

1           Linking the brain with behaviour: the neural dynamics of  
2           success and failure in goal-directed behaviour

3     Amanda K. Robinson<sup>1,2,3</sup>, Anina N. Rich<sup>1,2</sup>, Alexandra Woolgar<sup>1,2,4</sup>

4     <sup>1</sup>Department of Cognitive Science, Macquarie University

5     <sup>2</sup>Perception in Action Research Centre (PARC), Macquarie University

6     <sup>3</sup>School of Psychology, The University of Sydney

7     <sup>4</sup>MRC Cognition and Brain Sciences Unit, Cambridge University

8     Acknowledgements: We thank Christopher Whyte for assistance in data collection. This work was

9     funded by Australian Research Council Discovery Project 170101840, Australian Research Council

10    Future Fellowship FT170100105, Medical Research Council (UK) intramural funding

11    SUAG/052/G101400.

## Abstract

13 The human brain is extremely flexible and capable of rapidly selecting relevant information in  
14 accordance with task goals. Regions of frontoparietal cortex flexibly represent relevant task  
15 information such as task rules and stimulus features when participants perform tasks successfully,  
16 but less is known about how information processing breaks down when participants make  
17 mistakes. This is important for understanding whether and when information coding recorded  
18 with neuroimaging is directly meaningful for behaviour. Here, we used magnetoencephalography  
19 (MEG) to assess the temporal dynamics of information processing, and linked neural responses  
20 with goal-directed behaviour by analysing how they changed on behavioural error. Participants  
21 performed a difficult stimulus-response task using two stimulus-response mapping rules. We used  
22 time-resolved multivariate pattern analysis to characterise the progression of information coding  
23 from perceptual information about the stimulus, cue and rule coding, and finally, motor response.  
24 Response-aligned analyses revealed a ramping up of perceptual information prior to a correct  
25 response, suggestive of internal evidence accumulation. Strikingly, when participants made a  
26 stimulus-related error, and not when they made other types of errors, patterns of activity initially  
27 reflected the stimulus presented, but later reversed, and accumulated towards a representation of  
28 the *incorrect* stimulus. This suggests that the patterns recorded at later timepoints reflect an  
29 internally generated stimulus representation that was used to make the (incorrect) decision.  
30 These results illustrate the orderly and overlapping temporal dynamics of information coding in  
31 perceptual decision-making and show a clear link between neural patterns in the late stages of  
32 processing and behaviour.

## Introduction

34 A primary function of the human brain is to flexibly respond to relevant perceptual information in  
35 accordance with current context and task goals. The sound of a phone ringing, for example, should  
36 prompt a different response if it is your phone than if it belongs to someone else. This set of  
37 complex processes, termed cognitive control, involves interpreting incoming information given the  
38 current context to determine an appropriate action (Posner & Presti, 1987; Posner & Snyder,  
39 1975). Cognitive control involves dynamic information exchange at different levels of processing,  
40 from perceptual information processing to decision-making and response selection.  
41 Understanding how these different processes unfold could provide a great deal of insight into how  
42 the brain achieves goal-directed behaviour.

43 A large body of neuroimaging research implicates frontoparietal brain regions in goal-directed  
44 behaviour, which form a distributed network responsible for cognitive control (Duncan, 2010).  
45 This “multiple demand” (MD) network (Duncan, 2010), elsewhere referred to as the cognitive  
46 control network (Cole & Schneider 2007), frontoparietal control system (Vincent et al. 2008), or  
47 task-positive network (Fox et al. 2005), appears to flexibly represent different types of information  
48 depending on task context. For example, activity in these regions encodes task rules (e.g.,  
49 Crittenden et al., 2016; Woolgar et al., 2015, 2011), and auditory, visual, and tactile stimulus  
50 features (Bracci et al., 2017; Jackson et al., 2017; Long and Kuhl, 2018; Woolgar and Zopf, 2017;  
51 for a review see Woolgar et al., 2016). These regions particularly encode task elements that are  
52 demanding (Woolgar, Afshar, et al., 2015; Woolgar et al., 2011) or at the focus of attention (J.  
53 Jackson et al., 2017; J. B. Jackson & Woolgar, 2018; Woolgar, Williams, et al., 2015). Activity in  
54 some of these regions has also been causally implicated in selectively facilitating coding of task-  
55 relevant information (Jackson et al 2021). This lends support to the possibility that flexible

56 responses within the MD regions play a causal role in goal-directed behaviour (e.g., Duncan et al.,  
57 2020; Woolgar et al., 2019, 2018, 2010).

58 A characteristic feature of cognitive control is that it dynamically changes in response to task-  
59 relevant information. Research using fMRI has yielded insight into the brain networks involved in  
60 goal-directed behaviour, but the slow nature of the blood-oxygen-level response has limited the  
61 exploration of the corresponding dynamics. Time-resolved neuroimaging methods such as  
62 magnetoencephalography and electroencephalography (M/EEG) have been more fruitful in  
63 understanding how cognitive control unfolds over time (for review, see Gratton et al., 2017). For  
64 example, conflict-related processing involving incongruent task features elicits a larger evoked  
65 response than a congruent condition approximately 200-400ms after stimulus onset, which has  
66 been linked to activity within the anterior cingulate (Folstein & Van Petten, 2008), and task-  
67 switching involves a larger parietal positivity around 300ms relative to task repeats (Karayannidis et  
68 al., 2010). The newer method of multivariate pattern analysis (MVPA) in conjunction with M/EEG  
69 allows further insight into processing dynamics underlying cognitive control. MVPA uses pattern  
70 classification approaches applied to neuroimaging data to show what information is being coded  
71 within the brain (Haxby, 2012; Haxby et al., 2001). Time-resolved MVPA has been used to  
72 characterise how information coding changes over time (Carlson et al., 2011; Hebart & Baker,  
73 2018). For example, Hebart and colleagues (2018) had participants perform different tasks on  
74 visual object stimuli while measuring MEG, and showed that task-relevant object features were  
75 enhanced at late stages of processing, more than 500ms after the stimulus was presented. Other  
76 work has shown clear progression of task relevant information during complex tasks, with  
77 different dynamics for features such as stimulus, task and response (Hubbard et al., 2019;  
78 Kikumoto & Mayr, 2020; Wen et al., 2019). This line of research has also highlighted the  
79 importance of combining relevant task information for successful behaviour (Kikumoto & Mayr,

80 2020). These MVPA studies provided great insight into the neural dynamics of goal-directed  
81 behaviour, but all used designs where the task cue was presented prior to the target, allowing  
82 participants to prepare for the task in advance. Additionally, these studies focused on stimulus-  
83 aligned neural responses. It seems likely that tracking the dynamic coding of relevant task features  
84 relative to both stimulus onset and response, using a task that induces more flexible behaviour,  
85 might elucidate stronger links between dynamic neural responses and goal-directed behaviour.

86 Decades of neuroimaging research have focused on the neural correlates of behaviour, but  
87 assessing whether particular patterns of brain activity are *necessary* for behaviour has presented a  
88 challenge. In MVPA, a classifier algorithm is trained to distinguish between conditions using  
89 patterns of neural data from multiple trials of each condition. If a classifier can predict the  
90 conditions of new neural data better than chance, this demonstrates that the patterns of activity  
91 in the data must contain, or represent, information about the different conditions. However, the  
92 conclusion that decodable patterns represent “information” has been questioned on theoretical  
93 grounds (de-Wit et al., 2016): just because information is decodable using machine learning does  
94 not necessarily mean it is used by the brain to generate behaviour. This awareness has lead  
95 researchers to push for more explicit links between MVPA patterns and behaviour, for example,  
96 comparing details of patterns to reaction times or accumulation rates in models of behaviour  
97 (Grootswagers et al., 2018; Ritchie et al., 2015; Ritchie & Carlson, 2016).

98 Exploring how information coding changes when participants make errors is another way to  
99 establish how behaviourally meaningful patterns of activation are. For example, Williams et al  
100 (2007) demonstrated that multivariate fMRI patterns in lateral occipital cortex, but not those in  
101 early visual regions, reduced to chance when participants made errors on a shape discrimination  
102 task, indicating that patterns in early visual cortex were not directly read out in behaviour. In  
103 another study, participants performed a scene classification task (Walther et al., 2009), and

104 classifier prediction error patterns correlated with the types of errors in behaviour within high  
105 level object and scene-specific brain regions, but not within early visual cortex. Using  
106 magnetoencephalography (MEG), we have recently shown that this logic can even be used to  
107 predict behavioural errors before they occur (Karimi-Rouzbahani et al., 2021).

108 A stronger requirement for a behaviourally meaningful pattern of activity is that it should not just  
109 change on error, but change to something that predicts the particular error to be made (Woolgar  
110 et al., 2019). We tested this in fMRI, and found that patterns of activation in frontoparietal cortex  
111 indeed reversed on error, such that patterns of activation on error trials represented information  
112 that was not presented to the participant, in a manner that was diagnostic of the particular  
113 behavioural error they made (Woolgar et al., 2019). In that study, participants performed a  
114 difficult response-mapping task. When participants made a rule error, MD patterns of activity  
115 reflected the *incorrect* rule, and when participants made other errors, MD patterns of activity  
116 reflected the *incorrect* stimulus (Woolgar et al., 2019). Within visual cortex, in contrast, there was  
117 no evidence of relevant information (correct or incorrect) during errors. Thus, some multivariate  
118 patterns appear to be more directly relatable to behaviour than others, and there is a tight link  
119 between frontoparietal activity patterns and behavioural outcome.

120 In the current study, we used MEG and MVPA to examine the dynamics of this effect, asking  
121 whether information coding through the course of a trial was equally associated with behaviour.  
122 We aimed to (1) characterise the neural dynamics of multiple types of task relevant information,  
123 and (2) examine their relationship to behaviour over time. Participants performed a difficult  
124 response-mapping task which required different responses to a target stimulus depending on the  
125 current rule. To determine what aspects of this representation could be directly linked to  
126 behaviour, we examined information coding on *incorrect* trials: when the wrong rule was applied  
127 or when there were errors in perception. We found a clear progression in onset of information

128 coding, such that stimulus features are evident shortly after stimulus onset, followed by abstract  
129 rule coding and then the response, with the information about each task feature accumulating up  
130 to the time of response. When participants made stimulus errors, stimulus information was  
131 initially coded veridically but later accumulated in the opposite direction, towards a  
132 representation of the incorrect stimulus. By contrast, stimulus information was encoded correctly  
133 when participants made rule errors. The data reveal the dynamics with which information coding  
134 in the brain can be tightly linked to participant behaviour.

## 135 **Methods**

### 136 **Participants**

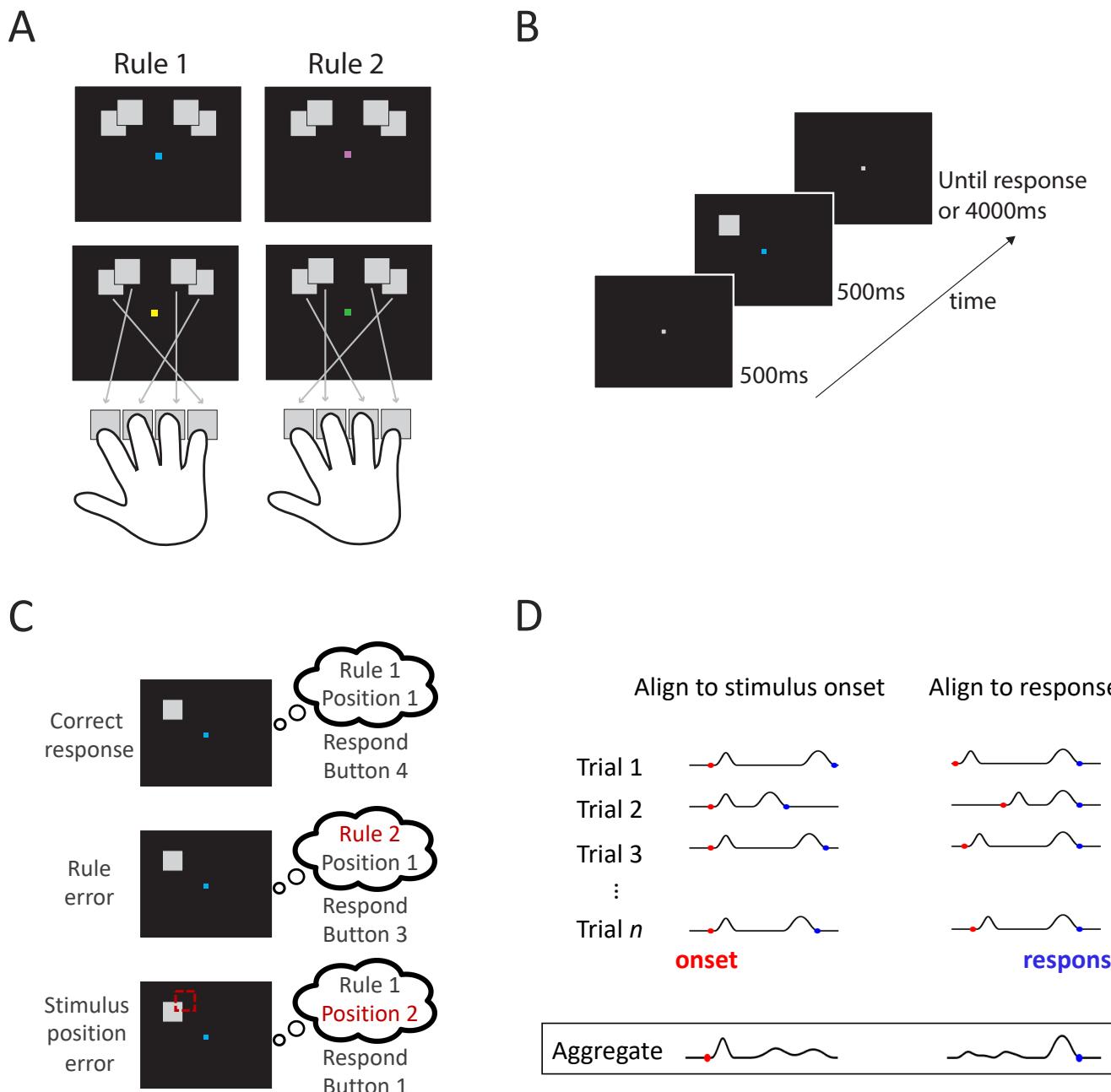
137 Participants were 22 healthy adults (14 females, age range 18-38 years) with normal or corrected-  
138 to-normal vision recruited from Macquarie University. This study was approved by the Macquarie  
139 University ethics committee and informed consent was obtained from all participants. Participants  
140 took part in two sessions: a 1-hour behavioural session, and a 2-hour MEG session, on separate  
141 days. They were compensated \$15 for the behavioural session and \$40 for the MEG session. For  
142 two participants, initial photodiode inaccuracies meant that the timing for 2 and 5 trials  
143 respectively was not adequately marked, so these trials were excluded from analyses. Data from  
144 an additional two participants (2 males) were collected and excluded: both participants had very  
145 few stimulus position errors during the MEG session (<10), and for one of the participants there  
146 was a recording error such that MEG data was recorded for only 680 out of 800 trials.

### 147 **Design and procedure**

148 Participants learned to apply two difficult response-mapping rules regarding the position of a  
149 target stimulus. The target was a grey square approximately 2x2 degrees of visual angle that

150 appeared in one of four positions. All positions were equidistant from fixation at an eccentricity of  
151 4 degrees of visual angle. Within the left and right visual fields, the two possible target locations  
152 overlapped by 60% horizontally and 65% vertically to create a high degree of position uncertainty  
153 (Figure 1a). Participants had to respond to the position of the stimulus using two possible  
154 response-mapping rules (Figure 1a). The two rules each comprised 4 unique position  
155 transformations and were mirror images of each other. The colour of a central fixation square  
156 acted as a cue for the rule. There were two cues per rule, in order to dissociate neural responses  
157 to cues from the neural responses to rules (e.g., blue and yellow = Rule 1, pink and green = Rule 2;  
158 counterbalanced across participants). Participants responded by pressing one of four response  
159 keys with their right hand. The stimuli were presented using PsychToolbox in MATLAB.

160 The stimulus-response rule mappings were designed to distinguish correct responses and specific  
161 types of errors (Figure 1c). An error was considered a *rule error* when the button press response  
162 reflected the combination of the correct stimulus position with the wrong rule. In contrast, a  
163 *stimulus error* was defined as a button press response consistent with the combination of the  
164 adjacent perceptually confusable position with the correct rule. For example, under Rule 1, if the  
165 stimulus appeared on the far left, the correct response would be button 4, a rule error (i.e., using  
166 Rule 2 applied to the correct position) would lead to a button 3 response, a stimulus error (i.e.,  
167 using Rule 1 correctly but confusing the stimulus with the other left position) would lead to a  
168 button 1 response, and confusing both the rule and the stimulus led to a button 2 response.



169

170 **Figure 1.** Experimental design. A) Response mapping rules. Participants had to indicate the position of a target  
 171 stimulus which appeared in one of four possible locations. There were two cues per rule, designated by blue/yellow  
 172 and green/pink squares at fixation. The button press associated with each position is indicated by the specific rule. B)  
 173 Trial timeline. After a fixation screen, the target stimulus and coloured fixation cue appeared simultaneously, and  
 174 participants had to apply the correct response-mapping rule using a button press. C) Behavioural response types. In  
 175 this example, the stimulus was in Position 1 and the rule was Rule 1 (blue cue), so the correct response was Button 4.  
 176 A rule error occurred if the rule was mistaken for Rule 2, leading to a response of button 3. A stimulus position error  
 177 occurred if the position was mistaken to be Position 2, leading to a response of button 1. D) Depiction of MEG data  
 178 collation: data aligned to stimulus onset (left) or response time (right). The temporal dynamics of stimulus-related and  
 179 decision-related neural responses vary across trials, with different processes aligned with onset and response. Aligning  
 180 the MEG data to stimulus onset versus response highlights different neural stages on the aggregate of all trials, even  
 181 though the content of each trial is identical.

182 *Training session*

183 Participants learned each rule in a separate session outside the MEG. They were trained to  
184 perform the task using increasingly difficult blocks of trials (see below). Feedback was given on  
185 every trial. For every incorrect response, participants were shown the correct response.

186 Initially, stimuli were presented in non-overlapping positions (i.e., further apart than the final  
187 paradigm) so there was no position uncertainty. Stimuli were presented on the screen until a  
188 response was made (i.e., not time-limited). In the first block, participants learnt the first rule (Rule  
189 1 or Rule 2, counterbalanced across participants). Each stimulus position was shown with its  
190 associated response four times (16 trials), and participants had to press the appropriate button for  
191 each stimulus. In each trial, cue colour was chosen randomly from the two possible cues for that  
192 rule. The second block followed the same protocol, but for the other rule (Rule 1/Rule 2). In the  
193 third block, participants had to perform the task by implementing both rules, but still with well  
194 separated stimuli. In the fourth block, the stimuli were presented in their final, overlapping  
195 experimental positions. Finally, in the fifth block, the stimuli were presented with the same  
196 procedure as the final experimental paradigm: the stimuli were overlapping and were presented  
197 for only 500ms. Blocks 3, 4 and 5 contained 32 trials each, consisting of 2 repeats of each cue and  
198 stimulus position, randomly ordered. In all blocks, participants had to perform at 60% accuracy or  
199 above to progress to the next block type. Blocks were repeated if they did not reach this  
200 threshold. On average, participants completed 8.61 training blocks ( $SD = 2.46$ ). Block 3 was most  
201 often repeated ( $M = 3.09$  repeats).

202 *Experimental session*

203 In the second session, participants performed the task while their neural activity was recorded  
204 using MEG. At the start of each block, participants were shown a graphical depiction of the rules

205 for at least 2 seconds. When they were ready, they pressed a button to begin the block. In each  
206 trial, participants were shown a grey fixation marker for 500ms, and then the square target  
207 stimulus and coloured rule cue were presented for 500ms (Figure 1b). The participants were  
208 instructed to respond as quickly as possible without sacrificing accuracy. After they responded,  
209 there was an inter-trial interval of 1000ms before the next trial started. There were 10 blocks of 80  
210 trials, each containing 5 trials per stimulus and cue colour combination. Within each block, the  
211 order of the trials was randomised. Instead of feedback on every trial, like in training, participants  
212 were given feedback about their mean accuracy and reaction times at the end of each block.

### 213 **MEG acquisition**

214 MEG data were collected at Macquarie University in the KIT-MQ MEG facility with a whole-head  
215 supine Yokogawa system containing 160 gradiometers (Kado et al., 1999). The participant's head  
216 was fitted with a cap containing five marker coils. The head shape and position of the marker coils  
217 was marked using a Polhemus digitisation system. Once inside the MEG, the position of the  
218 marker coils was measured to ensure the MEG sensors had good coverage over the participant's  
219 head. Marker position measurements were repeated halfway through the experiment and at the  
220 end of the session. Raw MEG data were collected at 1000 Hz with online 0.03 Hz highpass and 200  
221 Hz lowpass filters.

222 Stimuli were projected onto the ceiling of the magnetically shielded room. Stimulus timing was  
223 measured using a photodiode placed on the projection mirror and marked in an additional  
224 channel in the MEG recording. Participants indicated their response using a 4-Button Fiber Optic  
225 Response Pad (Current Designs, Philadelphia, USA). Response timing was marked in the MEG  
226 recording using a parallel port trigger.

227 **MEG data analysis**

228 MEG data were analysed using multivariate decoding, which is very sensitive to reliable effects in  
229 the data and resistant to artefacts such as eye blinks that are not consistent across time and  
230 condition (Carlson et al., 2020; Grootswagers et al., 2017). Due to the robustness of decoding to  
231 such artefacts, data were minimally preprocessed using EEGLAB (Delorme & Makeig, 2004). Data  
232 were filtered using a Hamming window FIR filter (default EEGLAB filter pop\_eegfiltnew) with  
233 highpass of 0.1Hz and lowpass of 100Hz, and then downsampled to 200 Hz before epoching. For  
234 separate analyses, trials were epoched relative to stimulus onset, marked by the photodiode, and  
235 to the button press response, marked by the parallel port trigger.

236 Data were analysed using time-resolved classification methods (e.g., Carlson et al., 2020;  
237 Grootswagers et al., 2017) and implemented using the CoSMoMVPA toolbox (Oosterhof et al.,  
238 2016). For each time point, data were pooled across all 160 MEG sensors, and we tested the ability  
239 of a linear discriminant analysis (LDA) classifier to discriminate between the patterns of neural  
240 responses associated with the different conditions. Trials were divided according to their  
241 associated behavioural responses: correct trials, rule errors and stimulus position errors (Figure  
242 1C). The classifiers were always trained on correct trials. To ensure that there were equal numbers  
243 of trials for each condition, correct trials were subsampled to be equal for each position and rule  
244 combination for each block per participant. To ensure adequate trial numbers for each of the  
245 analyses, blocks with fewer than two trials per rule x position combination were excluded; this  
246 amounted to 9 excluded blocks in total across 8 participants, with the remaining 14 participants  
247 having all blocks included. The total number of selected trials per participant was  $M = 437.09$  (min  
248 = 280, max = 600).

249 *Temporal dynamics of stimulus, cue, rule and response coding*

250 We performed pattern classification analyses to determine the time points at which stimulus  
251 position, cue, rule and response representations emerge in the brain. First, we decoded stimulus  
252 position by comparing neural representations of the inner two stimulus positions (Positions 2 and  
253 3) to those of the outer two stimulus positions (Positions 1 and 4). Separating position in this  
254 manner meant that motor responses and rule were balanced across the two position conditions  
255 and could not drive the classification results, ensuring we are detecting information related to  
256 stimulus position.

257 Next, we assessed the time course of rule coding by training a classifier to distinguish between  
258 Rule 1 and Rule 2. In having two colour cues per rule, this analysis focused on rule coding over and  
259 above the physical properties of the cues (Rule 1 (blue and yellow cues) versus Rule 2 (pink and  
260 green cues)).

261 We can also, however, decode cue coding separately from rule coding. To assess how cue  
262 decoding differed from rule decoding, we decoded between the two cues per rule (i.e., blue  
263 versus yellow colour cue for Rule 1, and pink versus green for Rule 2). Cue coding was quantified  
264 as the mean of the two pairwise analyses.

265 As a final analysis, we decoded motor response by comparing the inner two button presses to the  
266 outer two button presses. This comparison ensured that stimulus position and rule were balanced  
267 within each class, so that the classifier would be driven by the motor response alone.

268 For each decoding analysis, classification analyses were performed using a leave-one-block out  
269 cross-validation approach. This resulted in 10-fold cross-validation for participants with no  
270 excluded blocks ( $N = 15$ ). The remaining participants used 9-fold ( $N = 7$ ), 8-fold ( $N = 1$ ) and 7-fold  
271 cross-validation ( $N = 1$ ). For all decoding analyses, chance performance was 50%.

272 *Error representations*

273 The next set of analyses focused on decoding neural activity when participants made errors, to  
274 explore the relationship between patterns of activity and behaviour. To investigate the  
275 representation of rule and stimulus errors, we trained the classifier on the correct trials and tested  
276 on the error trials. This allowed us to decode what information was present in the patterns of  
277 response across sensors when participants made different kinds of mistakes. Specifically, the  
278 analyses assessed whether the error patterns resembled the *correct* stimulus and rule patterns  
279 (above chance decoding), or the neural patterns associated with the *incorrect* stimulus and rule  
280 (below chance decoding). Note that in this approach, **below chance classification is meaningful**: it  
281 indicates the representation of the pattern that is instantiated when the other (incorrect in this  
282 case) rule or stimulus position is encoded.

283 We performed error decoding for stimulus position and rule information. In a comparable  
284 procedure to the correct trial analysis, we used leave-one-block out cross-decoding analyses. In  
285 each fold, the classifier was trained on *correct* trials from all but one block, and tested on all *error*  
286 trials across the whole session. This ensured the same training data was used as in the correct trial  
287 analyses, but allowed well-characterised results for the relatively small number of error trials.  
288 Participants made an average of 5.71% rule errors (number of trials:  $M = 45.68$ ,  $SD = 23.92$ , min =  
289 10, max = 99) and 9.82% stimulus position errors (number of trials:  $M = 78.50$ ,  $SD = 35.64$ , min =  
290 26, max = 160; Figure 2A).

291

## 292 Statistical testing

293 To assess performance of the classifier, we used Bayesian statistics to determine the evidence that  
294 decoding performance was different from chance (Dienes, 2011, 2016; Jeffreys, 1961; Rouder et

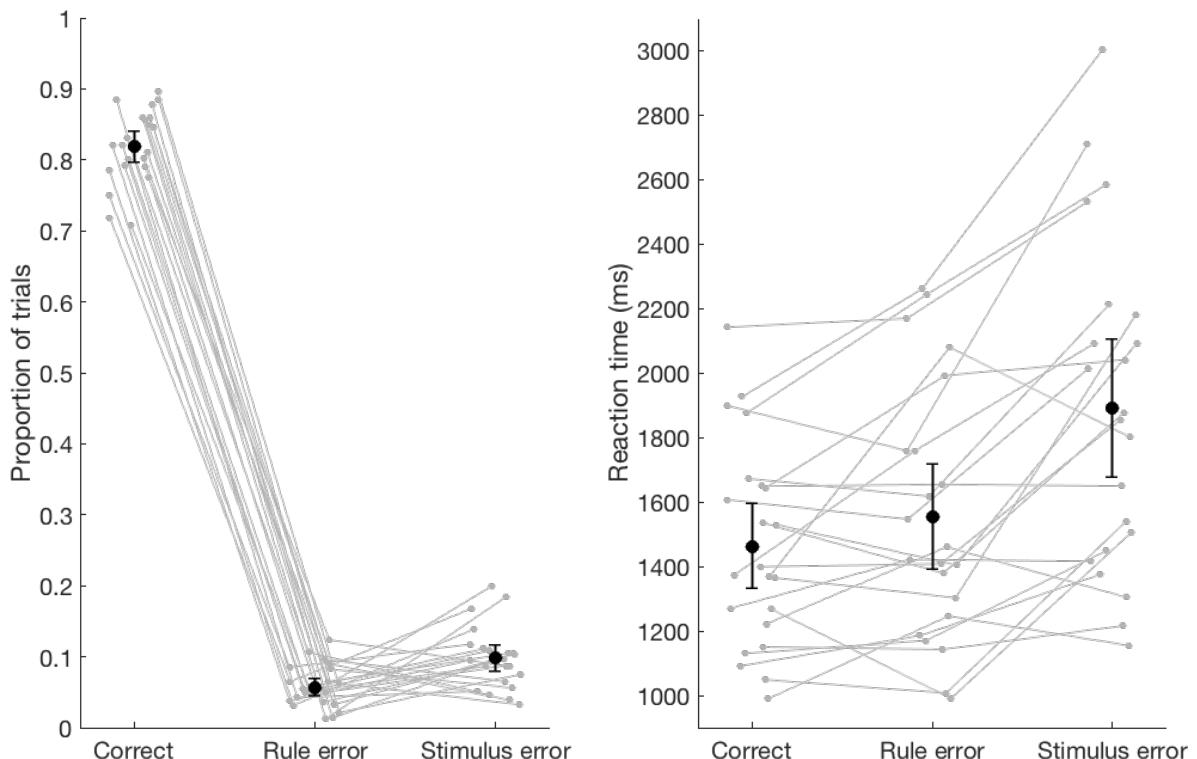
295 al., 2009; Wagenmakers, 2007). In all analyses, the alternative hypotheses of above- and below-  
296 chance (50%) decoding were tested using the ‘BayesFactor’ package in R (Morey et al., 2018).  
297 Bayes Factors were calculated using a JZS prior, centred around chance decoding of 50% (Rouder  
298 et al., 2009) with default scale factor of 0.707, meaning that for the alternative hypotheses of  
299 above- and below- chance decoding, we expected to see 50% of parameter values falling within -  
300 .707 and .707 standard deviations from chance (Jeffreys, 1961; Rouder et al., 2009; Wetzels &  
301 Wagenmakers, 2012; Zellner & Siow, 1980). A null interval was specified as a range of effect sizes  
302 between -0.5 to 0.5. A Bayes factor (BF) is the probability of the data under the alternative  
303 hypothesis relative to the null hypothesis. We consider  $BF > 3$  as evidence for the above- or below-  
304 chance decoding, and we refer to periods of time with sustained evidence (two consecutive time  
305 points of  $BF > 3$ ) as times when information “could be decoded”, indicating information was  
306 represented in the brain, and we refer to the first period of sustained evidence as the “onset” of  
307 decoding. We interpreted  $BF < 1/3$  as evidence in favour of the null hypothesis (Jeffreys, 1961;  
308 Wetzels et al., 2011).

## 309 Results

### 310 Behavioural results

311 All participants performed above 60% on the final block of the response-mapping task in the  
312 training session and therefore participated in the experimental MEG session. In the MEG session,  
313 participants performed well above chance ( $M = 81.89\%$ ) but still made both rule errors and  
314 stimulus position errors (Figure 2a). Reaction times were slowest for stimulus error trials (Figure  
315 2b).

A



316

317 Figure 2. Behavioural results in the MEG session ( $N = 22$ ). A) Proportion of trials, and B) Median reaction time for  
318 correct, rule errors and stimulus errors. Grey lines denote individual participants and black markers denote group  
319 means. Error bars are 95% confidence intervals across participants.

### 320 Temporal dynamics of goal-directed behaviour

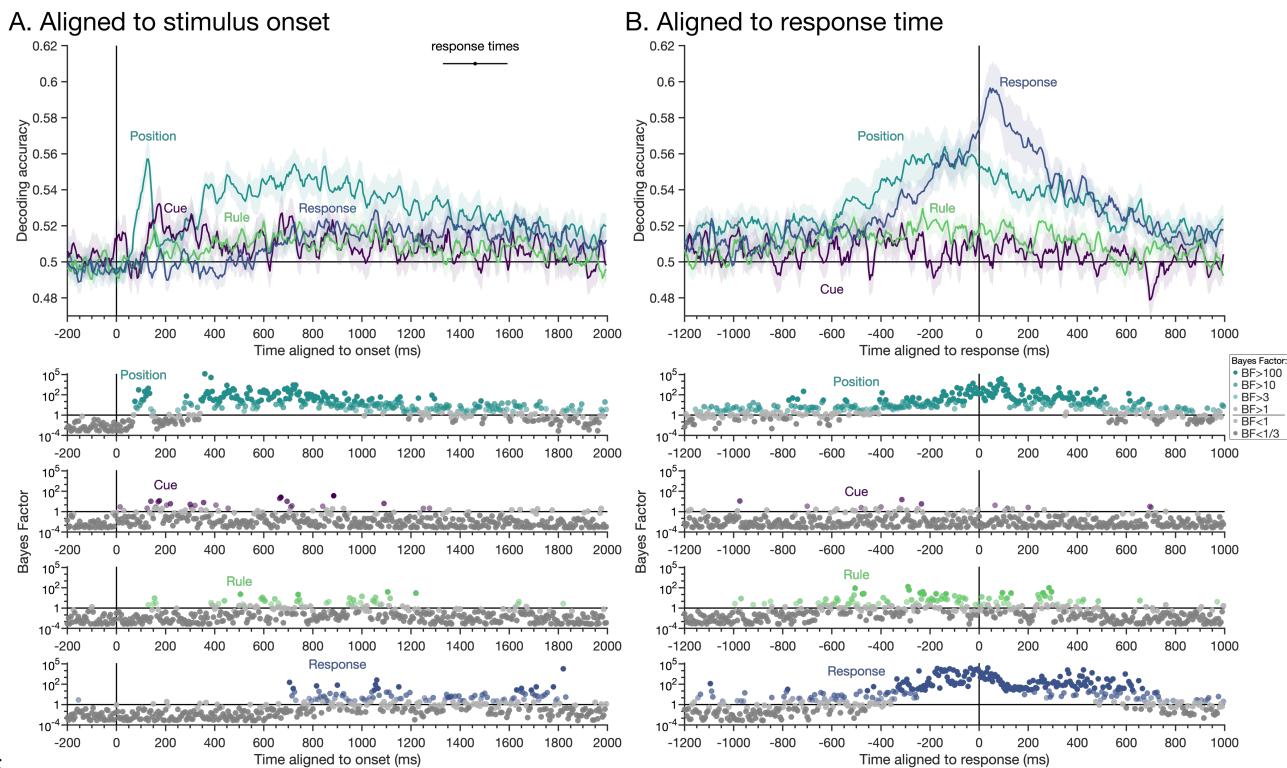
321 First, we investigated neural coding during correct trials by decoding different task-related  
322 information from the MEG signal when each trial was aligned to stimulus onset (analogous to  
323 classic event-related analyses). We then realigned the MEG signal of each trial to the response and  
324 performed the same decoding analyses (see Figure 1D for depiction of realignment). This gives us  
325 a unique insight into the time course of the processing stages during goal-directed behaviour.

326 Time-resolved decoding performed relative to stimulus onset revealed a progression of relevant  
327 information over time (Figure 3a). Stimulus position information was represented in the neural  
328 signal from approximately 75ms after the stimulus appeared, with a double-peak response. Cue  
329 information could be decoded from 170ms. These timings are consistent with early visual stages of

330 retinotopic position (Battistoni et al., 2020; Carlson et al., 2011) and colour processing (Teichmann  
331 et al., 2019), respectively. Rule information was briefly evident at about the same time as the cues  
332 (around 150ms), but also prominent from approximately 400 to 1200ms, coinciding with higher-  
333 level cognitive stages of processing. Finally, the button press response was represented from  
334 about 715ms. Decoding accuracy peaked at 125ms for stimulus position, 175ms for cue colour,  
335 600ms for rule and 1055ms for response button, showing a clear progression in information  
336 processing.

337 As is typical in difficult tasks, there was a wide variation in response times across trials and  
338 participants, indicating that the dynamics of high-level task-related processes such as decision-  
339 making vary trial to trial with respect to stimulus onset. Time-resolved decoding relies on  
340 processes occurring at the same time across trials, so this temporal jitter can mask results  
341 (Vidaurre et al., 2018). In order to capture processes that are more closely aligned with the  
342 response, we next realigned the MEG data to the response time (Figure 1D) and performed the  
343 same decoding analyses. The temporal dynamics of relevant task-related information were  
344 markedly different compared with onset-aligned results (Figure 3B). Notably, cue information  
345 could no longer be reliably detected, presumably because cue colour representations were  
346 transient and tightly stimulus-locked since, by design, the cue-distinctions were irrelevant as soon  
347 as rule information could be extracted from them. In contrast, decoding of stimulus position  
348 coding was evident more than 1000ms before the response, and rule coding was evident more  
349 than 600ms before the response, with evidence for stimulus processing earlier than rule  
350 processing. Motor response coding was sporadically present from more than 1000ms before the  
351 response, but was sustained from around 485ms prior to when the response was made. Response  
352 coding peaked after the response was given, potentially reflecting the contribution of  
353 somatosensory feedback from the different button presses. Interestingly, the representation of

354 stimulus position and response information appeared to ramp up before the response, plausibly  
355 reflecting the accumulation of evidence leading to a decision.



356

357 Figure 3. The temporal dynamics of correct stimulus position, cue colour, rule and response information coding. A)  
358 Decoding analyses conducted relative to stimulus onset. B) Decoding analyses conducted relative to response time.  
359 Shaded areas show standard error across participants ( $N = 22$ ). Decoding accuracy is smoothed with a 20ms window  
360 for visualisation. Line at top of plot A marks the mean and 95% confidence interval of median response times per  
361 participant. Bayes factors (BF) for above-chance decoding are displayed below the x-axes for every time point using a  
362 log scale and colour coded according to the evidence for above chance decoding (see inset).

### 363 Error representations

364 We were particularly interested in understanding whether and how the task-related information  
365 we can decode with MVPA is related to participant performance. Specifically, we investigated how  
366 information coding changes when an error is made. Recall that our design explicitly allows us to  
367 identify the likely source of the error based on the behavioural response (Fig. 1C). We assessed  
368 stimulus and rule information in the neural signal when participants made stimulus errors versus

369 rule errors. Based on our previous work with fMRI (Woolgar et al., 2019), we hypothesised that  
370 the brain would represent the incorrect stimulus prior to a stimulus error, and the incorrect rule  
371 prior to a rule error. Classifiers were trained to classify stimulus position and rule using correct  
372 trials and tested on incorrect trials. Therefore, for each time point on each error trial, the analysis  
373 reveals whether activation patterns were more similar to the usual patterns for the presented rule  
374 and stimulus (“correct” rule and stimulus) or the alternate one (“incorrect”) corresponding to the  
375 participant’s decision (as shown by the behaviour response).

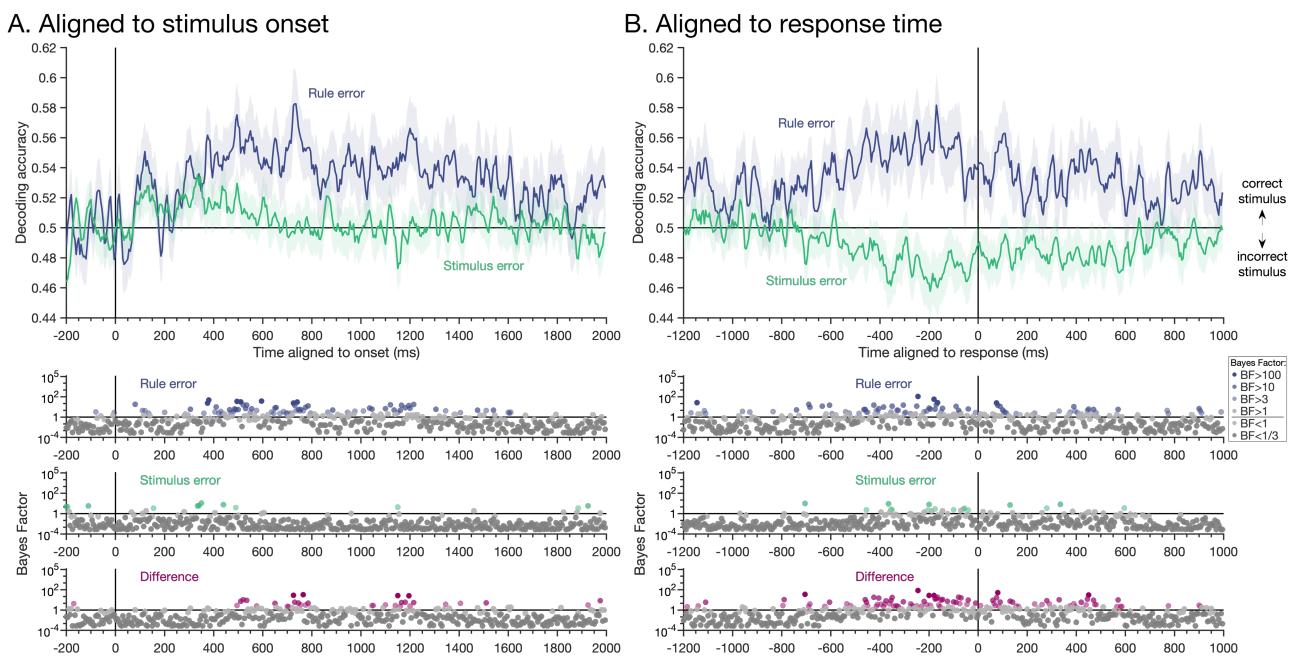
376 *Stimulus decoding – aligned to stimulus onset*

377 In this analysis, we looked at how stimulus position was coded on error trials. We found that when  
378 participants made rule errors, in which behaviour suggested that the stimulus was encoded  
379 correctly but the incorrect rule had been used, there was sustained stimulus decoding with similar  
380 dynamics to that on correct trials (Figure 4A; blue line). Stimulus coding on stimulus errors,  
381 however, was present only transiently at 335 ms after which coding attenuated (Figure 4A; green  
382 line). After 495 ms, there was substantial evidence that stimulus coding was higher when  
383 participants made errors based on applying the wrong rule than when the response suggested  
384 they had misperceived the stimulus. This indicates that when participants made stimulus errors,  
385 correct stimulus information was lost.

386 *Stimulus decoding - aligned to response time*

387 Next, we asked the same question but with the data re-aligned to the response time. On rule  
388 errors, there was a gradual ramping up of stimulus coding in the lead up to the response (Figure  
389 4B; blue line), as we had observed for correct trials (Figure 3B). In contrast, on stimulus errors,  
390 activation patterns ramped towards the patterns encoding the *incorrect* stimulus, as indexed by  
391 below-chance decoding from approximately 355ms before the response (Figure 4B; green line).  
392 Given that the correct stimulus had been encoded in the early part of these same trials (Figure

393 4A), this suggests an evolution of information coding towards the incorrect stimulus decision.  
 394 Stimulus decoding accuracy on rule errors was higher than that on stimulus errors for the bulk of  
 395 the epoch, particularly from about 795ms before the response to 600ms after the response.  
 396 Together, this finding shows that stimulus coding in the latter part of the trial reflected the  
 397 decision ultimately made by participants, rather than the stimulus presented.

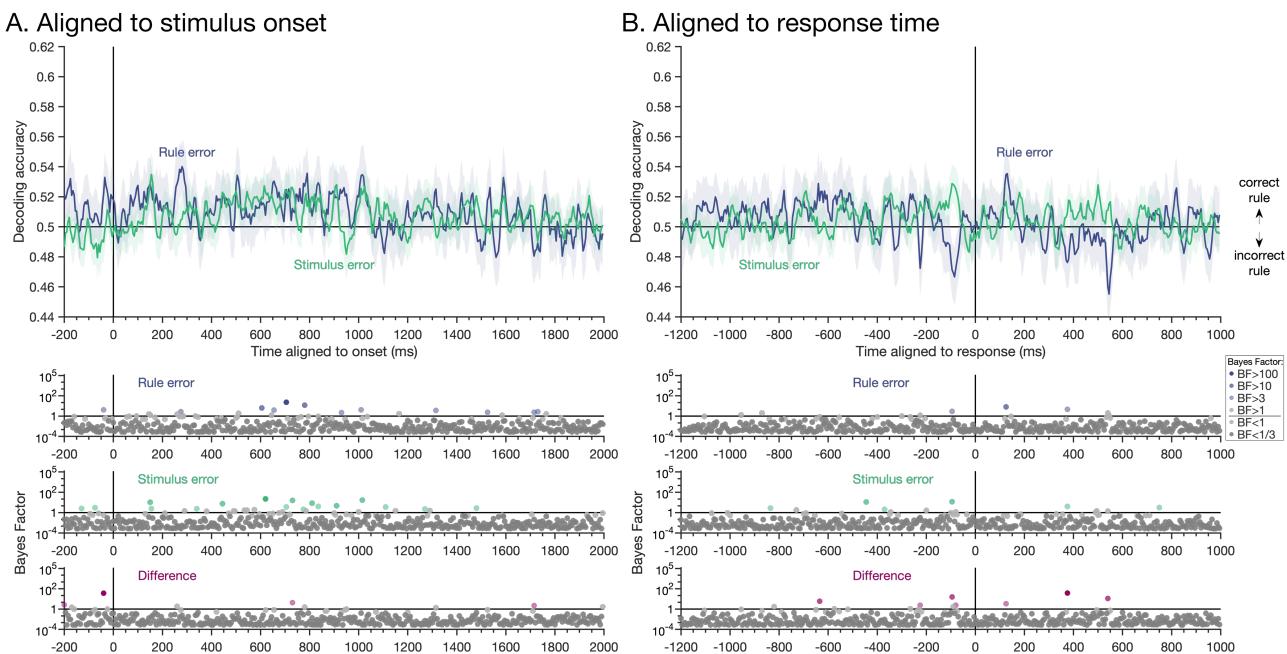


398  
 399 Figure 4. **Stimulus position decoding** on rule and stimulus error trials. A) Decoding analyses conducted relative to  
 400 stimulus onset revealed initial stimulus coding on both rule and stimulus error trials, with sustained stimulus coding  
 401 on rule error trials (similar to correct trials, Figure 3), but no stimulus coding at later timepoints on stimulus error  
 402 trials. The interaction, shown by BF difference (pink) confirmed that at later timepoints there was more evidence for  
 403 stimulus coding on rule errors than stimulus errors. B) Decoding analyses conducted relative to response time  
 404 revealed evidence for correct stimulus coding on rule error trials and evidence for *incorrect* stimulus coding on  
 405 stimulus error trials, evident as below-chance decoding accuracy. Decoding accuracy is smoothed with a 20ms window  
 406 for visualisation. Bayes factors are shown on a log scale and colour coded according to amount of evidence.

#### 407 *Rule decoding*

408 Next, we asked whether the representation of task *rule* in the correct trials would also generalise  
 409 to error trials. However, there were no sustained periods of evidence for rule information coding

410 on rule errors, and only very brief evidence for above-chance rule decoding on stimulus errors,  
 411 whether we aligned the MEG data to the stimulus onset (Figure 5A) or response (Figure 5B). There  
 412 was also no difference in rule coding between error types. For onset-aligned analyses, Bayes  
 413 Factors indicated evidence for above-chance decoding on rule and stimulus errors for some time  
 414 points, but it was not sustained. There were also some brief periods of below-chance decoding on  
 415 rule errors for response-aligned analyses, which would indicate coding of the incorrect rule,  
 416 consistent with behaviour, but this did not reach our interpretation levels for BFs (2 consecutive  
 417 timepoints  $BF > 3$ ). Overall, rule information that had been (weakly) present on correct trials was  
 418 largely absent on both types of behavioural error. Moreover, the *reversal* in coding – in this case  
 419 coding of the incorrect rule – was not evident as it was for stimulus coding.



420

421 Figure 5. **Rule decoding** on rule and stimulus error trials. A) Decoding analyses conducted relative to stimulus onset. B)  
 422 Decoding analyses conducted relative to response time. Decoding accuracy is smoothed with a 20ms window for  
 423 visualisation. There was substantial evidence for the null (i.e., that rule could not be decoded) on both stimulus error  
 424 and rule error trials, indicating the rule coding did not reverse for either type of error.

425 Taken together, the MEG decoding results show that on correct trials, all task-relevant aspects  
426 (stimulus position, cue colour, rule, response) could be decoded. The dynamics of coding varied  
427 such that analyses revealed stimulus-locked coding of perceptual features (stimulus position, cue)  
428 early in the time course, and analyses aligned to the response revealed coding of relevant task  
429 aspects (position, rule) and the resulting motor response ramping up prior to the response being  
430 given. Strikingly, the error decoding analysis showed that the increased stimulus position coding in  
431 the latter part of the epoch reflected the participant's decision about the stimulus more closely  
432 than the physical stimulus presented to them, providing strong evidence for the connection  
433 between specific neural responses decoded with MVPA and behaviour.

## Discussion

435 In this study we used MVPA with MEG to characterise the neural dynamics of stimulus, cue, rule  
436 and response coding in a difficult response-mapping task, and the link between these codes and  
437 behaviour. Our results showed a clear and orderly progression of task-relevant information coding  
438 after the stimulus was presented, while analyses aligned to the response time revealed that  
439 information coding for the stimulus and motoric response ramped up over the ~1 second before  
440 the response was given, in a manner reminiscent of evidence accumulation (e.g., Pisauro et al.,  
441 2017; Tagliabue et al., 2019). Strikingly, for trials on which participants made an error in the  
442 stimulus position, information coding initially corresponded to physical stimulation, but later  
443 accumulated in the opposite direction, so that activity patterns at later timepoints resembled  
444 those encoding the *incorrect* stimulus. This provides a crucial demonstration that patterns of  
445 neural activity recorded and classified in this way can be predictive of behaviour. These findings  
446 give insight into the dynamics of processes underlying cognitive control and provide a clear link  
447 between neural responses and behaviour.

448 The difficult response-mapping task implemented in this study required complex processing for  
449 successful performance. The task involved processing different types of perceptual information  
450 (cue, stimulus), conversion of the cue into the appropriate rule, application of the relevant rule to  
451 the stimulus position, and selection of the correct button-press response. Using MEG decoding  
452 with a carefully balanced experimental design, we were able to investigate the coding of each of  
453 these types of relevant information over time, and observe the succession of the different task-  
454 related features. We summarise and consider the findings below.

455 **Information coding after stimulus onset: correct trials**

456 Our results demonstrated different dynamics for perceptual, rule-related and motor processes for  
457 analyses aligned to stimulus onset (Figure 3A). Stimulus position was represented early in the time  
458 course (<100ms after stimulus onset), consistent with early retinotopic visual processes (Di Russo  
459 et al., 2005; Im et al., 2007). The cue was represented shortly thereafter at a time that is  
460 consistent with general colour processing and in line with previous work that found colour  
461 decoding was most evident from 135-155ms after image onset (Teichmann et al., 2019, 2020).  
462 Rule information, by contrast, was most evident at around 600ms and maintained for longer than  
463 cue information, perhaps reflecting the ongoing process of combining the stimulus and rule  
464 information to derive the response. Motor responses emerged last and exhibited a broad, shallow  
465 peak, perhaps reflective of the wide range of response times in the task. Our data thus emphasise  
466 the progression of coding for different types of task-relevant information over time, despite the  
467 relevant sensory information (stimulus and cue) being presented simultaneously.

468 These onset-aligned analyses are consistent with previous time-resolved multivariate analyses  
469 that showed progression of task-related information after stimulus presentation (Hebart et al.,  
470 2018; Hubbard et al., 2019; Kikumoto & Mayr, 2020; Wen et al., 2019). For example, Hebart et al  
471 (2018) showed that task-related information was evident in the MEG signal shortly after a task cue  
472 was presented, but ramped up again after the target object was presented. In a more complex  
473 design, Hubbard et al (2019) used cued task-switching with electroencephalography (EEG) which  
474 allowed them to look at coding of multiple task-related aspects over time using oscillatory power  
475 in the neural signal. Like our cue and rule results, they showed cue information preceded task rule  
476 information, and task decoding was prolonged throughout the trial period. Relevant and irrelevant  
477 stimulus information was evident after the stimuli were presented, and response information was  
478 present later in the signal. Here, we show that a similar cascade of information arises when the

479 cue and target stimulus are presented simultaneously. Our results and this previous work show  
480 that different task features are represented in the brain with different temporal dynamics, but  
481 there are time periods when multiple types of features are represented, potentially giving the  
482 neural correlates for the information integration needed on these tasks.

483 The cue colour decoding we observed is indicative of transient cue processing before the relevant  
484 rule was selected, at which point the colour distinctions (e.g., blue vs yellow, when both indicate  
485 rule 1) becomes irrelevant. In contrast, the prolonged coding of stimulus position and rule (far  
486 exceeding the stimulus presentation time of 500ms) likely reflects position and rule information  
487 being maintained, accumulated and/or manipulated as it is combined to achieve the correct  
488 response. Previous fMRI research has shown that the MD regions in frontoparietal cortex  
489 represent a range of information including details of stimuli and task rules (Woolgar et al., 2016),  
490 with particular emphasis on information that is task relevant (Woolgar et al. 2015a, Jackson et al  
491 2016), perceptually confusable (Woolgar et al., 2011) or difficult, like our task rules (Woolgar,  
492 Afshar, et al., 2015). The representations of stimulus position and rules we observed at late stages  
493 of processing (> 500ms after stimulus onset) would be theoretically consistent with processing  
494 within higher-level frontoparietal regions. For example, previous combined MEG-fMRI research  
495 has shown task-related enhancement of relevant features occurred after 500ms following stimulus  
496 presentation, and task coding in posterior parietal cortex and lateral prefrontal cortex seemed to  
497 peak from 500ms (Hebart et al., 2018). Moreover, attention enhances task-relevant information in  
498 frontoparietal regions from 500ms (Moerel et al., 2021). On the other hand, task-relevant stimulus  
499 information also seems to persist in occipital regions until these late timepoints (Goddard et al.,  
500 2016; Hebart et al., 2018; Moerel et al., 2021). Unfortunately, the limited spatial resolution of  
501 MEG makes it difficult to localise where our neural information arose through time. Future work

502 could address this, perhaps using computational methods to combine fMRI and MEG data such as  
503 similarity-based fusion (Cichy et al., 2014).

504 **Information coding prior to the response: correct trials**

505 The analyses aligned to the response time provided rich additional information about the  
506 temporal dynamics of stimulus, cue, rule and response coding (Figure 3B). We expected that these  
507 response-aligned analyses would emphasise higher-level decision-related processes required for  
508 behaviour which might not be so salient in data aligned to stimulus onset because of variability in  
509 their timing (Vidaurre et al., 2018). Previous EEG work has shown neural signals ramp up during  
510 perceptual decision-making, which has been described as evidence accumulation (e.g., Pisauro et  
511 al., 2017; Tagliabue et al., 2019), but these effects could be related to a general decision-making  
512 process rather than involving information about the stimulus of interest, and could be confounded  
513 with preparatory motor activity. Here, using decoding, we were able to assess the dynamics of  
514 different types of task-related information, separate from and in addition to response information,  
515 that was represented in the brain before the response. The results revealed an increase in  
516 stimulus information from approximately 1000ms prior to the response that peaked around the  
517 response time, a pattern which was noticeably absent in the onset-locked analyses. Response  
518 coding, by contrast, showed a later, sharper ramping in information that peaked just after the  
519 response was made. The ramping of stimulus position and response coding was, for the most part,  
520 when the stimulus and cue were no longer visible; the stimulus and cue were only presented for  
521 500ms, and the median response time was over 1400ms, so on a typical trial there was no  
522 stimulus presented in the 900ms prior to the response. Therefore, instead of perceptual  
523 accumulation, these pre-response representations appear to be internally generated codes that  
524 reflect the system moving towards different end states as the person arrives at their decision.

525 The results revealed concurrent coding of position, rule and response information prior to the  
526 response, which might reflect the need to combine position and rule information to select the  
527 appropriate response. Kikumoto & Mayr (2020) recently investigated the temporal dynamics of  
528 action selection using EEG in a cued rule selection task and provided evidence that conjunctions  
529 between task-relevant features are necessary for action selection. In addition to the succession of  
530 individual task features, they found rule-stimulus-response conjunctive representations could be  
531 decoded using stimulus-aligned EEG, and the strength of the conjunctive information was  
532 associated with faster responses, providing a link with behaviour. Other work has used temporal  
533 decomposition of EEG data and concluded that stimulus-response bindings have different  
534 temporal profiles to stimulus information, with gradual activation and decay over time (Takacs et  
535 al., 2020). We did not explicitly set out to look at conjunctive representations, but our results  
536 certainly fit with this account. During goal-directed behaviour, it seems that multiple task-relevant  
537 features are represented concurrently, presumably reflecting the need for this information to be  
538 maintained, and are then combined over time.

539 Our onset-aligned and response-aligned analyses revealed complementary aspects of the data.  
540 The pattern of results suggests that onset-aligned analyses may be most sensitive to perceptual  
541 responses, while response-aligned analyses may capture processes that are time-varied relative to  
542 stimulus onset and more closely yoked to the time of response, such as higher decision and motor  
543 preparation processes. Specifically, we found that stimulus position and cue colour had sharp  
544 initial decoding when aligned to stimulus onset, which was not visible after realignment to  
545 response. However, neural representations of stimulus and response exhibited a ramping  
546 accumulation before the button was pressed, that was not visible in onset-aligned data. This  
547 highlights the utility of including both approaches, perhaps particularly for difficult tasks with

548 substantial response time variability, to yield additional information about the dynamics  
549 underlying successful task performance.

550 **Information coding leading to incorrect behaviour: error trials**

551 To test whether the neural coding of task-relevant information detected with MVPA reflects  
552 activity necessary to successfully perform a task, we examined how these codes changed when  
553 participants made errors. We focused on decoding stimulus and rule information during stimulus  
554 errors and rule errors, situations in which the decision made could be dissociated from the  
555 stimulus and rule cue presented. Stimulus errors consisted of trials on which participants correctly  
556 applied the rule but confused the stimulus. Despite the behavioural evidence for correct rule use  
557 on these trials, there was only some evidence of rule coding, perhaps reflecting weak rule coding  
558 in general (on correct trials) and the limited number of error trials. However, stimulus position  
559 coding on stimulus error trials revealed a striking result: initial stimulus coding showed some  
560 fleeting evidence of the correct stimulus neural pattern, but prior to the response, stimulus coding  
561 became consistent with the incorrect stimulus. Thus, onset-aligned analyses and responses at  
562 early timepoints reflected perception, while response-aligned analyses and coding at later  
563 timepoints reflected behaviour. Recall in the paradigm that stimulus position was designed to be  
564 confusable, and a stimulus position error, by definition, means participants confused two (out of  
565 four) stimulus positions. When the stimulus was presented, participants would see it, which is  
566 consistent with brief veridical stimulus position decoding, but the insufficient maintenance of this  
567 information correlates with the behavioural performance: participants could not localise the  
568 stimulus precisely, which led to a decision in favour of the wrong stimulus. It is this internal  
569 decision-related process that seems to be reflected in below-chance (incorrect stimulus) decoding  
570 before the response. Previous work has shown that higher, but not early, perceptual regions,  
571 reflect behaviour in terms of accuracy (Walther et al., 2009; Williams et al., 2007) and reaction

572 time (Grootswagers et al., 2018), although none of these studies revealed the code reversal  
573 needed for a strong link with behaviour. Here, we used the temporal domain to show *what* was  
574 coded on error trials at different stages of processing. There was a dissociation between the  
575 coding of early perceptual information and the stimulus decision used to generate the behavioural  
576 response.

577 Rule errors consisted of trials on which participants appeared (in their behaviour) to apply the  
578 wrong rule to the correct stimulus. Accordingly, for stimulus position coding, the classifier trained  
579 on correct trials could successfully classify the stimulus position after onset and prior to the  
580 response on rule error trials. This indicates that the stimulus coding reversal seen above was  
581 diagnostic of the particular type of behavioural error, rather than reflective of errors in general,  
582 indicating a tight link with the specific decision made and reflected in behaviour.

583 For rule coding, we again found little evidence for rule coding on rule error trials. A couple of  
584 timepoints just prior to the response showed patterns of activity consistent with the incorrect  
585 rule, as we had predicted for a full double dissociation, but the effect was so transient that it is  
586 difficult to interpret with confidence. This may reflect the very small number of rule error trials,  
587 and/or the relatively weak coding of rule information in general in our data (potentially  
588 attributable to more variability in timing of this task aspect, and/or relatively poor signal from  
589 frontal regions that are further from the sensors in our supine MEG system). This limitation means  
590 that, in contrast to the stimulus information, we cannot conclude with confidence whether or not  
591 the rule patterns we decoded were closely linked to behaviour.

592 Our research contributes to the growing literature drawing links between neural responses and  
593 behaviour using MVPA. Using spatial and temporal neuroimaging, classifier prediction errors and  
594 distances from the classifier boundary have been shown to correlate with behavioural error

595 patterns and reaction times (e.g., Carlson et al., 2014; González-García et al., 2021; Grootswagers  
596 et al., 2018; Walther et al., 2009). Here, we argue that a tighter link between brain and behaviour  
597 can be found by testing *what* is represented on error trials when an incorrect decision is made.  
598 The results parallel fMRI work showing frontoparietal MD regions represent the correct stimulus  
599 but wrong rule during a rule error, and the correct rule but wrong stimulus during other errors  
600 (Woolgar et al., 2019). The current study extends this work by elucidating the dynamics with which  
601 the incorrect representations evolve over time, with early representations reflecting the stimuli  
602 presented, but a late gradual accumulation towards the opposite stimulus at timepoints just prior  
603 to behavioural response. We also show here that there is a dissociation in the perceptual coding  
604 (indexed by onset-aligned analyses) and high-level decision coding (indexed by response-aligned  
605 analyses) for error trials. Specifically, on stimulus errors, after a transient representation of the  
606 veridical stimulus, activity accumulated towards a pattern state reflecting the opposite and  
607 incorrect stimulus, apparently reflecting the internal generation of accumulation towards the  
608 wrong decision. This pattern was specifically diagnostic of behavioural errors attributable to  
609 stimulus misperception, as position information was coded correctly on other types of behavioural  
610 errors.

611 The results of this study provide new insights into how task-relevant information is processed in  
612 the human brain to allow successful goal-directed behaviour. There was a clear progression of the  
613 onset of task-relevant information in the brain, from stimulus position and cue, to rule and then  
614 response information. Complimentary response-aligned analyses, which highlight later high-level  
615 processes aligned in time to behaviour, additionally revealed dynamics of information coding  
616 resembling an accumulation of multiple types of task-relevant information. Moreover, when  
617 participants made behavioural errors, the direction of accumulation was reversed. Under these  
618 conditions, the trajectory of representation moved in the opposite direction such that the neural

619 pattern increasingly represented the incorrect stimulus (which had not been shown) in a manner  
620 diagnostic of the subsequent behavioural choice. The data highlight the orderly but overlapping  
621 dynamics with which several task elements can be represented in brain activity. Our findings  
622 emphasise a particular role for the trajectory of information coding at later time points in  
623 determining behavioural success or failure, and demonstrate the utility of aligning neural data  
624 differently to examine high-level complex cognitive processes.

## References

- 626 Battistoni, E., Kaiser, D., Hickey, C., & Peelen, M. V. (2020). The time course of spatial attention  
627 during naturalistic visual search. *Cortex*, 122, 225–234.  
628 <https://doi.org/10.1016/j.cortex.2018.11.018>
- 629 Bracci, S., Daniels, N., & Op de Beeck, H. (2017). Task Context Overrules Object- and Category-  
630 Related Representational Content in the Human Parietal Cortex. *Cerebral Cortex*, 27(1),  
631 310–321. <https://doi.org/10.1093/cercor/bhw419>
- 632 Carlson, T. A., Grootswagers, T., & Robinson, A. K. (2020). An introduction to time-resolved  
633 decoding analysis for M/EEG. In *The Cognitive Neurosciences*. MIT Press.
- 634 Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution  
635 decoding of object position and category. *Journal of Vision*, 11(2011), 1–17.  
636 <https://doi.org/10.1167/11.10.9>
- 637 Carlson, T. A., Ritchie, J. B., Kriegeskorte, N., Durvasula, S., & Ma, J. (2014). Reaction Time for  
638 Object Categorization Is Predicted by Representational Distance. *Journal of Cognitive  
639 Neuroscience*, 26(1), 132–142. [https://doi.org/10.1162/jocn\\_a\\_00476](https://doi.org/10.1162/jocn_a_00476)
- 640 Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and  
641 time. *Nature Neuroscience*, 17(3), 455–462. <https://doi.org/10.1038/nn.3635>
- 642 Crittenden, B. M., Mitchell, D. J., & Duncan, J. (2016). Task Encoding across the Multiple Demand  
643 Cortex Is Consistent with a Frontoparietal and Cingulo-Opercular Dual Networks  
644 Distinction. *The Journal of Neuroscience*, 36(23), 6147–6155.  
645 <https://doi.org/10.1523/JNEUROSCI.4590-15.2016>

- 646 de-Wit, L., Alexander, D., Ekroll, V., & Wagemans, J. (2016). Is neuroimaging measuring  
647 information in the brain? *Psychonomic Bulletin & Review*, 23(5), 1415–1428.  
648 <https://doi.org/10.3758/s13423-016-1002-0>
- 649 Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG  
650 dynamics including independent component analysis. *Journal of Neuroscience Methods*,  
651 134, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- 652 Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D., & Hillyard, S. A. (2005).  
653 Identification of the neural sources of the pattern-reversal VEP. *NeuroImage*, 24(3), 874–  
654 886. <https://doi.org/10.1016/j.neuroimage.2004.09.029>
- 655 Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on  
656 Psychological Science*, 6(3), 274–290. <https://doi.org/10.1177/1745691611406920>
- 657 Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical  
658 Psychology*, 72, 78–89. <https://doi.org/10.1016/j.jmp.2015.10.003>
- 659 Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for  
660 intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172–179.  
661 <https://doi.org/10.1016/j.tics.2010.01.004>
- 662 Duncan, J., Assem, M., & Shashidhara, S. (2020). Integrated Intelligence from Distributed Brain  
663 Activity. *Trends in Cognitive Sciences*, 24(10), 838–852.  
664 <https://doi.org/10.1016/j.tics.2020.06.012>
- 665 Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2  
666 component of the ERP: A review. *Psychophysiology*, 45(1), 152–170.  
667 <https://doi.org/10.1111/j.1469-8986.2007.00602.x>

- 668 Goddard, E., Carlson, T. A., Dermody, N., & Woolgar, A. (2016). Representational dynamics of  
669 object recognition: Feedforward and feedback information flows. *NeuroImage*, 128, 385–  
670 397. <https://doi.org/10.1016/j.neuroimage.2016.01.006>
- 671 González-García, C., Formica, S., Wisniewski, D., & Brass, M. (2021). Frontoparietal action-oriented  
672 codes support novel instruction implementation. *NeuroImage*, 226, 117608.  
673 <https://doi.org/10.1016/j.neuroimage.2020.117608>
- 674 Gratton, G., Cooper, P., Fabiani, M., Carter, C. S., & Karayavidis, F. (2017). Dynamics of cognitive  
675 control: Theoretical bases, paradigms, and a view for the future. *Psychophysiology*, June,  
676 1–29. <https://doi.org/10.1111/psyp.13016>
- 677 Grootswagers, T., Cichy, R. M., & Carlson, T. A. (2018). Finding decodable information that can be  
678 read out in behaviour. *NeuroImage*, 179, 252–262.  
679 <https://doi.org/10.1016/j.neuroimage.2018.06.022>
- 680 Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from  
681 Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series  
682 Neuroimaging Data. *Journal of Cognitive Neuroscience*, 29(4), 677–697.  
683 [https://doi.org/10.1162/jocn\\_a\\_01068](https://doi.org/10.1162/jocn_a_01068)
- 684 Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *NeuroImage*, 62(2),  
685 852–855. <https://doi.org/10.1016/j.neuroimage.2012.03.016>
- 686 Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed  
687 and overlapping representations of faces and objects in ventral temporal cortex. *Science*,  
688 293(5539), 2425–2430. <https://doi.org/10.1126/science.1063736>
- 689 Hebart, M. N., & Baker, C. I. (2018). Deconstructing multivariate decoding for the study of brain  
690 function. *NeuroImage*, 180(A), 4–18. <https://doi.org/10.1016/j.neuroimage.2017.08.005>

- 691 Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational  
692 dynamics of task and object processing in humans. *eLife*, 7, 1–21.  
693 <https://doi.org/10.7554/eLife.32816>
- 694 Hubbard, J., Kikumoto, A., & Mayr, U. (2019). EEG Decoding Reveals the Strength and Temporal  
695 Dynamics of Goal-Relevant Representations. *Scientific Reports*, 9(1), 9051.  
696 <https://doi.org/10.1038/s41598-019-45333-6>
- 697 Im, C.-H., Gururajan, A., Zhang, N., Chen, W., & He, B. (2007). Spatial Resolution of EEG Cortical  
698 Source Imaging Revealed by Localization of Retinotopic Organization in Human Primary  
699 Visual Cortex. *Journal of Neuroscience Methods*, 161(1), 142–154.  
700 <https://doi.org/10.1016/j.jneumeth.2006.10.008>
- 701 Jackson, J. B., & Woolgar, A. (2018). Adaptive coding in the human brain: Distinct object features  
702 are encoded by overlapping voxels in frontoparietal cortex. *Cortex*, 108, 25–34.  
703 <https://doi.org/10.1016/j.cortex.2018.07.006>
- 704 Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Feature-selective Attention in  
705 Frontoparietal Cortex: Multivoxel Codes Adjust to Prioritize Task-relevant Information.  
706 *Journal of Cognitive Neuroscience*, 29(2), 310–321. [https://doi.org/10.1162/jocn\\_a\\_01039](https://doi.org/10.1162/jocn_a_01039)
- 707 Jeffreys, H. (1961). *Theory of probability* (Third). Oxford University Press.
- 708 Kado, H., Higuchi, M., Shimogawara, M., Haruta, Y., Adachi, Y., Kawai, J., Ogata, H., & Uehara, G.  
709 (1999). Magnetoencephalogram systems developed at KIT. *IEEE Transactions on Applied  
710 Superconductivity*, 9(2), 4057–4062. <https://doi.org/10.1109/77.783918>
- 711 Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., & Forstmann, B. U. (2010).  
712 Advance Preparation in Task-Switching: Converging Evidence from Behavioral, Brain

- 713 Activation, and Model-Based Approaches. *Frontiers in Psychology*, 1, 25.
- 714 <https://doi.org/10.3389/fpsyg.2010.00025>
- 715 Karimi-Rouzbahani, H., Woolgar, A., & Rich, A. N. (2021). Neural signatures of vigilance
- 716 decrements predict behavioural errors before they occur. *eLife*, 10, e60563.
- 717 <https://doi.org/10.7554/eLife.60563>
- 718 Kikumoto, A., & Mayr, U. (2020). Conjunctive representations that integrate stimuli, responses,
- 719 and rules are critical for action selection. *Proceedings of the National Academy of Sciences*,
- 720 117(19), 10603–10608. <https://doi.org/10.1073/pnas.1922166117>
- 721 Long, N. M., & Kuhl, B. A. (2018). Bottom-Up and Top-Down Factors Differentially Influence
- 722 Stimulus Representations Across Large-Scale Attentional Networks. *Journal of*
- 723 *Neuroscience*, 38(10), 2495–2504. <https://doi.org/10.1523/JNEUROSCI.2724-17.2018>
- 724 Moerel, D., Rich, A. N., & Woolgar, A. (2021). *Selective attention and decision-making have*
- 725 *separable neural bases in space and time* [Preprint].
- 726 <https://doi.org/10.1101/2021.02.28.433294>
- 727 Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018, May 19). *Package*
- 728 “*BayesFactor*.” <https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf>
- 729 Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate
- 730 Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in*
- 731 *Neuroinformatics*, 10. <https://doi.org/10.3389/fninf.2016.00027>
- 732 Pisauro, M. A., Fouragnan, E., Retzler, C., & Philiastides, M. G. (2017). Neural correlates of
- 733 evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI.
- 734 *Nature Communications*, 8(1), 15808. <https://doi.org/10.1038/ncomms15808>

- 735 Posner, M. I., & Presti, D. E. (1987). Selective attention and cognitive control. *Trends in*  
736 *Neurosciences*, 10(1), 13–17. [https://doi.org/10.1016/0166-2236\(87\)90116-0](https://doi.org/10.1016/0166-2236(87)90116-0)
- 737 Posner, M. I., & Snyder, C. R. R. (1975). Attention and Cognitive Control. In R. L. Solso (Ed.),  
738 *Information Processing and Cognition: The Loyola Symposium*. Lawrence Erlbaum.
- 739 Ritchie, J. B., & Carlson, T. A. (2016). Neural decoding and “inner” psychophysics: A distance-to-  
740 bound approach for linking mind, brain, and behavior. *Frontiers in Neuroscience*, 10(APR),  
741 1–8. <https://doi.org/10.3389/fnins.2016.00190>
- 742 Ritchie, J. B., Tovar, D. A., & Carlson, T. A. (2015). Emerging Object Representations in the Visual  
743 System Predict Reaction Times for Categorization. *PLOS Computational Biology*, 11(6),  
744 e1004316. <https://doi.org/10.1371/journal.pcbi.1004316>
- 745 Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for  
746 accepting and rejecting the null hypothesis. *Psychon Bull Rev*, 16(2), 225–237.  
747 <https://doi.org/10.3758/PBR.16.2.225>
- 748 Tagliabue, C. F., Veniero, D., Benwell, C. S. Y., Cecere, R., Savazzi, S., & Thut, G. (2019). The EEG  
749 signature of sensory evidence accumulation during decision formation closely tracks  
750 subjective perceptual experience. *Scientific Reports*, 9(1), 4949.  
751 <https://doi.org/10.1038/s41598-019-41024-4>
- 752 Takacs, A., Mückschel, M., Roessner, V., & Beste, C. (2020). Decoding Stimulus–Response  
753 Representations and Their Stability Using EEG-Based Multivariate Pattern Analysis.  
754 *Cerebral Cortex Communications*, 1. <https://doi.org/10.1093/texcom/tgaa016>
- 755 Teichmann, L., Grootswagers, T., Carlson, T. A., & Rich, A. N. (2019). Seeing versus knowing: The  
756 temporal dynamics of real and implied colour processing in the human brain. *NeuroImage*,  
757 200, 373–381. <https://doi.org/10.1016/j.neuroimage.2019.06.062>

- 758 Teichmann, L., Quek, G. L., Robinson, A. K., Grootswagers, T., Carlson, T. A., & Rich, A. N. (2020).  
759           The influence of object-colour knowledge on emerging object representations in the brain.  
760           *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.0158-20.2020>
- 761 Vidaurre, D., Myers, N. E., Stokes, M., Nobre, A. C., & Woolrich, M. W. (2018). Temporally  
762           unconstrained decoding reveals consistent but time-varying stages of stimulus processing.  
763           *Cerebral Cortex*, 1–12, 260943–260943. <https://doi.org/10.1101/260943>
- 764 Wagenmakers, E. (2007). A practical solution to the pervasive problems of p values. *Psychonomic  
765 Bulletin and Review*, 14(5), 779–804. <https://doi.org/10.3758/BF03194105>
- 766 Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural Scene Categories Revealed in  
767           Distributed Patterns of Activity in the Human Brain. *Journal of Neuroscience*, 29(34),  
768           10573–10581. <https://doi.org/10.1523/JNEUROSCI.0559-09.2009>
- 769 Wen, T., Duncan, J., & Mitchell, D. J. (2019). The time-course of component processes of selective  
770           attention. *NeuroImage*, 199, 396–407. <https://doi.org/10.1101/511022>
- 771 Wetzel, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011).  
772           Statistical evidence in experimental psychology: An empirical comparison using 855 t tests.  
773           *Perspectives on Psychological Science*, 6(3), 291–298.  
774           <https://doi.org/10.1177/1745691611406923>
- 775 Wetzel, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and  
776           partial correlations. *Psychonomic Bulletin & Review*, 19(6), 1057–1064.  
777           <https://doi.org/10.3758/s13423-012-0295-x>
- 778 Williams, M. A., Dang, S., & Kanwisher, N. G. (2007). Only some spatial patterns of fMRI response  
779           are read out in task performance. *Nature Neuroscience*, 10(6), 685–686.  
780           <https://doi.org/10.1038/nn1900>

- 781 Woolgar, A., Afshar, S., Williams, M. A., & Rich, A. N. (2015). Flexible Coding of Task Rules in  
782 Frontoparietal Cortex: An Adaptive System for Flexible Cognitive Control. *Journal of*  
783 *Cognitive Neuroscience*, 27(10), 1895–1911. [https://doi.org/10.1162/jocn\\_a\\_00827](https://doi.org/10.1162/jocn_a_00827)
- 784 Woolgar, A., Dermody, N., Afshar, S., Williams, M. A., & Rich, A. N. (2019). *Meaningful patterns of*  
785 *information in the brain revealed through analysis of errors* (p. 673681) [BioRxiv preprint].  
786 <https://doi.org/10.1101/673681>
- 787 Woolgar, A., Duncan, J., Manes, F., & Fedorenko, E. (2018). Fluid intelligence is supported by the  
788 multiple-demand system not the language system. *Nature Human Behaviour*, 2(3), 200–  
789 204. <https://doi.org/10.1038/s41562-017-0282-3>
- 790 Woolgar, A., Jackson, J., & Duncan, J. (2016). Coding of Visual, Auditory, Rule, and Response  
791 Information in the Brain: 10 Years of Multivoxel Pattern Analysis. *Journal of Cognitive*  
792 *Neuroscience*, 28(10), 1433–1454. <https://doi.org/10.1162/jocn>
- 793 Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., Roca, M., Antoun,  
794 N., Manes, F., & Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of  
795 damage within frontal and parietal cortex. *Proceedings of the National Academy of*  
796 *Sciences*, 107(33), 14899–14902. <https://doi.org/10.1073/pnas.1007928107>
- 797 Woolgar, A., Williams, M. A., & Rich, A. N. (2015). Attention enhances multi-voxel representation  
798 of novel objects in frontal, parietal and visual cortices. *NeuroImage*, 109, 429–437.  
799 <https://doi.org/10.1016/j.neuroimage.2014.12.083>
- 800 Woolgar, A., & Zopf, R. (2017). Multi-sensory coding in the multiple-demand regions: Vibrotactile  
801 task information is coded in frontoparietal cortex. *Journal of Neurophysiology*,  
802 jn.00559.2016-jn.00559.2016. <https://doi.org/10.1152/jn.00559.2016>

- 803 Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive Coding of Task-Relevant  
804 Information in Human Frontoparietal Cortex. *Journal of Neuroscience*, 31(41), 14592–  
805 14599. <https://doi.org/10.1523/JNEUROSCI.2616-11.2011>
- 806 Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. *Trabajos de  
807 Estadistica Y de Investigacion Operativa*, 31(1), 585–603.  
808 <https://doi.org/10.1007/BF02888369>

809