analysis, is then fully described.

2.1. Experimental overview and rationale

Our study comprised reanalysis of a large behavioural data set from Lindstedt and Gray (2019) (Experiment 1), followed by a novel experiment involving simultaneous behavioural and MEG data (Experiment 2). The task in both experiments was "Meta-T", a laboratory adaptation of Tetris. In Experiment 1, we validated a set of orthogonal features describing Tetris performance using Principal Component Analysis (PCA) and an analysis of expert-novice differences. After collecting novel behavioural and MEG data in Experiment 2, we used the PCA weights from Experiment 1 to extract behavioural features from Meta-T, before applying Hidden Markov modelling on the resultant time series to identify distinct behavioural states during gameplay. These states were then related to contemporaneous neural activity recorded via MEG, focusing on occipital alpha power as an attentional marker.

The rationale for this two-part approach was to establish robust behavioural features in a large sample before extracting and analysing these features in combination with neural data collected from a smaller sample. Fig. 1 provides an overview of the experimental and analysis pipelines.

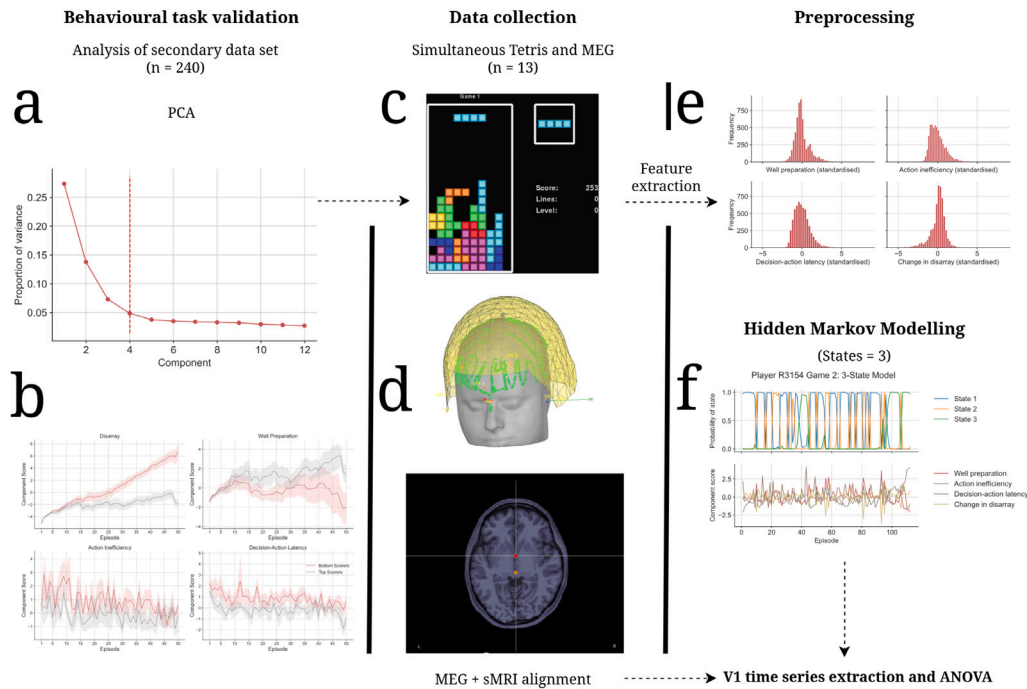


Fig. 1. Overview of experimental and analysis pipelines. **a**, We extracted 4 orthogonal behavioural features using an archival data set of 240 participants. **b**, Feature extraction was validated by comparing performance trajectories of players in the top and bottom deciles of performance, confirming Meta-T reliably produces meaningful behavioural information relating to motor planning and execution. **c**, Simultaneous behavioural and MEG data were collected using an identical version of Meta-T. **d**, Structural MRI scans and MEG sources were estimated using minimum-norm estimation after co-registration. **e**, PCA weights extracted from Experiment 1 were then applied to the behavioural data obtained in Experiment 2. **f**, After fitting a three-state HMM to the resulting principal component scores, we aligned MEG and sMRI data for each participant, and computed alpha power in the visual cortex across each HMM state.

2.2. Experimental design

2.2.1. Stimuli

The behavioural task was an implementation of Tetris called “Meta-T”, developed by Lindstedt and colleagues to study human expertise and learning (Lindstedt and Gray, 2015). Meta-T is a near-identical representation of the original Nintendo Entertainment System (NES) Tetris, with the exception of minor visual differences relating to the use of Python (using pygame (Shinners, 2011)) as the development language. Importantly, Meta-T possesses several features that make it suitable for cognitive science, and has been used as a task environment in several published studies on human and machine expertise (Lindstedt and Gray, 2013; Sibert et al., 2017; Sibert and Gray, 2018; Lindstedt and Gray, 2019).

First, Meta-T outputs data files at the end of each session that, in addition to detailing the participant’s ID and other session-specific information, include a log of post-game summary statistics for each game, a log of game-state (e.g., pile structure) and behavioural (e.g., action latencies) information describing performance for each “episode” of play (i.e., the time between tetromino appearance and tetromino drop), as well as a complete log containing key-press information at the millisecond level (See Lindstedt and Gray (2015) for an exhaustive description of logged variables). Second, researchers can modify game parameters such as screen size, game length, or difficulty curve, by editing the default configuration file. In doing so, researchers can constrain participant behaviour to bespoke experimental conditions according to the requirements of their research question.

We note minor differences in the stimulus code between experiments due to adjustments made for MEG scanning. In Experiment 2, participants played an updated Python3 version of the Tetris implementation used in Experiment 1 (Fig. 2). The code was further adapted to accommodate a fibre optic response interface (Cambridge Research Systems 905 package) connecting the stimulus computer and MEG scanner, allowing participants to use a non-electronic, non-magnetic

five button response pad to play Meta-T without adding additional noise to the scans. Finally, Meta-T was configured to send triggers to the MEG record upon the occurrence of salient events: button inputs, tetromino appearances/drops, line clears, and game start/end.

2.3. Experimental procedure

2.3.1. Experiment 1: Participants

We used a secondary, experimental data set of Meta-T gameplay made public by Lindstedt and colleagues through the Open Science Framework (<https://osf.io/78ebg/>). We describe the data set here following the original experimenters’ (Lindstedt and Gray, 2019) reports as well as our own examination of its contents. These data were collected from 240 participants under laboratory conditions and informed consent was provided by the Rensselaer IRB.

2.3.2. Experiment 1: Stimulus task

Participants were seated in front of a computer and instructed to play Meta-T for 50 min using a provided NES controller, connected to the computer via USB. Players repeatedly engaged in successive games of Meta-T until the 50-minute period elapsed, restarting games upon failure. The data set comprise three log files, each detailing all 240 participants’ task engagement at three different levels of time: one describing behaviour at the time of each button input, one describing behaviour in the time spanning the appearance to dropping of each tetromino, and one summarising behaviour at the level of the entire game. We concentrated our analyses on logs of tetromino drops at each game, as these provided the highest density of measures across all log files.

2.3.3. Experiment 2: Participants

15 healthy, right-handed participants were recruited through the York Neuroimaging Centre (YNiC) participant pool (United Kingdom).

All participants provided informed consent, and the study was approved by the York Neuroimaging Centre ethics committee. All participants were familiar with playing Tetris, and provided a self-report of their proficiency on a 5 point Likert scale ($M=3.08$, $SD=1.04$), as well as their proficiency in digital games in general ($M=3.38$, $SD=1.19$). Data from 2 participants were excluded from analysis due to poor MEG data quality, resulting in a final sample of $n = 13$ participants (4 female, $M_{age} = 33$, $SD_{age} = 11.31$).

2.3.4. Experiment 2: MEG and MRI set-up

MEG scanning was conducted using a 4D Neuroimaging Magnes WH3600 scanner (248 channels + 23 reference channels) at YNIC. Data were recorded at a sampling rate of 500 Hz and then downsampled to 200 Hz. Prior to scanning, five fiducial head-coils were attached to each participant's head with hypoallergenic tape. Facial landmarks (nasion, left and right preauricular) and head shape were then recorded using a Polhemus Traktrak 3D digitiser. To assess head movement inside the scanner helmet, we measured the position of the head-coils before and after every scan, and then compared these measurements to the spatial relation between head-coils recorded outside of the scanner. Movement < 0.5 cm was our acceptance threshold for head movement, beyond which we reran our coil-on-head (CoH) scan to confirm any discrepancies in coil position and to subsequently recalculate coil positions using non-displaced coils. To estimate the neural sources of our MEG recordings, MEG data were co-registered with high-resolution structural MRI images. These were $1 \times 1 \times 1$ mm MPRAGE T1-weighted structural MRI scans, acquired for each participant using a Siemens Prisma 3T MRI scanner. The Freesurfer pipeline (Dale et al., 1999; Fischl et al., 2004) was used to perform image segmentation and cortical reconstruction.

2.3.5. Experiment 2: Stimulus task

After being briefed and prepared for scanning, each participant was seated in a dark, magnetically shielded room under the scanner. Participants were given some time to practise playing Meta-T in the scanner until they reported feeling well-adjusted to the button inputs, during which time the data acquisition software was configured for scanning. Participants played Meta-T in a seated position and were instructed to minimise head movement, while playing until game over. Meta-T was configured to run in a full-screen environment without in-game music. Tetromino sequences were randomised between games but standardised across participants using fixed numerical seeds. Games were played until loss, with all other settings matching Classic Tetris defaults.

Each scan was initiated five seconds before the start of each game and scan duration varied for each participant depending on their performance across games (i.e., better performance resulted in longer games). Each scan and concurrent game was preceded and followed by a CoH scan, allowing us to assess head movement while the participant took a short break. Each participant typically played two or three games, resulting in an average acquisition duration of $M = 7.84$ min per game ($SD = 2.88$) and an average total acquisition duration of $M = 21.55$ min per participant ($SD = 4.54$).

2.4. Analysis

Here, we describe analysis of behavioural data, describing extraction and validation of behavioural features through PCA on Experiment 1. This is followed by a description of analogous feature extraction for Experiment 2 and subsequent HMM analysis. We then describe pre-processing of MEG data from Experiment 2, and analysis of neural activity across HMM states.

Pre-processing and analysis was performed using both Python 3.8 (Van Rossum and Drake, 2009) and MatLab/Brainstorm (Tadel et al., 2019). All analysis pipelines and supporting software are publicly available at <https://github.com/ozvar/tetrisMEG>, together with details regarding requisite software dependencies.

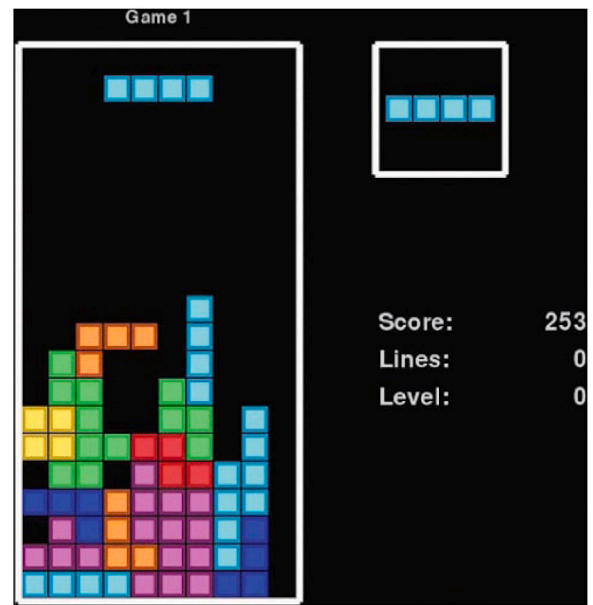


Fig. 2. Depiction of Meta-T user interface. The left side of the screen shows the current board, including the tetromino that is currently being controlled above it. The top right side of the screen shows the next tetromino that will be played after the current one is dropped. The player is also presented with the current game number, their current score, number of cleared lines, and the current level.

Source: Image taken from Lindstedt and Gray (2019).

2.4.1. Experiment 1: Feature extraction and behavioural analysis

As mentioned previously, Meta-T captures a wealth of information throughout gameplay. Each row in the episodic log file details over 60 variables for the current tetromino drop, including:

1. Features summarising the session (e.g., participant ID, game number, timestamp),
2. Game state features relating to the tetromino (current and next tetromino, current tetromino position),
3. Game state features describing the pile (e.g., height, circumference, number of unplayable cells)
4. Motor execution features (e.g., number of control inputs, latency before and between actions),
5. Features describing tetromino placement (e.g., number of lines cleared, landing height).

Our first aim was to validate a dimensionality reduction protocol using the behavioural data set from Experiment 1. We performed a PCA using the sklearn package (Pedregosa et al., 2011), focusing on episodic logs describing behaviour and game-state at the level of each tetromino drop, as these logs provided the greatest breadth of information relating to moment-to-moment performance. Each row in these logs corresponded to one tetromino drop, describing the input behaviours from tetromino appearance to drop, changes to the game-state following the tetromino drop, as well as summary variables describing participant and session related information.

After inspecting data for outliers and removing two players who never progressed past level 0, we extracted a subset of variables related to performance, excluding session (e.g., subject ID, time stamp) and game-state descriptors (e.g., current tetromino, tetromino orientation). PCA was performed on this trimmed data set (see Table A.1 for a list of variables retained for PCA), initially with an unconstrained number of components. We identified the optimal number of components to retain using a scree plot (Figure A.7), retaining four components that explained 53.3% of the variance: Disarray, Well preparation, Action inefficiency, and Decision-action latency (see Section 3.1 for a detailed description of retained features following PCA).

To validate the meaningfulness of our measures, we compared top and bottom performers, hypothesising significant differences in component scores across time between the groups. Players were sorted based on their average scores in their first three games and we selected the top and bottom quintiles for comparison. We then visualised the averaged trajectory of each component over the first 50 episodes of the first game for both groups, allowing us to examine how these measures differentiate between skill levels. These differences were tested for statistical significance through a mixed ANOVA of each performance component, with scoring quintile as the between subjects factor, and tetromino drop as the within-subjects factor.

2.4.2. Experiment 2: Feature extraction and validation

After validating feature extraction on the archival data set from Experiment 1, we fit the weights of our PCA model to transform the behavioural data from Experiment 2, producing for each participant from our novel sample an analogous time series of four components describing Meta-T performance. We used the mathematical difference of disarray rather than its absolute value to capture moment-to-moment impacts on pile structure. The scores of each component were then standardised to permit comparison between components with different scales. Distributions of performance component scores across all tetromino drops in the sample are depicted in the Appendix (Figure A.8).

To verify that these behavioural measures remained meaningful in the scanner environment of Experiment 2, we examined correlations between participants' game performance and component scores. For each of the first two games of each participant ($n = 13$ as every participant played at least two games), we computed average z-scored component scores across all tetromino drops and analysed their relationships with final game scores using Spearman rank correlations.

2.4.3. Experiment 2: Hidden Markov modelling

Trigger timestamps from the MEG record were imported into the behavioural record for use in HMM analysis, as the MEG record reflected the ground truth of event timings. We adjusted for the ~5 s difference in initiation times between MEG scans and Meta-T games by subtracting the game start trigger in the MEG record from the episode duration column in the behavioural record. This was to avoid exaggerating the duration of each tetromino drop in the behavioural data by the initial task-free scan time, which was crucial for accurate analysis of state temporal dynamics (see below).

Model fitting. We used the Python `hmmlearn` package (an open source module with an API similar to `scikit-learn`; Anon (2022)), to fit a three-state HMM to the time series of PCA-derived performance variables, where each point in the time series describes participant performance at each tetromino drop. Our model was fit to our data at the group-level by concatenating the data across all of our participants and games. This approach aligns with established methods for group-level HMM analysis of behavioural and neural data across both human and animal studies (Ashwood et al., 2022; Calhoun et al., 2019; Karapanagiotidis et al., 2020; Vidaurre et al., 2018). We fit a Gaussian HMM (the observations are assumed to be well-described by a Gaussian distribution) with a diagonal covariance matrix and a 200 iteration upper bound for training, ensuring that the Expectation Maximisation (EM) algorithm stopped either after 200 iterations or on convergence to a maximally likely solution before reaching the iteration limit. As the EM algorithm is gradient-based and may therefore converge to local optima, we ran multiple courses of model fitting with different initialisations (a random initial transition matrix for the states) but otherwise identical parameters. We then chose the model with the highest log-likelihood for the remainder of our analyses. As an additional check of model robustness, we compared the log-likelihood of our true model to a randomised model that we produced by fitting an HMM with identical parameters to a randomly shuffled time series of our observations.

We observed a consistently higher model fit in our true model as compared to our chance model, with Bayesian Information Criterion (BIC) values of 89,525.36 and 98,337.01 respectively (where lower values are indicative of the preferred model) (Visser et al., 2002).

We chose a three-state model assuming three modes of engagement with Meta-T: a default state where participants were engaged and attentive, a performant state where participants were both engaged and playing optimally, and a “panic” state involving suboptimal moves and blunders, potentially relating to inattention. States were assigned to each point in the behavioural time series using the Viterbi algorithm, which finds the most likely sequence of hidden states given the observations. It is worth noting that due to the variable amount of time taken by players to drop each tetromino, each point in the resulting HMM state time series was also of variable length.

State temporal dynamics. We first assessed the temporal dynamics of the model by interpreting the transition matrix. This is the central output of the model and describes the probability of participants switching between each pair of states from one tetromino to the next. We then computed the fractional occupancy of each state, that is, the fraction of total time that is spent by our sample in each state, both in the data set as a whole as well as in each individual game (Fig. 5). Previous applications of HMMs to the analysis of human brain dynamics have evaluated HMM validity by examining how state occupancy is distributed across participants (Karapanagiotidis et al., 2020; Vidaurre et al., 2018). An effective HMM would be expected to output state sequences that show participants occupying multiple states without large discrepancies in state occupancy (suggestive of single states overwhelming entire participants or recordings). One statistic that reflects this requirement is the maximum fractional occupancy, that is, the fraction of time taken by the state that occupies the most amount of time in a given data acquisition or participant. To examine this criterion, we visualised our transition matrix together with a bar chart depicting fractional occupancy in each state, as well as a histogram of maximum fractional occupancy across all data acquisitions. In addition to a group level plot of maximum fractional occupancy, we produced a plot of state fractional occupancies unique to each participant to check for outliers who may have been skewing the overall model.

State performance dynamics. We analysed state performance dynamics by visualising group-level state profiles as well as individual game sequences (Fig. 6). To interpret state-performance dynamics across the entire sample, we referred to the demeaned score of each component within each state, which is the primary Gaussian emission returned by our model. These scores provide insight into the relative importance of each component within each state. These were tabulated together with violin plots of component scores for each state, which show how component scores were distributed across states. To inspect state sequences and performance for individual games, we visualised the time series of observed performance components individual games in parallel to the time series of posterior probabilities; a secondary output of our model that describes the probability of each of the three states being active given our observations for any given participant and game. By plotting these two time series in parallel for exemplar games, it becomes possible to visually relate patterns of performance to particular states in any given segment of our data. Additionally, to assess potential differences in eye movements across states, we estimated blink events from the MEG data using ICA (see below) to compare blink rates across HMM states. These results are presented in the Appendix (Figure A.13).

2.4.4. Experiment 2: MEG data pre-processing

Analysis and pre-processing of MEG data were performed in MatLab using Brainstorm (Tadel et al., 2019). Data were first band-pass filtered between 1 and 40 Hz using a finite impulse response filter. We performed an Independent Component Analysis (using the `infomax` algorithm; Bell and Sejnowski (1995)) to identify and reject components capturing physiological artefacts such as blinks and heartbeats.

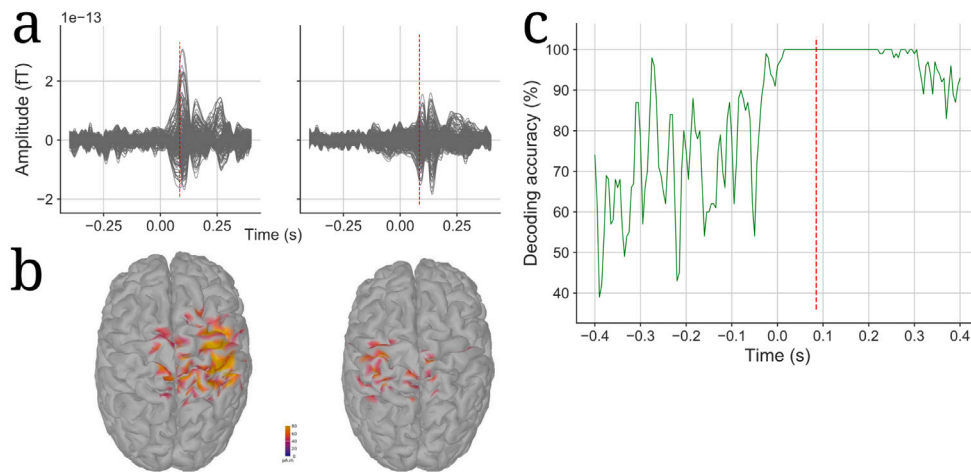


Fig. 3. SVM decoding of left versus right button inputs for participant R3154. **a**, Averaged 244 channel MEG time series from -400 ms to 400 ms for left and right translation inputs respectively. **b**, Heat maps of cortical source estimates at 0.085 s after the input was registered by Meta-T, corresponding to temporal peaks of amplitudes in (a). **c**, Averaged percentage accuracy of the classifier at each millisecond of the MEG time series. The red vertical dotted line on each line plot shows the time point corresponding to the presented cortical activity.

Raw time series data for each scan were then inspected manually in epochs of 50 s, and any periods contaminated by additional artefacts were removed manually.

As a test of data synchronicity, we checked that we were able to distinguish between neural responses to button presses executed with the left versus right thumb (i.e., the buttons used to translate the tetromino left and right respectively), as this would indicate that our MEG and behavioural data were synchronised and contained meaningful information. For each participant, we trained a linear Support Vector Machine using the `libsvm` library (Chang and Lin, 2011) to decode left versus right translation inputs using the MEG time series extracted from -400 ms to 400 ms relative to each button press. To improve computational efficiency and signal-to-noise ratio, trials from each class (i.e., left versus right translation) were randomly assigned to five folds. Trials in each fold were then sub-averaged, yielding a total of five sub-averaged trials per class. Decoding was then performed on the five sub-averaged trials following a leave-one-out cross-validation procedure, and the process was iterated 50 times. Classification accuracy was averaged across the 50 iterations for each millisecond across the trial time range, and plotted for each participant (Fig. 3). We attained a classification accuracy of 100% for every participant approximately 0.1 s after the response.

2.4.5. Experiment 2: Neural activity across HMM states

To examine brain activity across HMM states, we first aligned digitised head surfaces with reconstructed MRI images using fiducial landmarks and applied a minimum-norm source estimation algorithm to estimate source amplitudes across our regions of interest (ROIs) (Hämäläinen and Ilmoniemi, 1994). MEG source time series were extracted from ROIs parcellated using the Brodmann atlas (A.11), focusing on the primary visual cortex (V1) and motor cortex (M1) in both hemispheres.

HMM state onsets and switch times from the behavioural analysis were imported into the MEG record. MEG time series were then binned into one-second intervals, each labelled with the corresponding active HMM state. We computed the Fourier transform for each bin using the Fast Fourier Transform algorithm, with uniform bin length ensuring consistent transform lengths.

We calculated root-mean-square (RMS) amplitudes for alpha ($8\text{--}12$ Hz) in V1 and mu ($8\text{--}13$ Hz) rhythms in M1. These were computed for each HMM state and participant. To analyse the neural distinctness of HMM states within participants, we conducted one-way repeated measures ANOVAs for V1 alpha and M1 mu RMS separately. Due to observed differences between left and right hemispheres (Figure A.12), we performed separate statistical analyses for each hemisphere.

3. Results

3.1. PCA

Following PCA, we extracted four principal components that explained up to 53.3% of the variance in Meta-T performance. Table (A.1) displays the PCA loadings, describing the correlation between each variable and principal component (only correlations past 0.20 are displayed). Each component was provided a meaningful label based on how it individually loaded on Meta-T variables and the unique aspect of Tetris gameplay that it related to. These labels are in line with a previous report that identified a similar factor structure (Lindstedt and Gray, 2019). We describe each component here in detail, with variables in parentheses referring to the original names of Meta-T variables as found in the table of PCA loadings.

1. *Disarray*. Players that fail to clear lines as their Tetris pile increases in size are prone to developing an unfavourable Tetris pile. Disarray is the first principal component and it primarily loads on variables capturing this deficiency in pile structure, such as pits (e.g., “pits”, “pit_rows”, “pit_depth”), randomness of the pile (e.g., “col_trans”, “row_trans”), and pile height.
2. *Well preparation*. Achieving a high score in Tetris requires capitalising on opportunities to score bonus points, typically by clearing multiple lines with a single tetromino. Well preparation relates to the forward planning required to achieve multiple line clears, such as by reserving a single, empty column at either edge of the pile. This phenomenon is uniquely captured by the second principal component through its correlation with variables relating to well properties (e.g., “wells”, “deep_wells”) and the height difference at edge columns (i.e., “cd_1”, “cd_9”).
3. *Action inefficiency*. Action inefficiency captures inputs (e.g., rotations, translations) that are made in excess of the minimum number of inputs required to place a tetromino at its final destination (e.g., “min_rots_diff”, “min_trans_diff”). This relates to poor motor execution and planning.
4. *Decision-action latency*. This component corresponds to the initial lag (“initial_lat”) and average lag between actions (“avg_lat”) associated with each tetromino placement. It also corresponds to the local quality of placement for each tetromino (i.e., the reduction in pile height caused by placement and amount of contact with tetrominoes in the pile). Taken together we view this component as capturing both the speed and quality of decision-making as it relates to identifying optimal tetromino placement.

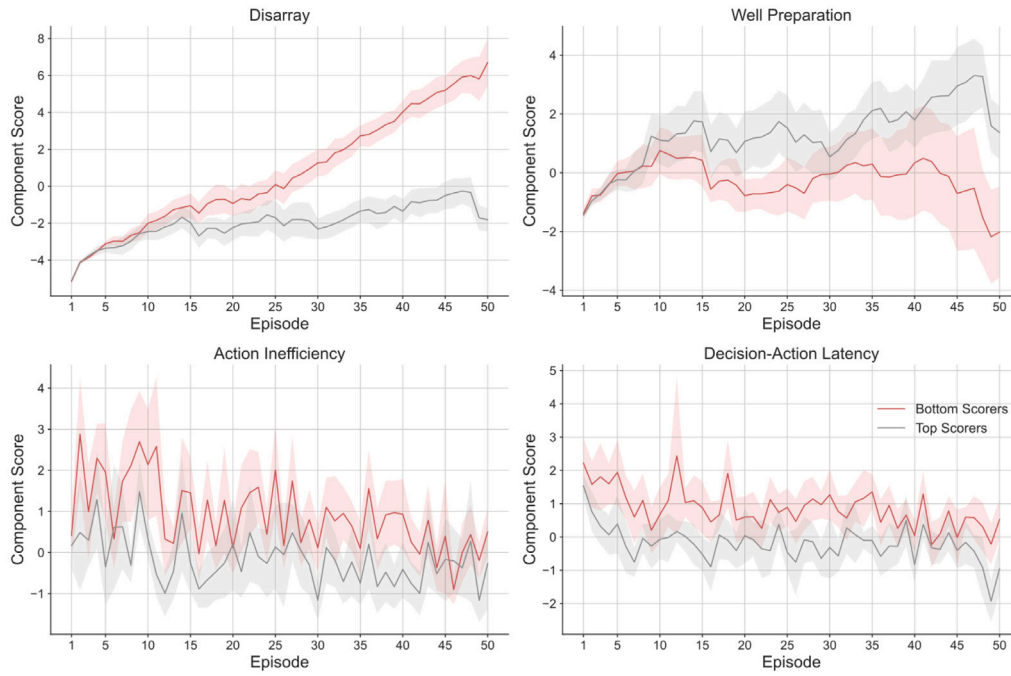


Fig. 4. Comparison of moment-to-moment performance between top and bottom scorers. Each panel depicts the mean trajectory of the respective performance component for bottom (red line) versus top scorers (grey line) for the first 50 tetrimino drops from game 1. From top left going clockwise to the bottom left panel, the subplots show trajectories for mean disarray, well preparation, decision-action latency and action inefficiency respectively. Shaded regions depict 95% confidence intervals of the mean.

3.1.1. Validation of behavioural features

Visualisation and statistical analysis confirmed that top and bottom scoring players in Experiment 1 differed significantly in their component scores across time (Fig. 4). As a reminder, participants are exposed to identical tetrimino sequences for each successive game that is played. It is therefore striking that while the peaks and troughs in action inefficiency and decision-action latency appear similar between top and bottom scorers, the top scorers were more efficient and faster in their gameplay across almost all depicted tetrimino drops. Moreover, while disarray in both groups appears to trend upwards, the upward trend is much more pronounced in bottom scorers than in the top scoring group. Conversely, bottom scoring players trend downwards in their well preparation while top scoring players trend upwards.

Between-subjects effects for disarray [$F(1, 34) = 81.93, p < 0.001$, partial $\eta^2 = 0.71$], well preparation [$F(1, 34) = 10.90, p = 0.002$, partial $\eta^2 = 0.2427$], action inefficiency [$F(1, 34) = 26.38, p < 0.001$, partial $\eta^2 = 0.44$], and decision-action latency [$F(1, 34) = 21.75, p < 0.001$, partial $\eta^2 = 0.39$] were all statistically significant.

Additionally, interaction effects between scoring decile and tetrimino drop were statistically significant for disarray [$F(1, 34) = 36.79, p < 0.001$, partial $\eta^2 = 0.52$], well preparation [$F(1, 34) = 3.84, p < 0.001$, partial $\eta^2 = 0.11$], and action inefficiency [$F(1, 34) = 1.42, p = 0.30$, partial $\eta^2 = 0.04$], but not for decision-action latency [$F(1, 34) = 1.09, p = 0.31$], suggesting that between-groups differences in the former three components grew statistically significantly more pronounced as games went on. Taken together these results demonstrated that extracted features probe meaningful aspects of Meta-T performance.

Results of the feature extraction analysis from Experiment 2 were consistent with those from Experiment 1, and correlations from Game 2 were generally stronger than in Game 1. Pile disarray showed a strong negative correlation with final game scores ($\rho = -0.79, p = 0.001$). Well preparation and action inefficiency also showed significant correlations with performance in expected directions ($\rho = 0.58, p = 0.037$ and $\rho = -0.59, p = 0.035$ respectively). While decision-action latency showed a non-significant correlation ($\rho = 0.35, p = 0.239$) with overall performance, we retained this component in the subsequent HMM analysis as it captures a theoretically distinct aspect of gameplay

— the temporal dynamics of decision-making and execution — that complements our other behavioural measures. Scatter plots of these relationships annotated with relevant statistics for both Games 1 and 2 are presented in the Appendix (Figures A.9 and A.10).

3.2. Hidden Markov model analysis

3.2.1. State temporal dynamics

The HMM transition matrix reveals distinct probabilities for state transitions (Fig. 5). In particular, the probability of switches from State 1 to State 1 and State 3 to State 3 were high (0.69 and 0.79 respectively) showing that participants have an affinity to remain in these states once they enter them. The probability of switching from State 1 to State 2 was also relatively high, while the self-transitions for State 2 and State 3 were relatively low (0.2 and 0.13 respectively), suggesting that State 2 was a transient state that participants switched to mostly from State 1 but seldom remained in.

In our case, the majority of games had maximum fractional occupancy below 0.6 (mean fractional occupancy was 0.54), demonstrating that our participants' time was shared across all states in our model. Our plot showed that a little over half of all time (~52%) on task was spent in State 1, making this the dominant state throughout task performance. This was followed by State 3, accounting for 28% of state occupancy, and State 2 with 20%. A bar plot of fractional occupancies (Figure A.14) for each participant additionally shows that participants stayed for similar durations of time in each state, indicating that there were no outliers in the sample skewing the overall model.

3.2.2. State performance dynamics

Together these visualisations inform us about how participants transition between and how frequently they occupy states, but they do not tell us how behaviour and cognition vary across states. Here, we examine state performance dynamics through two complementary approaches: analysis of group-level state profiles and visualisation of individual game sequences (Fig. 6).

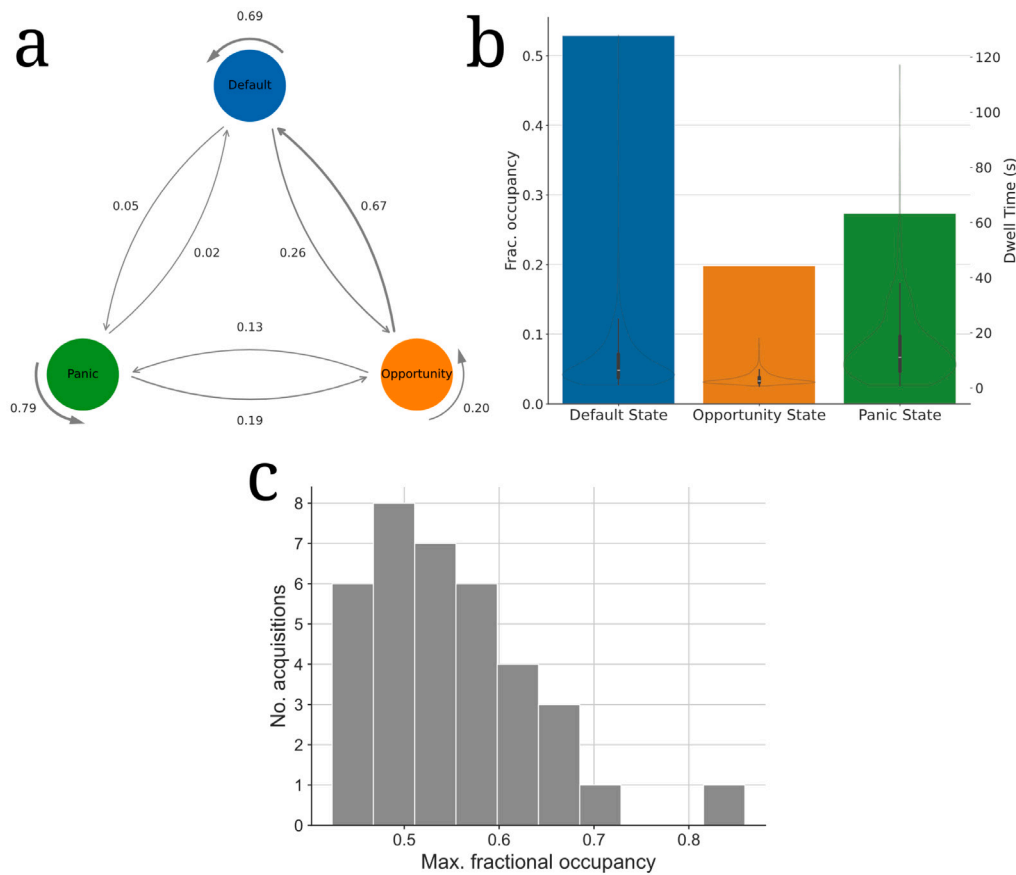


Fig. 5. Overview of three-state HMM temporal dynamics. **a**, Inferred transition matrix of the model, where states are depicted as nodes and directed connections describing the probability of switching between each pair of states. **b**, Fractional occupancy of each state, defined as the fraction of total time spent in that state, as well as violin plots depicting distributions of dwell times (the amount of time resided in each state occupancy) across each state. **c**, Distribution of maximum fractional occupancy across acquisitions in the sample. That is, for each data acquisition in the sample, the maximum fractional occupancy represents the fraction of total time spent in the state that the participant occupied for the most amount of time for that acquisition.

Default state (State 1). We first refer to the table of demeaned component scores, which offers a quantitative summary of how each performance aspect deviates from the overall mean within each state. Together with the corresponding violin plot, these scores characterise the unique performance profile of each state. We observed that State 1 is mostly characterised by efficient motor execution, slight increases to pile disarray at each tetromino drop. In line with this state being the most occupied state across the data set, we view State 1 as the “default” state, corresponding to usual, attentive Tetris gameplay.

Opportunity state (State 2). State 2 is characterised by high well preparation, large reductions to pile disarray, and moderate motor inefficiency and low decision-action latency. Participants appear to make quick decisions with relatively poor inputs in this state, corresponding to large changes to pile structure. Additionally, given that dwell times in State 2 appear to be short, we interpreted State 2 as a transient “opportunity” state, during which the participant is prepared to either score significant points through line clears, or fumble and compromise the established pile structure. We pursued this idea by calculating the percentage of tetromino drops in the opportunity state that resulted in at least one line clear. This number was 97%, confirming our initial intuition. The remaining 3% of drops in the opportunity state that did not result in a line clear were distributed across 11 players in the sample, indicating that this state does not exclusively capture cleared lines, but rather pile structure conducive to line clears that most players in the sample occasionally failed to take advantage of.

Panic state (State 3). Finally, and in contrast to the opportunity state, State 3 was characterised by the lowest well preparation, increases to pile disarray, as well as relatively high motor inefficiency and moderate but variable decision-action latency. We interpreted State 3 by considering these trends in tandem with aforementioned temporal dynamics. That is, instances of State 3 showed higher dwell times than the opportunity state, and transitions to State 3 were over twice as likely from the opportunity state than from the default state. Taken together, we interpreted State 3 as the “panic” state characterised by poor motor execution and planning, during which participants attempt to resolve difficult pile structures that likely arise from sudden and significant changes to structure that may occur in the opportunity state.

State dynamics in a single game. To illustrate these state dynamics, we examined a game from participant R3154 (top panel, Fig. 6). When well preparation is high and pile disarray is reduced, the participant enters the opportunity state which, consistent with our interpretation of the transition matrix, appears as a transient state with short dwell times. Conversely, the panic state is associated with low well preparation and increases in pile disarray. The participant’s motor executions are most efficient during the default state, evident from dips in action inefficiency following transitions to this state. This individual example corroborates our group-level analysis and provides a concrete illustration of how these states manifest during gameplay.

3.2.3. Endogenous rhythms across states

We first conducted a one-way repeated measures ANOVA to test for within-participants differences in V1 RMS alpha between states. These

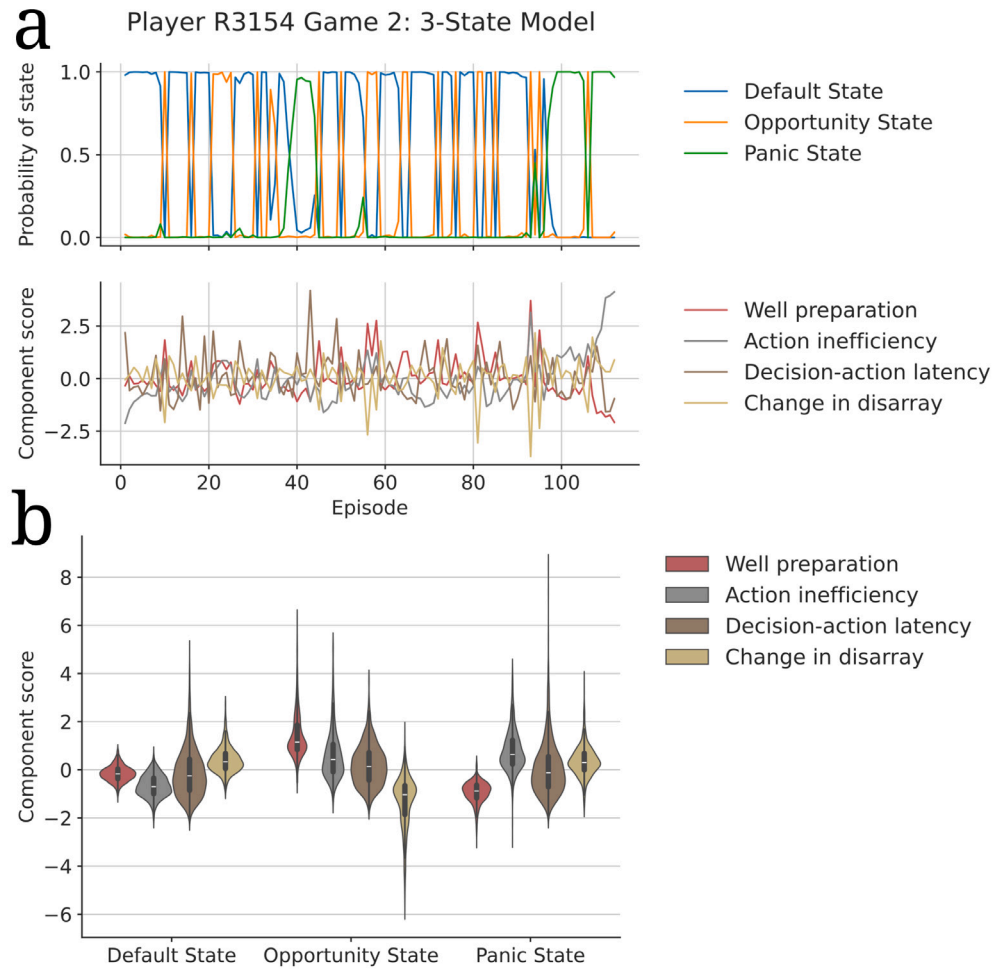


Fig. 6. Overview of behavioural profiles of each HMM state. Time series of state posterior probabilities (top) and performance component z-scores (bottom) for each tetromino drop for an example game. **b** Violin plots of performance component distributions across states. Each violin includes a box plot (black) indicating median and quartiles, with the coloured area representing the kernel density estimate. Tails show extreme values. The adjoining table shows demeaned component scores within each state.

tests were significant for both left V1 [$F(2, 24) = 3.6317$, $p = 0.0419$] and right V1 [$F(2, 24) = 4.2665$, $p = 0.0260$], effect sizes ($\eta^2 = 0.0024$ and $\eta^2 = 0.0046$ respectively). We also conducted one-way repeated measures ANOVAs to test for within-participants differences in M1 RMS mu between states. Differences in neither left M1 [$F(2, 24) = 0.7357$, $p = 0.4896$] nor right M1 [$F(2, 24) = 0.8488$, $p = 0.4404$] were statistically significant.

Post-hoc differences for within-participants differences in occipital alpha across states showed significant differences in alpha activity in the left primary visual cortex between states 1 and 3 [$t(12) = -1.44$, $p = 0.0374$, Cohen's $d = -0.1036$], as well as significant differences in the right primary visual cortex between states 1 and 2 [$t(12) = 2.70$, $p = 0.0194$, Cohen's $d = 0.0835$] and states 2 and 3 [$t(12) = -2.25$, $p = 0.0436$, Cohen's $d = -0.1585$]. These results suggest that, in addition to our states displaying distinct patterns of alpha activity, participants manifest the highest levels of alpha in the panic state, followed by the default state.

4. Discussion

Drawing on recent advances in behavioural neuroscience, we used an HMM to identify hidden states in multivariate psychomotor data obtained from an ecologically valid task, showing that humans shift between latent states during psychomotor performance that differ in behavioural and neural characteristics. Our task was a laboratory version of Tetris that logs granular performance metrics through time, and was performed in an MEG scanner. We identified three distinct states: (1) a “default” state with efficient motor execution and variable pile structure, (2) a transient “opportunity” state characterised by high scoring potential and rapid decision-making, and (3) a “panic” state involving challenging pile structures and inconsistent performance. Comparisons of neural activity between our three states revealed statistically significant differences in amplitudes of occipital alpha-band activity, a signal associated with attentional state (Foster and Awh,

2019; Peylo et al., 2021), indicating that differences in cognition across states may relate to attention. Taken together, our findings show that humans switch between behaviourally and neurally distinct states as they engage in complex psychomotor performance. We show that the dynamics of these state transitions can be captured using synchronised behavioural and neural measurements, and modelled using unsupervised learning techniques to describe the relationship between states and performance.

Previous latent state models of behaviour have concentrated on animal behaviour in relatively well-studied task environments, such as courtship behaviours in fruit flies (Calhoun et al., 2019), visual detection in mice (Chen, 2015; Roy et al., 2021; Ashwood et al., 2022), or swim bouts in larval zebrafish (Sharma et al., 2018). These paradigms lend themselves well to models of latent states as the resultant observations are intuitively discretisable. Additionally, many of these studies are high in ecological validity, modelling behaviours that would be natural to observe in an animal's usual behavioural repertoire. In comparison, the application of sequence classification techniques to identify latent states in humans has predominantly involved artificial tasks (e.g., motion coherence task (Ashwood et al., 2022)) or resting-state fMRI (Vidaurre et al., 2018, 2019; Karapanagiotidis et al., 2020). Here, we used a laboratory adaptation of a highly popular and commercially successful video game, paralleling growing interest in the use of naturalistic stimuli within the domain of cognitive neuroscience (Sonkusare et al., 2019; Reggente et al., 2018). Specifically, participants played a laboratory adaptation of Tetris (Lindstedt and Gray, 2013, 2015) that collects numerous cognitive-behavioural variables relating to game state, motor execution, and motor planning. Our analysis included a feature engineering component whereby behavioural measurements were decomposed into four performance components based on data obtained by an independent laboratory using the same task (Lindstedt and Gray, 2019). Thus, using a tried and tested version of a video game explicitly tailored for laboratory research, we add to a growing body of literature that uses video games for research in cognitive neuroscience (Voss et al., 2012; Bavelier et al., 2012; Boot, 2015; Zhang et al., 2015).

One related implication for our analysis is that phases of gameplay that are more demanding for less skilled players may place lower demands on best players in the sample. Having concatenated all observations to produce our input time series for model fitting, our model would not have accounted for the potential effects of variation in skill. This is an important consideration, given previous evidence highlighting that variables discriminating between less versus more skilled players are not the same across skill brackets (Thompson et al., 2013). In parallel research on animal models, this issue is either resolved through extensive training, or it is completely bypassed by observing naturally ingrained behaviours. Here, we made efforts to recruit participants who reported familiarity with Tetris, but we were unable to control for how proficient they were. Additionally, we realised during data collection that many participants were familiar with modern versions of Tetris with nuanced differences that confounded their initial experiences for the game. For instance, our configuration of Meta-T emulates Classic Tetris and therefore lacks visual guidelines indicating each tetromino's destination, and prohibits rotating tetrominoes at the very edge of the well; both of these are mechanics that some of our more experienced participants reported relying on in their usual recreational gameplay. These issues may have introduced some additional noise to our model.

Compared to previous latent state models of low-level psychophysical phenomena, we opted for a complex behavioural environment that is high in ecological validity. In doing so, we show that hidden Markov modelling can be used to identify state shifts in tasks that approach real world behaviour. However, we acknowledge our position in the trade-off between simple behavioural data suitable for predictive modelling versus rich behavioural data that makes prediction much more difficult. Given the nature of our input data (i.e., our time series of performance

components), our model infers parameters that describe the temporal dynamics of our states, and generates emission probabilities describing the probabilities of observations given the state time series. In the case of our Gaussian HMM, the emission probability parameters of each state were the mean and standard deviation parameters describing the Gaussian probability density function of each performance component in the respective state. This can lead to expectations about how participants may perform across states based on how the distributions of performance components shift across states but it does not permit the more fine-grained prediction mechanisms that other studies have used, such as the GLM-HMM approach in Calhoun et al. (2019).

The validity of our model is supported by the correspondence between the behavioural characterisations of our states and the underlying neural signatures of each state. In a model that failed to distinguish between cognitively meaningful states, we would expect to observe no differences in neural signatures associated with cognition. Instead, comparisons of neural activity across our inferred states revealed statistically significant differences in occipital alpha, a signal that has been previously linked to attention (Foster et al., 2017; Foster and Awh, 2019; Peylo et al., 2021). In particular, post-hoc tests revealed elevated occipital alpha in the panic state as compared to the default state, and higher occipital alpha in the default state as compared to the opportunity state.

These findings align with previous research on the relationship between occipital alpha power and attentional processes. The elevated alpha power observed in the “panic” state compared to the default state is consistent with studies showing increased alpha activity during periods of high cognitive load or task difficulty. For instance, Klimesch and colleagues (Klimesch et al., 1999) found that alpha power increased during the retention period of a working memory task, particularly when task demands were high. In contrast, the opportunity state might represent moments when players have identified clear strategic moves, requiring less broad attentional deployment and more focused execution.

State-dependent variations in occipital alpha power may relate to other neural signatures, such as activity in the default-mode network (DMN), which is known to be active during internally directed cognition (Raichle, 2015). Mo and colleagues (Mo et al., 2013) found a positive correlation between visual alpha power and DMN activity during eyes-open resting state, suggesting that increased alpha power may serve to protect internal information processing by gating out sensory input. This framework could explain our observation of elevated alpha in the panic state, where players focus internally on strategic planning while inhibiting distracting visual input.

Furthermore, the state transitions we observed in Tetris gameplay may be analogous to the spontaneous switching between externally-oriented and internally-oriented states. It is important to note that the relationship between alpha power and attention is complex and can be task-dependent. Foxe and Snyder (2011) reviewed evidence suggesting that alpha oscillations can serve as both an attentional suppression mechanism and an active processing mode, depending on the specific cognitive demands. Our results, showing distinct alpha patterns across the three identified states, underscore the nuanced nature of attentional modulation during complex, ecologically valid tasks like Tetris.

While our model likely reflects attention shifts, without manipulating attention as in previous studies (Cohen and Maunsell, 2009; Mitchell et al., 2009), we cannot precisely determine which task aspects are attended to in different states. To address this limitation, we performed a supplementary analysis estimating blink events across states using ICA, revealing significant differences in participants' blink rates between states. This finding aligns with recent research suggesting that blinking may be a source of attentional shifts (Yang et al., 2024). Future work could combine eye tracking with this task to allow more precise inferences about attention across states, potentially exploring the relationship between blinking and attentional shifts as a promising avenue for investigation.

It is also possible that attention shifts continuously, and not discretely as assumed by our model. Ashwood and colleagues (2022) found superior model fit in their discrete model as compared to a model with continuous latent states (Roy et al., 2021), albeit in the context of a different task. Additionally, these authors found that a two-state discrete model fit human data from a motion coherence task better than a three-state model. We are open to the possibility that models with different assumptions may describe performance in the present context better than our three-state gaussian HMM, for instance a two-state, engaged versus disengaged model. However, this is a question for future work.

4.1. Limitations

Many studies of cognition that use video games, in particular commercial video games, have analysed univariate measures of performance such as end of game summary metrics (e.g., win/loss, points scored), or time-bound measures of performance (e.g., points scored per minute). We show here that the analysis of fine-grained, multivariate behavioural time series can generate inferences and research questions that may be difficult to access with summary metrics alone. Relatedly, and partly as a consequence of this limitation, studies of video games that involve repeated measures often aggregate data within and across sessions of engagement. Previous work has advised against this on theoretical (Towne et al., 2014; Gobet, 2017; Stafford and Vaci, 2022) as well as empirical grounds, demonstrating how certain insights into individual differences (Harwell et al., 2018) or skill acquisition (Towne et al., 2016; Rahman and Gray, 2020) can only be achieved after disaggregating data and considering behaviour in a more detailed fashion. Although this study involved detailed analysis of behaviour through time, we have aggregated data across subjects despite variation in players' average scores indicating a heterogeneity in skill level. Additionally, follow-up studies should be conducted with a prior power analysis and a sample size sufficient for achieving adequate statistical power.

In addition, and in contrast to previous latent state models of behaviour, we did not train our participants on the task. Tetris is a complex psychomotor task requiring both rapid perceptual decision-making and skilled motor inputs. Although we recruited participants who all indicated ample prior experience with Tetris, as mentioned before, we are nonetheless conscious of large variation in participant skill, as well as noise arising from unfamiliarity with our specific configuration of Meta-T, the controller, and the scanner environment in which the task was performed. In addition, we note the absence of a “ground truth” model with which to validate our model. Instead, we compared the log-likelihood of our model to a randomised chance model, which indicated superior fit of the true model. However, we acknowledge as a limitation that due to the nature of our input data and the type of HMM that we used, the predictive capacity of our model may be restricted. Future work, perhaps using recording systems that are more tolerant to continuous recording in more naturalistic environments (for example, EEG, OPM-MEG or chronically-implanted electrode arrays) would allow us to collect more extensive data sets on a larger population and integrate other physiologically-relevant measures such as pupilometry, actigraphy, skin conductance and heart and respiration monitoring. Such dense, multivariate recordings could allow us to characterise the neurophysiological biomarkers associated with different states and make predictions of future performance.

4.2. Conclusions

Using simultaneous behavioural and neural recordings of participants playing a laboratory version of Tetris, we extend previous work by demonstrating that individuals switch between latent states during performance in an ecologically valid task. Individuals in our sample

shifted between three states each with unique performance characteristics during gameplay. Further, MEG analysis revealed differences in occipital alpha across states, suggesting that differences across states may be related to attention. Our results show that analysing sessions of data by averaging summary statistics alone may mask a wealth of information describing the dynamics of performance and cognition. We demonstrate how these dynamics can be uncovered using unsupervised learning techniques and granular, multivariate data.

CRediT authorship contribution statement

Ozan Vardal: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Theodoros Karapanagiotidis:** Writing – review & editing, Methodology, Conceptualization. **Tom Stafford:** Supervision. **Anders Drachen:** Writing – review & editing, Supervision, Funding acquisition. **Alex Wade:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the EPSRC Doctoral Training Programme for Intelligent Games and Games Intelligence (IGGI) EP/S022325/1 for funding this work. We thank John Lindstedt and Wayne Gray for providing us with the source code for the original Meta-T task, as well as an archival data set that launched our preliminary analyses. We also thank Richard Aveyard and Rebecca Lowndes for assisting with data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2025.121134>.

Data availability

All code and data including anonymised T1 anatomical data are archived at the York Neuroimaging Centre. They will be made available upon publication on OpenNeuro (<https://openneuro.org>).

References

- Anderson, J.R., Betts, S., Bothell, D., Hope, R., Lebiere, C., 2019. Learning rapid and precise skills. *Psychol Rev* 126 (5), 727–760.
- Anderson, J.R., Bothell, D., Fincham, J.M., Anderson, A.R., Poole, B., Qin, Y., 2011. Brain regions engaged by part- and whole-task performance in a video game: a model-based test of the decomposition hypothesis. *J. Cogn. Neurosci.* 23 (12), 3983–3997.
- Anon, 2022. hmmlern. URL <https://hmmlern.readthedocs.io/en/latest/>.
- Ashwood, Z.C., Roy, N.A., Stone, I.R., International Brain Laboratory, Urai, A.E., Churchland, A.K., Pouget, A., Pillow, J.W., 2022. Mice alternate between discrete strategies during perceptual decision-making. *Nat. Neurosci.* 25 (2), 201–212.
- Bavelier, D., Green, C., Pouget, A., Schrater, P., 2012. Brain plasticity through the life span: Learning to learn and action video games. *Annu. Rev. Neurosci.* 35, 391–416.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7 (6), 1129–1159.
- Boot, W., 2015. Video games as tools to achieve insight into cognitive processes. *Front. Psychol.* 6, 3.
- Boot, W.R., Basak, C., Erickson, K.I., Neider, M.B., Simons, D.J., Fabiani, M., Gratton, G., Voss, M.W., Prakash, R.S., Lee, H., Low, K.A., Kramer, A.F., 2010. Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychol.* 135 3, 349–357.

- Calhoun, A., Pillow, J., Murthy, M., 2019. Unsupervised identification of the internal states that shape natural behavior. *Nature Neurosci.* 22, 1–10.
- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* 2 (3), 1–27.
- Chen, Z., 2015. Estimating latent attentional states based on simultaneous binary and continuous behavioral measures. *Comput. Intell. Neurosci.* 2015, 493769.
- Cohen, M., Maunsell, J., 2009. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neurosci.* 12, 1594–1600.
- Dale, A., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 9 (2), 179–194.
- Eyjoisdottir, E., Branson, K., Yue, Y., Perona, P., 2017. Learning recurrent representations for hierarchical behavior modeling. In: *International Conference on Learning Representations*.
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. *Cerebral Cortex* 14 (1), 11–22.
- Foster, J.J., Awh, E., 2019. The role of alpha oscillations in spatial attention: limited evidence for a suppression account. *Curr. Opin. Psychol.* 29, 34–40.
- Foster, J.J., Sutterer, D.W., Serences, J.T., Vogel, E.K., Awh, E., 2017. Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention. *Psychol. Sci.* 28 (7), 929–941.
- Foxe, J., Snyder, A., 2011. The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front. Psychol.* 2.
- Gobet, F., 2017. Allen Newell's program of research: The video-game test. *Top. Cogn. Sci.* 9 (2), 522–532.
- Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32 (1), 35–42.
- Harwell, K.W., Boot, W.R., Ericsson, K.A., 2018. Looking behind the score: Skill structure explains sex differences in skilled video game performance. *PLOS ONE* 13 (5), 1–18.
- Karapanagiotidis, T., Vidaurre, D., Quinn, A., Vatansever, D., Poerio, G., Turnbull, A., Ho, N., Leech, R., Bernhardt, B., Jefferies, E., Margulies, D., Nichols, T., Woolrich, M., Smallwood, J., 2020. The psychological correlates of distinct neural states occurring during wakeful rest. *Sci. Rep.* 10 (1).
- Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., Winkler, T., 1999. 'Paradoxical' alpha synchronization in a memory task. *Cogn. Brain Res.* 7 (4), 493–501.
- Lee, H., Voss, M.W., Prakash, R.S., Boot, W.R., Vo, L.T.K., Basak, C., Vanpatter, M., Gratton, G., Fabiani, M., Kramer, A.F., 2012. Videogame training strategy-induced change in brain function during a complex visuomotor task. *Behav. Brain Res.* 232 (2), 348–357.
- Lindstedt, J.K., Gray, W.D., 2013. Extreme expertise: Exploring expert behavior in tetris. *Cogn. Sci.* 35.
- Lindstedt, J., Gray, W., 2015. Meta-T: Tetris as an experimental paradigm for cognitive skills research. *Behav. Res. Methods* 47.
- Lindstedt, J., Gray, W., 2019. Distinguishing experts from novices by the mind's hand and mind's eye. *Cogn. Psychol.* 109, 1–25.
- Mitchell, J., Sundberg, K., Reynolds, J., 2009. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63, 879–888.
- Mo, J., Liu, Y., Huang, H., Ding, M., 2013. Coupling between visual alpha oscillations and default mode activity. *NeuroImage* 68, 112–118.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Peylo, C., Hilla, Y., Sauseng, P., 2021. Cause or consequence? Alpha oscillations in visuospatial attention. *Trends Neurosci.* 44 (9), 705–713.
- Rahman, R., Gray, W., 2020. SpotLight on dynamics of individual learning. *Top. Cogn. Sci.* 12.
- Raichle, M.E., 2015. The brain's default mode network. *Annu. Rev. Neurosci.* 38, 433–447.
- Reggente, N., Essoe, J.K., Aghajan, Z.M., Tavakoli, A.V., McGuire, J.F., Suthana, N.A., Rissman, J., 2018. Enhancing the ecological validity of fMRI memory research using virtual reality. *Front. Neurosci.* 12 (JUN).
- Renart, A., Machens, C.K., 2014. Variability in neural activity and behavior. *Curr. Opin. Neurobiol.* 25, 211–220.
- Roy, N.A., Bak, J.H., Akrami, A., Brody, C.D., Pillow, J.W., 2021. Extracting the dynamics of behavior in sensory decision-making experiments. *Neuron* 109 (4), 597–610.e6.
- Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W.R., Birbaumer, N., 2005. A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur. J. Neurosci.* 22 (11), 2917–2926.
- Sharma, A., Johnson, R., Engert, F., Linderman, S., 2018. Point process latent variable models of larval zebrafish behavior. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (Eds.), *Advances in Neural Information Processing Systems*, vol. 31, Curran Associates, Inc..
- Shinners, P., 2011. Pygame. <http://pygame.org/>.
- Sibert, C., Gray, W., 2018. The tortoise and the hare: Understanding the influence of sequence length and variability on decision-making in skilled performance. *Comput. Brain & Behav.* 1.
- Sibert, C., Gray, W., Lindstedt, J., 2017. Interrogating feature learning models to discover insights into the development of human expertise in a real-time, dynamic decision-making task. *Top. Cogn. Sci.* 9.
- Sonkusare, S., Breakspear, M., Guo, C., 2019. Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends Cogn. Sci.* 23.
- Stafford, T., Haasnoot, E., 2017. Testing sleep consolidation in skill learning: a field study using an online game. *Top. Cogn. Sci.* 9 (2), 485–496.
- Stafford, T., Vaci, N., 2022. Maximizing the potential of digital games for understanding skill acquisition. *Curr. Dir. Psychol. Sci.* 31 (1), 49–55.
- Tadel, F., Bock, E., Niso, G., Mosher, J.C., Cousineau, M., Pantazis, D., Leahy, R.M., Baillet, S., 2019. MEG/EEG group analysis with brainstorm. *Front. Neurosci.* 13.
- Tenison, C., Anderson, J., 2016. Modeling the distinct phases of skill acquisition. *J. Exp. Psychol. [Learn. Mem. Cogn.]* 42, 749–767.
- Thompson, J.J., Blair, M.R., Chen, L., Henrey, A.J., 2013. Video game telemetry as a critical tool in the study of complex skill learning. *PLOS ONE* 8 (9), 1–12.
- Thompson, J., Blair, M., Henrey, A., 2014a. Over the hill at 24: Persistent age-related cognitive-motor decline in reaction times in an ecologically valid video game task begins in early adulthood. *PLoS ONE* 9 (4).
- Thompson, J.J., Blair, M.R., Henrey, A.J., 2014b. Over the hill at 24: Persistent age-related cognitive-motor decline in reaction times in an ecologically valid video game task begins in early adulthood. *PLOS ONE* 9 (4), 1–10.
- Thompson, J.J., McColeman, C.M., Blair, M.R., Henrey, A.J., 2019. Classic motor chunking theory fails to account for behavioural diversity and speed in a complex naturalistic task. *PLOS ONE* 14 (6), 1–24.
- Thompson, J.J., McColeman, C.M., Stepanova, E.R., Blair, M.R., 2017. Using video game telemetry data to research motor chunking, action latencies, and complex cognitive-motor skill learning. *Top. Cogn. Sci.* 9 (2), 467–484.
- Towne, T., Boot, W., Ericsson, K., 2016. Understanding the structure of skill through a detailed analysis of individuals' performance on the space fortress game. *Acta Psychol.* 169, 27–37.
- Towne, T., Ericsson, K., Sumner, A., 2014. Uncovering mechanisms in video game research: Suggestions from the expert-performance approach. *Front. Psychol.* 5, 161.
- van Maanen, L., Brown, S.D., Eichele, T., Wagenmakers, E.-J., Ho, T., Serences, J., Forstmann, B.U., 2011. Neural correlates of trial-to-trial fluctuations in response caution. *J. Neurosci.* 31 (48), 17488–17495.
- Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA.
- Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., 2018. Discovering dynamic brain networks from big data in rest and task. *NeuroImage* 180, 646–656.
- Vidaurre, D., Myers, N.E., Stokes, M., Nobre, A.C., Woolrich, M.W., 2019. Temporally unconstrained decoding reveals consistent but time-varying stages of stimulus processing. *Cerebral Cortex* 29 (2), 863–874.
- Visser, I., Raijmakers, M.E., Molenaar, P.C., 2002. Fitting hidden Markov models to psychological data. *Sci. Program.* 10, 185–199.
- Voss, M.W., Prakash, R.S., Erickson, K.I., Boot, W.R., Basak, C., Neider, M.B., Simons, D.J., Fabiani, M., Gratton, G., Kramer, A.F., 2012. Effects of training strategies implemented in a complex videogame on functional connectivity of attentional networks. *NeuroImage* 59 (1), 138–148.
- Wiltchko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abaira, V.E., Adams, R.P., Datta, S.R., 2015. Mapping sub-second structure in mouse behavior. *Neuron* 88 (6), 1121–1135.
- Yamagishi, N., Callan, D.E., Goda, N., Anderson, S.J., Yoshida, Y., Kawato, M., 2003. Attentional modulation of oscillatory activity in human visual cortex. *NeuroImage* 20 (1), 98–113.
- Yang, B., Intoy, J., Rucci, M., 2024. Eye blinks as a visual processing stage. *Proc. Natl. Acad. Sci.* 121 (15), e2310291121.
- Zhang, Y., Du, G., Yang, Y., Qin, W., Li, X., Zhang, Q., 2015. Higher integrity of the motor and visual pathways in long-term video game players. *Front. Hum. Neurosci.* 9, 98.