

1 **Excitatory neurons of the anterior cingulate cortex encode chosen**
2 **actions and their outcomes rather than cognitive state**

3

4 Martin M. Jendryka^{1*}, Uwe Lewin^{1,6}, Sampath K.T. Kapanaiiah¹, Hartmut Dermutz², Birgit
5 Liss^{1,3}, Anton Pekcec⁴, Thomas Akam^{5,#}, Benjamin F. Grewe^{2,#}, Dennis Kätzel^{1,*}

6

7 ¹Institute of Applied Physiology, Ulm University, Ulm, Germany

8 ²Institute for Neuroinformatics, ETH Zürich, Zürich, CH

9 ³Linacre College and New College, University of Oxford, Oxford, UK

10 ⁴Boehringer Ingelheim Pharma GmbH & Co. KG, Div. Research Germany, Biberach an der
11 Riss, Germany

12 ⁵Department of Experimental Psychology and Wellcome Centre for Integrative Neuroimaging,
13 University of Oxford, Oxford, UK

14 ⁶Present address: Max-Planck-Institute for Biological Intelligence, Martinsried, Munich,
15 Germany

16 #equal contribution

17 * Correspondence: martinjendryka@googlemail.com (M.M.J.) and dennis.kaetzel@uni-
18 [ulm.de](mailto:dennis.kaetzel@uni-ulm.de) (D.K.); +49 731 500 33770; Fax +49 731 500 33779; Institute of Applied Physiology,
19 Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany;

20

21

22

23 **Abstract**

24 The anterior cingulate cortex (ACC) causally influences cognitive control of goal-directed
25 behaviour. However, it is unclear whether ACC directly encodes cognitive variables like
26 attention or impulsivity, or implements goal-directed action selection mechanisms that are
27 modulated by them. We recorded ACC activity with miniature endoscopic microscopes in mice
28 performing the 5-choice-serial-reaction time task, and applied decoding and encoding
29 analyses. ACC pyramidal cells represented specific actions before and during the behavioural
30 response, whereas the response type (e.g. correct/incorrect/premature) – indicating the state
31 of attentional and impulse control – could only be decoded during and after the response with
32 high reliability. Devaluation and extinction experiments further revealed that action encoding
33 depended on reward expectation. Our findings support a role for ACC in goal-directed action
34 selection and monitoring, that is modulated by cognitive state, rather than in tracking levels of
35 attention or impulsivity directly in individual trials.

36

37

38 Introduction

39 Maintenance of high levels of sustained attention and inhibition of impulsive responding are
40 key to successful goal-directed behaviour, and impaired in a variety of psychiatric disorders
41 [1,2]. Both aspects can be measured by the 5-choice-serial-reaction-time task (5-CSRTT) [3,4]
42 in both humans and rodents. In this task, subjects can make four different types of response,
43 indicative of different cognitive states: (i) they can correctly respond to a stimulus presented
44 briefly and after a considerable waiting time to earn a reward (correct response), requiring
45 high attentional and impulse control. (ii) they can, instead, follow the impulsive urge to respond
46 before cue-presentation (premature response, indicating reduced impulse control), or (iii)
47 respond into a non-cued hole (incorrect response, indicating reduced attentional control). (iv)
48 Alternatively, they may not respond at all (omission, indicating reduced task engagement or
49 inattention). Therefore, the measurement of neurophysiological correlates of these four
50 response options, promises to identify circuits that regulate attention, impulse control, and
51 possibly other aspects of deterministic goal-directed behaviour.

52 Several rodent studies have implicated the anterior cingulate cortex (ACC) in this regulation.
53 Manipulations of rodent ACC have been shown to produce shifts in the relative occurrence of
54 these behavioural outcomes in the 5-CSRTT which support a causal role of this brain
55 structure. For example, the activation of G_i-protein signalling in excitatory pyramidal cells,
56 either in all layers or in layer 5 exclusively, may reduce premature and, partly, increase correct
57 responding [5]. In contrast, the chemogenetic inhibition of a subgroup of ACC neurons
58 projecting to the visual cortex may induce a shift from correct responding to response omission
59 [6], whereas their pre-cue stimulation at 30 Hz after such errors may have the opposite effect
60 [7]. The chemogenetic activation of ACC parvalbumin interneurons, in turn, reduces both
61 premature and incorrect responses, but not response omissions [8].

62 Studies with physiological measurement of neural activity in rodent ACC during the 5-CSRTT
63 and related tasks, have partly supported the possibility of such a causal role. One study
64 revealed that excitatory and inhibitory neurons in rat ACC may change their firing rate

65 differently both before and after correct vs. incorrect choices in the 5-CSRTT [9]. Specifically,
66 ramping neural activity in the ACC and the adjacent prelimbic cortex (PrL, upper part of the
67 rodent medial prefrontal cortex, mPFC) before cue-presentation has been interpreted as
68 preparatory signal under conditions that require high sustained attention, as this activity
69 increase was smaller before incorrect (low attention) responses and lowest during omissions
70 [9,10].

71 Several other studies using physiological measurements have, however, failed to find such an
72 indication of a causal role of ACC for modulating the occurrence of response options on a trial-
73 by-trial basis in the form of distinct pre-choice activity. They rather suggest that ACC may
74 monitor ongoing behaviour, and potentially provide feedback or error signals. For example,
75 neurons in rat ACC were shown to encode behaviour-related information mostly during and
76 after a choice, in a deterministic lever-based working memory task, thereby monitoring action
77 and outcome [11]. A specific subpopulation of ACC neurons that project to visual cortex was
78 selectively excited *after* incorrect choices or omissions (i.e. they conduct error-monitoring), but
79 their activity did not differ between those erroneous and correct choices while they were made
80 [7]. Imaging during a head-fixed Go/No-Go paradigm even found no evidence for a selective
81 recruitment of these neurons and projections for enhanced stimulus discrimination, but rather
82 that they simply represent rewarded action and stimuli [12]. Using a Go/No-go paradigm with
83 visual cues in mice, another group confirmed that ACC neurons are generally more likely
84 activated by cues that imply reward than those that do not, but also suggested that these cells
85 fire selectively either to signals that imply action or action restraint [13]. Another study in the
86 5-CSRTT also failed to detect much increase of firing rates of pyramidal neurons in the dorsal
87 PrL/ACC region before cue-onset, but found the modulation of their firing times by gamma-
88 oscillations in this period [14]. The role of the ACC may also depend on the task structure, as
89 it was shown that, in a probabilistic task, rat ACC neurons represent expected outcome first,
90 before switching to actual outcome in case of a mismatch between the two [11], which could

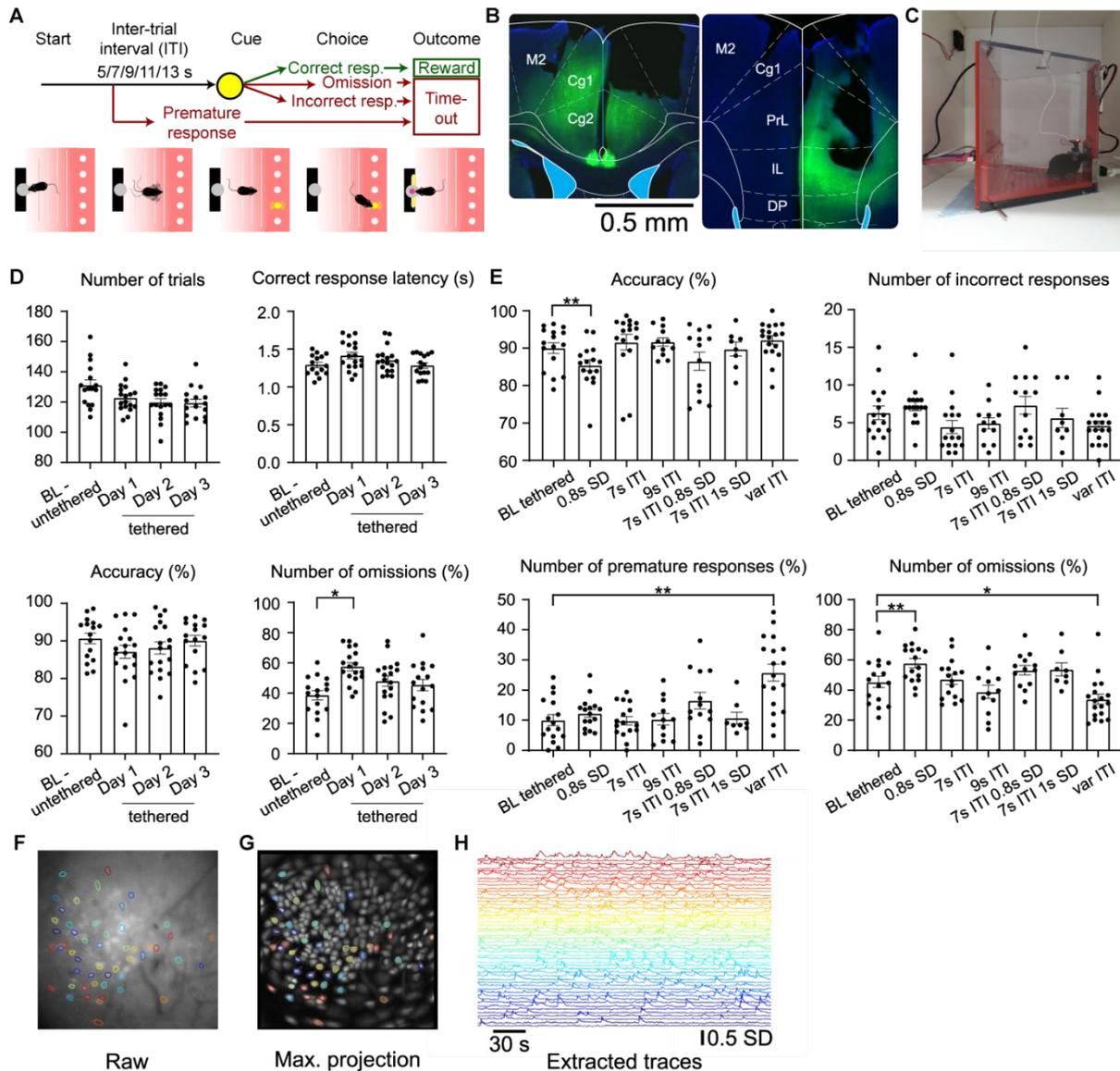
91 constitute a feedback signal for updating prior beliefs. This is in line with selective activity of
92 some ACC neurons after incorrect choices in the 5-CSRTT, constituting an error signal [7,9].

93 In summary, no clear picture of the mechanistic role of the rodent ACC in attention and impulse
94 control has emerged yet; whereas some studies found representations several seconds
95 before the choice event, which could indicate a causal role for choices or, generally, the
96 present cognitive state, others have found representations rather around the time of choice
97 itself and most profoundly during rewarded or after incorrect responses, which is more in line
98 with action- and outcome-monitoring, at least in deterministic tasks. Likewise, in monkeys,
99 ACC activity has been linked to error-monitoring, value representation and belief-updating
100 [15,16] rather than to attention *per se* [12]. Therefore, we here use simultaneous monitoring
101 of dozens of excitatory neurons with miniature endoscopic microscopes (miniscopes) in the 5-
102 CSRTT, in mice, in combination with time-resolved encoding and decoding analysis to reveal
103 which aspects of attentional, impulse and motor control are represented in the ACC at which
104 point in time.

105 **Results**

106 **Miniscope-based recording of neocortical activity in the 5-CSRTT**

107 To monitor activity of individual pyramidal neurons, we transduced ACC with an AAV5-vector
108 expressing the fluorescent calcium sensor GCaMP6m under the *CamKII α* -promoter [17], in
109 male C57BL/6J wildtype mice ($N = 12$), and implanted a gradient refractive index lens in a
110 separate surgery at the same location (Figure 1B). For comparison, we also generated a
111 smaller, second subgroup ($N = 6$), where activity was monitored in the ventral mPFC (Figure
112 1B), a region that was previously shown to represent rewarded choices [18]. Mice had been
113 pre-trained in the 5-CSRTT, and their training was continued after recovery from the second
114 surgery, until they reached a stable baseline.



115

Figure 1

116 **Figure 1. Behavioural performance with simultaneous miniscope recording.** (A)
 117 Structure of an individual trial of the 5-CSRTT (see Methods for description). (B) Selective
 118 transfection of the ACC (comprised of regions Cg1 and Cg2; left, AP 1 mm) and the ventral
 119 mPFC (at the border between the regions PrL and IL; right, AP 2 mm) with GCaMP6m
 120 expressed in excitatory cells; black gap in the right hemispheres indicates GRIN lens location.
 121 (D) Measures of task engagement and performance in the 5-CSRTT (as indicated above panels)
 122 during training sessions without miniscope or dummy (baseline, BL-untethered) and
 123 during the first three sessions performed with tethered miniscope (Day 1-3). Dots indicate
 124 individual animals, bars show mean \pm s.e.m.. Asterisks represent Dunnett pairwise post-hoc
 125 test comparing tethered days against baseline after significant effect of day in a one-way
 126 ANOVA. (E) Key performance indicators of the 5-CSRTT (as indicated above panels)
 127 measuring attention (accuracy, incorrect responses), impulse control (premature responses)
 128 and task engagement (omissions) for the baseline protocol and six challenge conditions during
 129 which miniscope recordings were conducted. Same display of mean \pm s.e.m. and statistics
 130 (comparison of each challenge against baseline) as in (D). (F) Example of a 400 μ m x 400 μ m
 131 raw image obtained from ACC of an individual mouse during the 5-CSRTT with a miniscope,
 132 with exemplary identified active neurons encircled in different colours corresponding to traces

133 shown in (H). **(G)** Similar display as in (F), but overlay of fields of view as maximum projection
134 from 12 animals as imaged in the same 5-CSRTT protocol in ACC (circles indicate the same
135 neurons as in (F)). **(H)** Example of z-scored calcium activity traces over 10 min measured with
136 GCaMP6m in the FOV shown in (F) from exemplary individual active neurons.

137

138 The recording of neural activity with miniscopes during operant tasks constitutes a challenge
139 due to the relatively large and protruding form factor of such microscopes and a
140 disadvantageous design of many operant box systems with deep and low recesses
141 constituting the reward receptacle and poke holes. To enable miniscope recordings during the
142 5-CSRTT, we designed a custom-made operant box system with shallow and elevated poke-
143 holes that reside in a protruding inner wall-layer (Figure 1C) [19]. This allowed mice to conduct
144 the task with little disturbance by the mounted and tethered miniscope (UCLA model v3 or v4;
145 Supplementary Video 1), as was confirmed by a lack of changes of achieved trial numbers,
146 response latency, attentional accuracy (number of correct responses/(number of correct and
147 incorrect response), and omissions (number of trials with omitted responses relative to total
148 number of trials, %) beyond the first day of tethered training (Figure 1D). With repeated
149 tethered training, animals performed well over 100 trials with less than 50% omissions on
150 average, providing sufficient numbers of active responses for further analysis (Figure 1D).

151 In order to maximally engage attentional and impulse control – and to obtain sufficient
152 numbers of incorrect and premature responses per session for later analysis - we performed
153 six behavioural challenges with simultaneous miniscope recordings; this included a further
154 shortening of the stimulus duration (SD) from 2 s at baseline to 0.8 or 1.0 s in challenge
155 conditions, and/or an extension of the waiting time (inter-trial interval, ITI) before stimulus
156 presentation from 5 s at baseline to fixed durations of 7 or 9 s or to variable lengths (7, 9, 11,
157 or 13 s randomly at equal distribution, varITI). As expected, attentional performance, as
158 indicated by accuracy, was lower with decreased stimulus duration (0.8 s SD challenge),
159 which, however, also increased omissions, making it less suitable for analysis (Figure 1E).
160 Overall, the varITI challenge appeared to produce the most suitable dataset for further

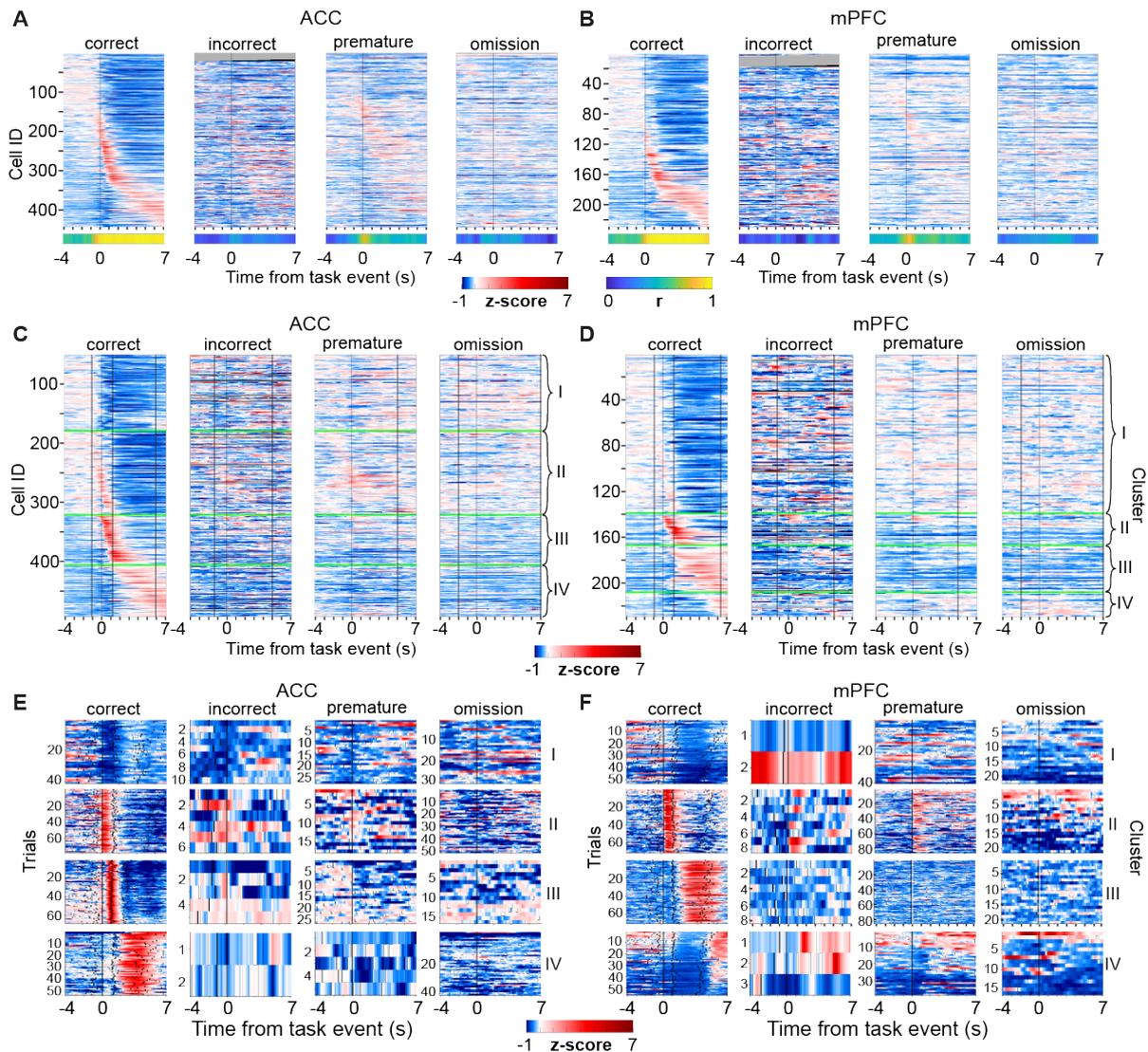
161 physiological analysis, given that the relative number of premature responses was increased
162 ($P < 0.0001$; Dunnett's post-hoc test after significant main effect of challenge in mixed-effects
163 ANOVA, $N = 18$) and the relative number of omissions was decreased ($P < 0.001$) compared
164 to the baseline protocol, whereas incorrect responses were still present at a level comparable
165 to the other test conditions (Figure 1E). We obtained stable recordings over 30 min sessions,
166 yielding 10-72 cells per field of view (FOV) in the ACC and 14-72 cells/FOV in the mPFC (see
167 Methods for details on trace extraction; Figure 1F-H).

168 **Individual ACC neurons have time-locked activity peaks around correct and** 169 **premature responses**

170 We first investigated qualitatively, if identified neurons display activity that is related to any of
171 the four behavioural response options. Therefore, the calcium signal traces of each neuron
172 were extracted from 4 s before until 7 s after each behavioural choice-poke event (note that,
173 for omissions, the end of the stimulus presentation, was used as reference time point of choice
174 for all time-locked analysis). Such individual episodic traces split into two populations of traces
175 from trials with either *even* or *odd* order number; the averages of traces from *even* trials were
176 then plotted in vertical order according to the peak latency of the averages of the
177 corresponding *odd* trials (Figure 2A-B; Supplementary Figure 1). This indicated that ACC
178 neurons often showed activity patterns that were time-locked to correct and premature
179 responses. In support of this conclusion, Pearson correlations between the averages of *odd*
180 and *even* trials indicated a high reproducibility of time-locked activity around *correct* and
181 *premature* responses, with particularly high correlations (> 0.8) around and after the time of
182 choice, in ACC and mPFC (Figure 2A-B, bottom; Supplementary Figure 1). Such correlated
183 patterns were largely absent for incorrect responses and omissions. This constitutes a first
184 indication that neural activity in both structures represents aspects of choices related to high
185 attention (correct responses) and impulsivity (premature responses).

186 Notably, correct and incorrect responses (in contrast to the other two event types) involve a
187 similar global sensory stimulation (one poke-hole illuminated at the time of responding) and

188 the same motor output (poking), suggesting that the time-locked activity seen only for correct
 189 responses, does likely not reflect sensory or motor aspects. However, we cannot fully exclude
 190 the possibility that deviations in the *local* stimulus - illuminated vs. non-illuminated hole into
 191 which the mouse pokes - may partially account for such differences.



192

Figure 2

193 **Figure 2. Event-locked activity of individual neurons in the variable ITI challenge. (A)**
 194 **Top:** Average z-scored calcium activity in individual ACC neurons time-locked to the onset of
 195 the behavioural event stated above each sub-panel, shown for -4 - +7 s around the event. For
 196 cross-validation, averaging was done across the even trials only and the cells sorted according
 197 to the average peak latency across the odd trials (see also Supplementary Figure 1). **Bottom:**
 198 Pearson correlations between the averaged z-scored activity of the odd and even trials at each
 199 time point. Note that temporal order of peaks is maintained for correct and premature
 200 responses with resulting high correlations, but not for incorrect choices and omissions. $N = 12$
 201 animals and 443 cells. **(B)** Same display and analysis as in (A) but for all neurons recorded in
 202 mPFC. $N = 6$ animals and 229 cells. **(C)** Same data as in (A) but clustered into four clusters

203 (I-IV, indicated on the right) according to their activity from -4 - 7 s around correct responses.
204 Equivalent to (A), clustering was done based on the average of *odd* trials, only, whereas the
205 plot shows the average of the corresponding *even* trials. Activity around incorrect, premature
206 and omitted responses for the same cells is shown in the same order as was determined
207 according to the clustering around correct responses resulting in a lack of emerging temporal
208 patterns. (D) Same analysis as in (C) but for cells in mPFC. Grey lines in (A-D) indicate cells
209 for which a response cannot be shown because the mouse made no incorrect response. (E)
210 Z-scored single-trial calcium activity of exemplary individual cells from each cluster (I-IV,
211 indicated on the right). The timepoint of cue presentation (for correct and incorrect responses
212 only), reward receptacle entry and exit (only correct responses) are shown by the white short
213 vertical lines. The consistent white line represents the timepoint of the choice. (F) Same as in
214 (E), but for cells from the mPFC.

215

216 To further investigate the temporal relationship of neural activity, we k-means-clustered the
217 cells [20] into four clusters according to their average activity in *odd* correct response trials,
218 sorted neurons within each cluster according to time of peak-activity, and then displayed the
219 corresponding average of *even* trials (Figure 2C, D). Qualitatively, this resulted in three
220 clusters with relatively clear activity peaks either before, during, or after correct responses,
221 respectively, in addition to a fourth cluster with increased activity during reward collection only,
222 in ACC (Figure 2C). In contrast, in mPFC a cluster with a well-defined peak at the time of the
223 correct choices was lacking, and the emerging clusters showed activity either before or after
224 the response (Figure 2D). To assess the response-specificity of these temporal patterns, we
225 conducted the same temporal alignment and averaging for the other three response options
226 but sorted the cells according to their order number obtained for clustering by correct
227 responses. For all three response types, this resulted in the loss of clear temporal response
228 patterns, indicating that the temporal relationship cells displayed for correct responses, were
229 largely specific for this one response type (Figure 2C-D). Finally, when plotting the activity in
230 individual trials of one randomly selected neuron for each cluster, the trial-to-trial reliability of
231 activity as time-locked to correct responses was qualitatively confirmed (Figure 2E-F).

232 **Distinct choices are represented by ACC population activity**

233 While the analyses described above confirm that individual ACC and mPFC neurons are
234 modulated by ongoing attention- and impulsivity-related choices or actions, a comprehensive
235 and multi-variate encoding of behaviour is expected only at the level of populations of multiple

236 neurons. To evaluate such behavioural representations in ACC, we performed a decoding
237 analysis, training linear support vector machine (SVM) classifiers to predict the type of
238 behavioural event performed in any given trial based on the population activity, at different
239 time points around the choice event, and during the first 4 s of the preceding ITI (starting with
240 the end of the time-out or of reward consumption). We first focused on binary discrimination
241 between correct and either omitted or premature responses, given that incorrect responses
242 occurred in insufficient numbers for this analysis, in the varITI-challenge in most mice (Figure
243 1E; Supplementary Table 1). To estimate significant predictions, we compared the resulting
244 accuracies with accuracies obtained from classifiers that were trained on data with randomly
245 shuffled labels (paired *t*-tests at each time point with Benjamini-Hochberg correction for
246 multiple comparisons).

247 No appreciable prediction of correct responses (vs. omissions or vs. premature responses)
248 was possible during the ITI, indicating that ACC activity did not reflect, if a mouse *was going*
249 *to* act in a goal-directed, attentive fashion or to avoid task engagement or to act impulsively in
250 an upcoming trial (Figure 3A, left). Although, there was a significant decoding of correct vs.
251 omitted or premature responses at low accuracy of around 60% (vs. 50% chance level),
252 already from at least 4 s before the choice poke onwards, average decoding accuracies only
253 started rising around cue-onset and reached their maximum of >90% only approx. 600 ms
254 *after* the choice-poke. They remained at >90% throughout the time of reward consumption
255 (Figure 3A, right). This indicates that the pre-cue and pre-choice representations were very
256 minor compared to the same representation around and after the choice. These results appear
257 inconsistent with the notion that the primary driver of variance in ACC activity are slowly
258 varying cognitive states of attention or impulsivity, but rather that ACC representations seems
259 to be tightly tied to actions and outcomes. In mPFC, decoding accuracies had a similar
260 temporal trajectory (Figure 3B).

261

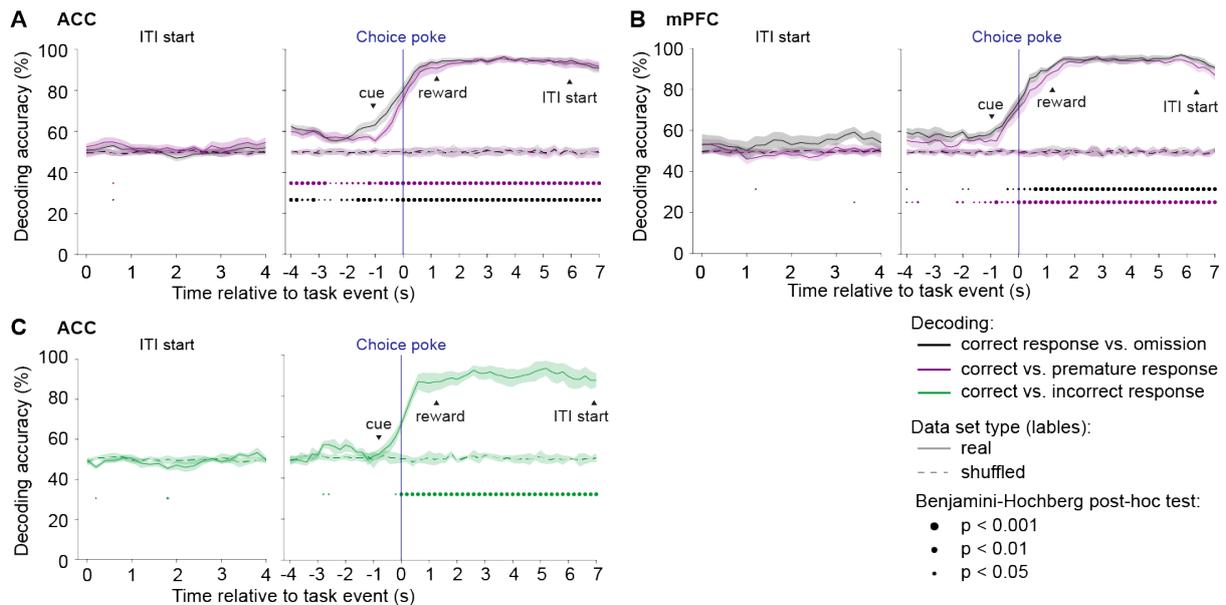


Figure 3

262

263 **Figure 3. Decoding of behavioural choice from population activity in ACC and mPFC.**
 264 (A-B) Cross-validated decoding accuracies derived from binary classification using linear
 265 SVMs calculated from z-scored amplitude values at each 200 ms time bin and predicting
 266 behavioural events from population activity in ACC (A; $N = 11$ mice; one mouse not analyzable
 267 due to low omission rate) or mPFC (B; $N = 6$ mice) in individual varITI sessions; solid lines
 268 and shading represent averages across animals \pm s.e.m., respectively. Decoding accuracies
 269 were first averaged across 100 classifiers calculated on data from each session, and then
 270 across sessions (i.e. animals). Dashed lines indicate results from the same analysis but
 271 performed on control data obtained by random shuffling of event-labels relative to neural
 272 activity data; dots at the bottom indicate a significant difference in the pairwise comparisons
 273 between those two accuracy values at each time point (t -test with Benjamini-Hochberg
 274 correction for multiple comparisons). Binary classification was done differentiating correct
 275 responses against omissions (black) or against premature responses (magenta). Chance level
 276 is 50%. (C) Same analysis as in (A) but classifying *correct vs. incorrect* responses by using
 277 sessions from across all challenge protocols, if more than 5 incorrect responses were made
 278 ($N = 6$). mPFC was not analysed because only 3 sessions had the sufficient number of
 279 incorrect responses. See Supplementary Table 1.

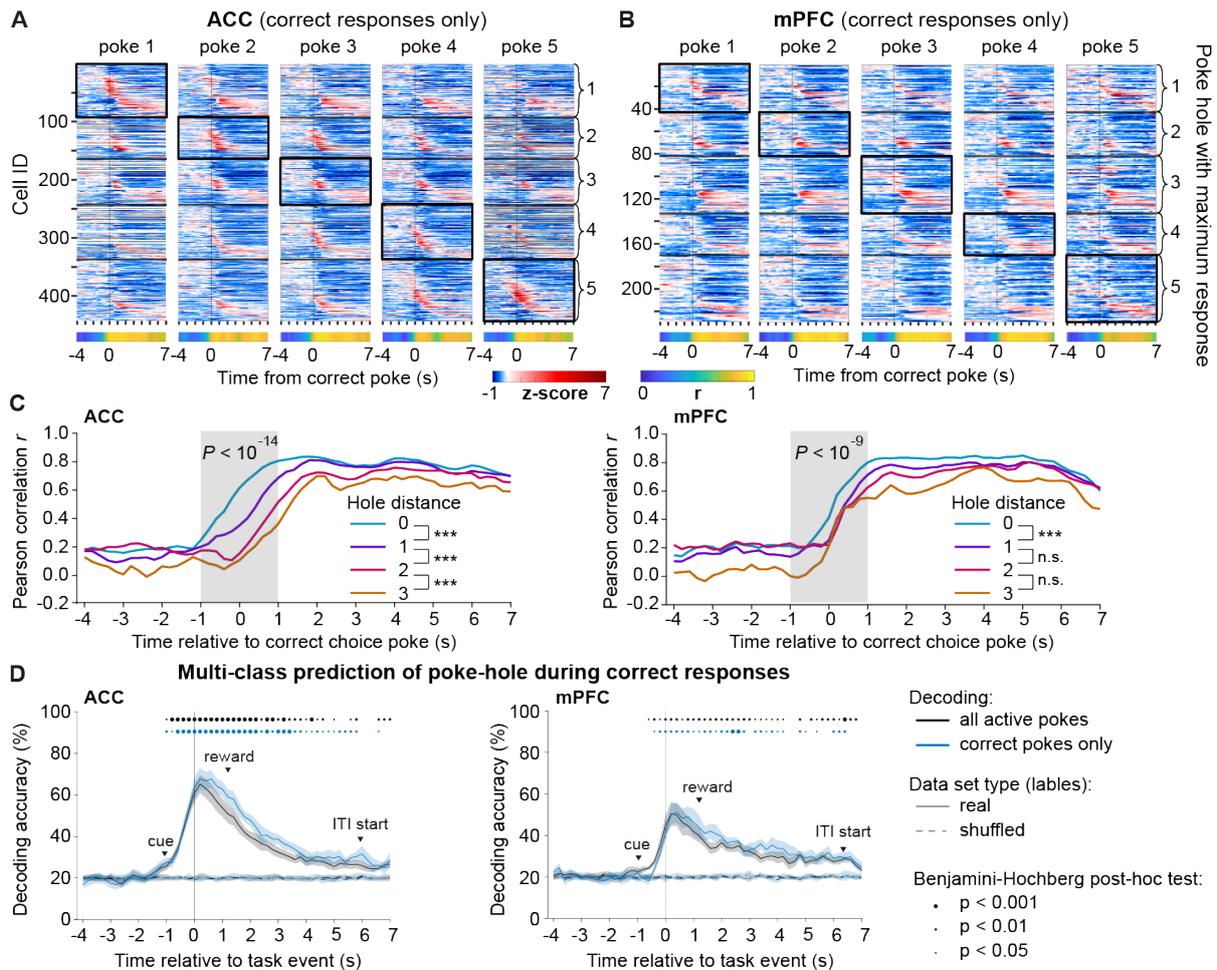
280

281 In the pairwise discriminations described above, a confound by non-choice-related aspects
 282 such as presence of the cue (*correct vs. premature*) or the motor response (*correct vs.*
 283 *omission*) cannot be ruled out. Only *correct* and *incorrect* responses are sufficiently similar in
 284 most parameters and differ mainly in the choice *per se*. To enable a cross-validated decoding
 285 analysis involving *incorrect* responses, we gathered sessions from all six behavioural
 286 challenge conditions given they had a sufficient number of incorrect responses (≥ 6). Using
 287 such data, we found that – in contrast to premature and omitted responses - incorrect choices

288 could be distinguished significantly from correct ones only from the time of the choice-poke
289 onwards (approximately 1 s after cue-onset), indicating a lack of a consistent signature of both
290 preparatory attention and of sensory stimuli in the overall ACC activity (Figure 3C). As seen
291 with the other decoding attempts, high and saturating average prediction accuracies (>80%)
292 could only be reached from around 600 ms after the choice-poke and were maintained
293 throughout reward consumption or its omission (Figure 3C). Overall, the temporal distribution
294 of decoding accuracies is more aligned with the notion that ACC represents ongoing action
295 and its (expected or actual) consequence rather than controlling levels of sustained attention
296 or impulsivity on a trial-by-trial basis.

297 **ACC neuron populations encode spatial aspects of ongoing action**

298 The observation that decoding of response types available in the 5-CSRTT was only possible
299 with high accuracy from the choice-poke onwards, suggests that ACC and mPFC may
300 represent selected actions rather than high-level cognitive states like attention and impulsivity.
301 To further scrutinize this hypothesis, we investigated whether these neurons encode a more
302 fine-grained representation of current action by analysing the responses to each individual
303 choice poke-hole. We aligned the average activity of each neuron in *even* trials to the time of
304 correct choice for each hole individually and sorted the neurons first according to the hole (1-
305 5) which evoked the strongest response during the poke (± 1 s) into five groups, and then
306 sorted by peak latency of the average of *odd* trials within each group. A reproducible pattern
307 emerged for *even* trials that correlated strongly to that of *odd* trials from around the time of
308 choice onwards, in ACC and mPFC (Figure 4A-B, bottom), and which appeared the more
309 dissimilar between poke-holes around the time of poking, the further the holes were apart from
310 each other (Figure 4A-B, top).



311

Figure 4

312 **Figure 4. Activity in the ACC represents spatial action selection.** (A-B) Same data as in
 313 Figure 2A-B (correct responses), but separated by poke-hole and arranged into five groups
 314 (indicated on the right, separated by black lines) based on the hole for which the response of
 315 a neuron had the maximum AUC in the period ± 1 s around the poke (indicated by a black
 316 rectangle around the cluster). *Top*: Average z-scored calcium activity in individual ACC (A) or
 317 mPFC (B) neurons time-locked to the correct choice-poke, shown for -4 - +7 s around the
 318 event. For cross-validation, averaging was done across the *even* trials only and the cells were
 319 sorted according to the average peak latency across the *odd* trials of poke 1. Grey lines
 320 indicate sessions in which the given hole was not poked into. *Bottom*: Pearson correlations
 321 between the averaged z-scored activity of the *odd* and *even* trials at each time point. Note that
 322 the qualitative similarity to the pattern of a given poke gets reduced the further away the poke-
 323 hole is, especially around the time of poking. Supplementary Figure 2A-B shows the same
 324 data without prior sorting into clusters. (C) Based on the data shown in (A-B), Pearson
 325 correlations between response patterns of pairs of poke holes, coded in colour according to
 326 the distance between the holes; correlation values for hole-combinations with the same
 327 distance (e.g. 1-4 and 2-5 for distance 3) were averaged; see Supplementary Figure 2C for
 328 the individual correlation values of each hole-pair. A repeated-measures ANOVA was
 329 calculated across the population of 11 observations in the time period ± 1 s around the poke
 330 (grey bar; P -value for main effect of hole-distance indicated at the top); results from paired
 331 Sidak post-hoc tests are indicated for adjacent hole distances in the colour-legend of each
 332 sub-panel. *** $P < 0.001$; n.s., $P > 0.5$. (D) Accuracies of the decoding of the identity of the
 333 poke-hole, either for all active responses (correct, incorrect, premature; black) or for correct
 334 responses only (blue) aligned to the time of poking (0, vertical line); average latency to cue

335 onset (response latency), reward-poke entry (reward latency) and start of the next ITI (reward
336 consumption time, indicated only for correct trials) are indicated by arrowheads. Decoding
337 accuracies were first averaged across 100 classifiers calculated on data from each session,
338 and then across sessions (i.e. animals). Dashed lines indicate results from the same analysis
339 but performed on control data obtained by random shuffling of event-labels relative to neural
340 activity data; dots at the top indicate significant pairwise comparisons between those two
341 accuracy values (*t*-test after correcting for multiple comparisons across time bins). Shaded
342 area, s.e.m. across mice.

343

344 To assess this quantitatively, we calculated Pearson correlations between such population
345 activity patterns for all pairs of poke-holes and averaged them across hole-pairs with the same
346 distance (Figure 4C; Supplementary Figure 2). Indeed, within approximately ± 1 s around the
347 choice-poke, correlations were the higher the closer the holes of the pair were to each other.
348 After the choice, in contrast, correlations were consistently high, irrespective of distance. This
349 suggested that ACC activity displays a certain similarity related to spatial proximity of poke
350 holes before the poke, while being dominated by non-spatial aspects of the choice after it is
351 made (Figure 4A-C).

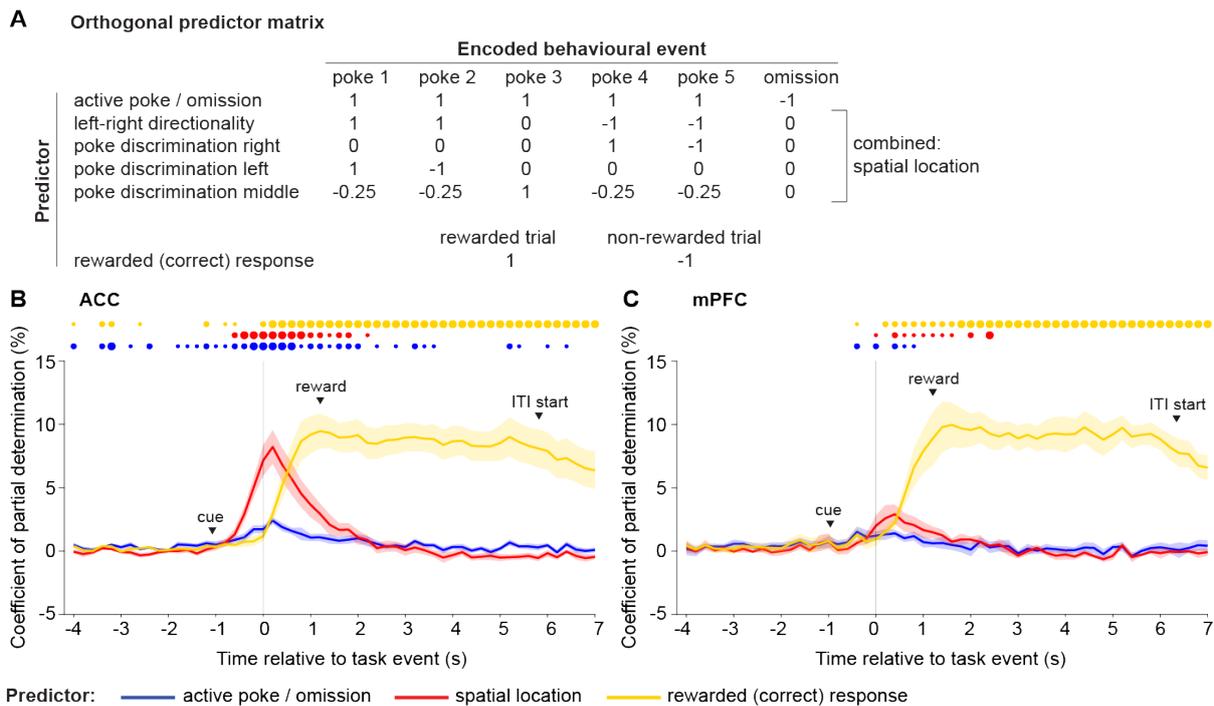
352 To further investigate this early spatial selectivity, we trained multi-class SVM classifiers to
353 decode the identity of the poke-hole. Surprisingly, this identity could be predicted from about
354 1 s before the poke (just after cue-onset) onwards, reaching an average peak accuracy of
355 close to 70% in ACC and close to 55% in mPFC (against a chance level of 20%) approximately
356 200 ms after the poke. In contrast to the representation of event-type (Figure 3A-C), average
357 accuracy decreased again immediately after the poke, suggesting that the representation of
358 the precise action fades after its occurrence (Figure 4D). Interestingly, the decoding accuracy
359 was virtually identical, when performing the same analysis on pokes of all three active
360 response types (correct, incorrect, premature) *combined*, suggesting the existence of a
361 representation of action that is independent from the representation of response-type and
362 outcome.

363 **Independent encoding of action and outcome in ACC**

364 The hypothesis of an independent encoding of motor action (poke hole identity) and choice or
365 outcome (event-type) entails the question to what extent each factor determines neural activity
366 in the ACC, at every time point. To answer this question, we trained linear regression models
367 predicting the activity of each neuron using predictors coding the active poking (vs. omissions),
368 rewarded choice (correct responses vs. absence of reward due to erroneous choices), and
369 poke-hole identity (spatial location; encoded by four predictors, see predictor matrix in Figure
370 5A). We ran a separate regression for each time-point in the aligned activity and calculated
371 the cross-validated coefficient of partial determination (CPD) for each predictor at each time-
372 point, which quantifies the share of the variation in the neural activity that is uniquely explained
373 by that predictor. To estimate statistical significance for a given predictor at a given timepoint,
374 the distribution of CPD values across subjects for that predictor and timepoint was compared
375 to 0% [21].

376 From the average onset of cue presentation (1 s before the poke), matching the time course
377 of the decoding analysis (Figure 4D), the spatial identity of the poke-hole started to gain ever
378 more influence over ACC-activity, and dominated it compared to the other predictors from
379 approximately 600 ms before until 400 ms after the choice poke (Figure 5B). This effect was
380 mostly driven by an encoding of the left (hole 1-2) vs. the right (hole 4-5) side of the 5-choice
381 wall, although most other tested predictors of the selected action (poke discrimination left,
382 right, and middle) and of active responding in general (vs. omission) displayed significant
383 CPDs during and after the time of poking, as well (Supplementary Fig. 3). From 600 ms after
384 the poke onwards, however, the factor of rewarded (correct vs. erroneous) response had the
385 single strongest influence on neural activity out of the tested predictors (Figure 5B). This
386 overall pattern suggests, that AAC *simultaneously* encodes fine-grained selected action and
387 a high-level representation of both active and correct responding from the time of choice-poke
388 onwards for almost 2 s, but action representation dominates around the time of execution
389 whereas high-level representation dominates subsequently (Figure 5B). Given that the

390 influence of outcome starts rising immediately after the choice poke - and hence more than a
 391 second before actual reward consumption - this activity reflects *expected* (vs. omitted) reward
 392 at this early stage, but it might later on reflect actual outcome, as previously shown [11]. mPFC
 393 neurons, in contrast, also encoded outcome (from 200 ms after the poke) but lacked a
 394 consistent encoding of motor action, especially before the poke (Figure 5C).



395

Figure 5

396 **Figure 5. Encoding of poking action and reward in the population activity of the ACC**
 397 **and mPFC.** (A) Orthogonal predictor matrix designed to indicate the representation of the
 398 poke (1: poke in either hole, 0: omission), the poke directionality (1: poke in left holes 1 or 2; -
 399 1: poke in right holes 4 or 5; 0: poke in middle hole 3 or omission), right poke discrimination
 400 (1: poke in hole 4; -1: poke in hole 5; 0: poke in holes 1,2 and 3 and omission), left poke
 401 discrimination (1: poke in hole 1; -1: poke in hole 2; 0: poke in holes 3,4 and 5 and omission),
 402 middle poke discrimination (1: poke in hole 3; -0.25: poke in holes 1,2,4,5; 0: omission) and
 403 reward (1: rewarded; 0: not rewarded). The four predictors representing the spatial location of
 404 the poke hole have been removed at once from the regression model in order to quantify the
 405 encoding of motor action. (B-C) Coefficient of partial determination (CPD) averaged across
 406 cells recorded in ACC (B, $N = 11$ mice) and mPFC (C; $N = 6$ mice). Time bins where CPDs for
 407 a given event were significantly higher than zero after cross-validated linear regression are
 408 indicated with a dot at the top of each panel, colour-coded for the respective predictor (one
 409 sample t -test with Benjamini-Hochberg post-hoc correction). CPDs were determined for each
 410 event-type by subtracting the sum squared errors of the full linear regression model
 411 (incorporating every event type as predictor) from the sum squared error of the reduced
 412 regression model where one predictor (corresponding to the event type for which the
 413 population activity should be explained) was removed.

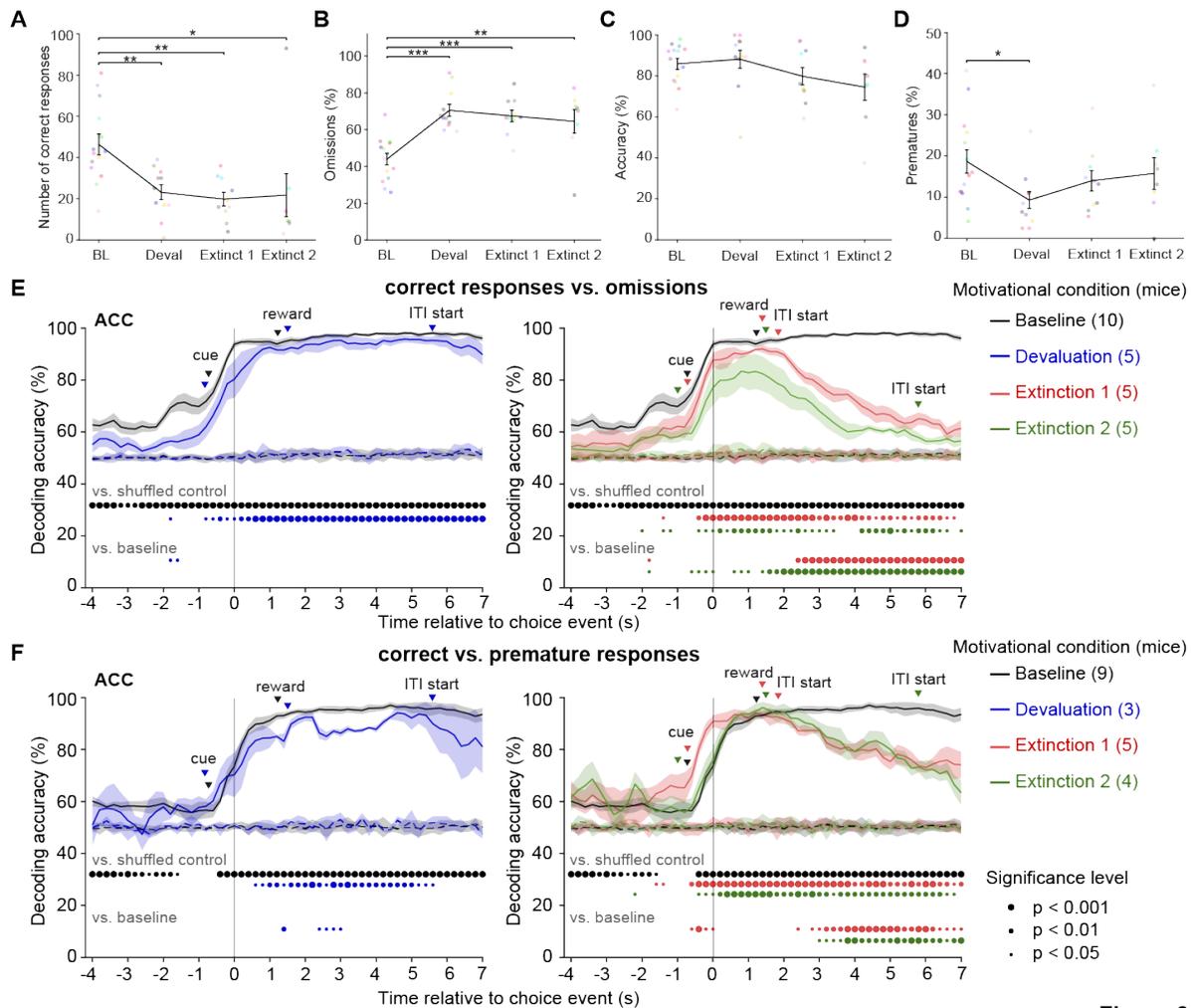
414

415 **Devaluation and omission of reward alter representation of choices**

416 The prominent representation of *correct* – and, hence, *rewarded* – responses in ACC,
417 particularly during and after the response, which is also in line with earlier studies [9,10],
418 suggests that these activities may represent the expectation of reward rather than elevated,
419 preparatory attention. This is also supported by the fact that incorrect and correct responses
420 cannot be discriminated from population activity *before* the choice-poke is made (Figure 3C).
421 To assess this conclusion further, we conducted four further 5-CSRTT experiments where the
422 value or expectancy of the reward was altered, using the combined 0.8s/7s ITI challenge (to
423 obtain more incorrect responses, Figure 1E), at the end of the sequence of tests: first we
424 recorded a baseline session with normal food-deprivation and reward delivery. In a second
425 session, the reward was *devalued* by pre-feeding (by providing 6 g of food overnight and 2 ml
426 of milk reward 1 h before session start). In the third and fourth recording sessions the food-
427 deprivation (i.e. value of reward) was normal, but the delivery of reward was omitted (*extinction*
428 1 and 2). Normal training sessions in the baseline protocol were conducted after the first two
429 sessions, but not between the extinction sessions.

430 At the behavioural level, both devaluation and extinction caused a significant decrease of the
431 number of correct responses, driven by an increase in omissions, whereas accuracy remained
432 relatively stable ($P < 0.05$, Dunnett's post-hoc test, performed after significant main effect of
433 condition in RM-ANOVA; $N = 15$; Figure 6A-C). During devaluation, also active *erroneous*
434 (incorrect and premature) responses decreased, and reward latency increased, in line with a
435 reduced motivation (Figure 6D; Supplementary Figure 4); such effects appeared qualitatively
436 also during extinction sessions, but mostly without reaching significance. Response latencies
437 were not significantly altered indicating unperturbed responsiveness in any of the conditions
438 (Supplementary Figure 4).

439



440

Figure 6

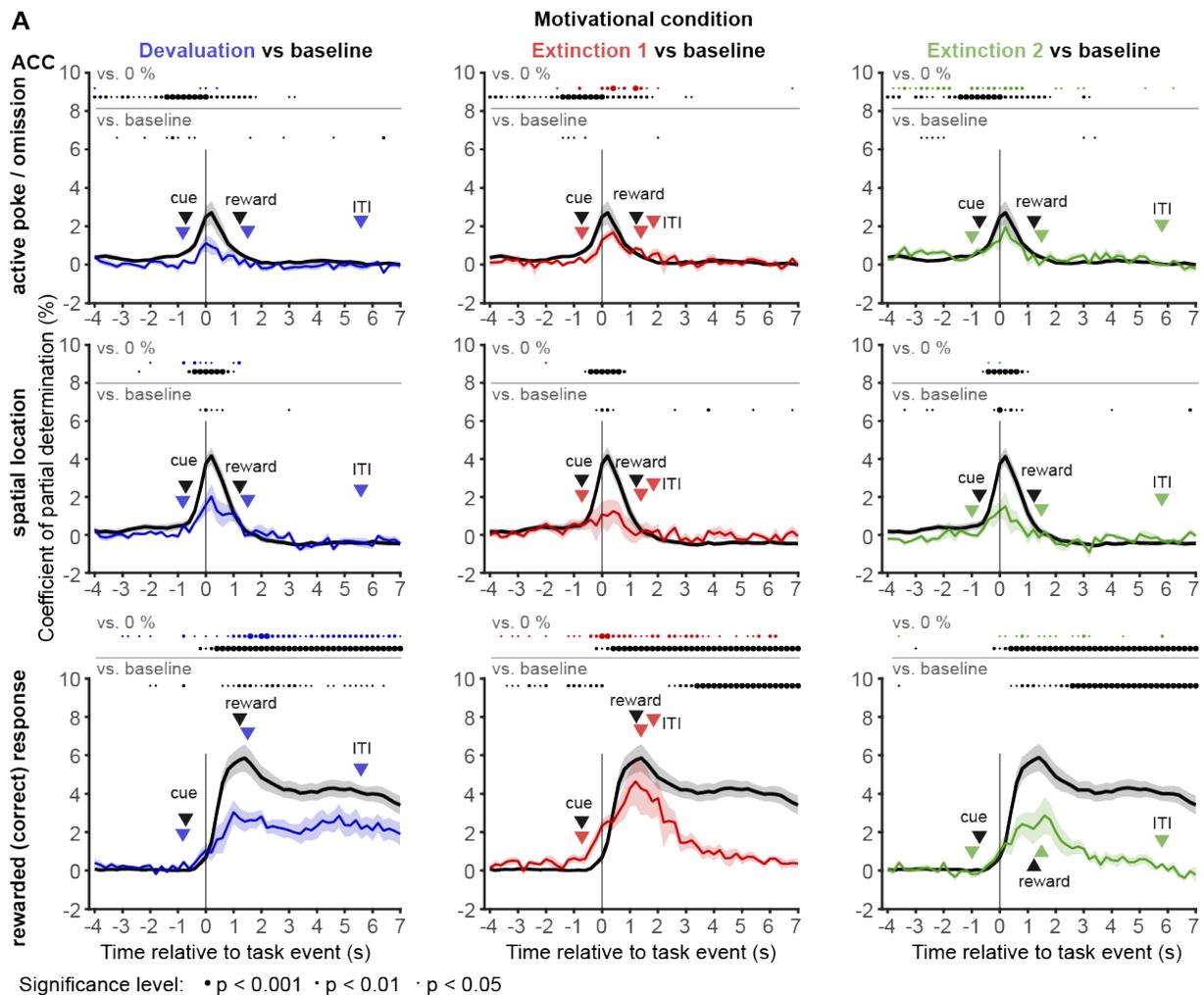
441 **Figure 6. Decoding of behavioural choice from population activity in ACC during reward**
 442 **devaluation and extinction experiments. (A-D)** Measures of task engagement and
 443 performance in the 5-CSRTT (as indicated on y-axes) during training sessions in the 0.8s-
 444 SD/7s-ITI combined challenge at normal reward conditions (baseline, BL), after devaluation
 445 of reward (Deval), or with omission of reward (extinction; Extinct 1/2), as indicated on x-axes
 446 with tethered miniscope. Dots indicate individual animals coded by colour, bars show mean \pm
 447 s.e.m.. Asterisks represent Dunnett pairwise post-hoc test comparing BL condition against the
 448 other conditions after RM-ANOVA. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. See Supplementary
 449 Figure 4 for further behavioural measures. **(E-F)** Cross-validated decoding accuracies derived
 450 from binary classification using linear SVMs calculated at each 200 ms time bin and predicting
 451 correct vs omitted (E) or vs. premature (F) responses from population activity in ACC during
 452 test sessions in the 0.8s-SD/7s-ITI combined challenge with normal reward conditions
 453 (baseline, black), after devaluation of reward (blue) or with omission of reward on two
 454 consecutive test sessions (extinction 1, red; extinction 2, orange). Whereas 10 mice
 455 participated in these test sessions, actual N-numbers for the analyses (stated in figure legends
 456 on the right and in Supplementary Table 1) vary mainly due to mice that did not perform
 457 sufficient numbers of correct or premature responses, in rare cases also due to technical
 458 failures. Solid lines and shading represent averages across animals \pm s.e.m., respectively.
 459 Decoding accuracies were first averaged across 100 classifiers calculated on data from each
 460 session, and then across sessions (i.e. animals). Dashed lines indicate results from the same
 461 analysis but performed on control data obtained by random shuffling of event-labels relative
 462 to neural activity data. Dots below the traces of each panel indicate time bin and alignment

463 (coded by colour) where decoding accuracies from these two classifiers differed significantly
464 (paired *t*-test with Benjamini-Hochberg correction; coded by colour for each reward condition).
465 Dots at the bottom represent Dunn-Sidak post-hoc tests comparing accuracy values from
466 devaluation or extinction sessions (coded by colour) to those from the baseline condition
467 (conducted after significant effect of reward condition in the repeated-measures ANOVA).
468 Significance level is encoded by dot size. Triangles indicate average latency to cue-
469 presentation, reward receptacle entry and the start of the next ITI (the latter was often close
470 after reward receptacle entry in extinction conditions as no reward was provided). Chance
471 level is 50%. Decoding of incorrect responses could not be performed due to their small
472 numbers in most of the respective sessions.

473

474 To evaluate if the change of the value or contingency of the reward affected the representation
475 of rewarded responses, we repeated the time-resolved binary decoding of correct vs.
476 premature or omitted responses, as conducted previously for the varITI-challenge (Figure 3),
477 in each of the four conditions. Whereas, under normal reward conditions, the trajectories
478 looked like those found before, with increasing representation of the choice and its outcome
479 with the cue-onset, somewhat altered decoding accuracies were found in the other conditions
480 (Figure 6E-F). Generally, decoding accuracy was lower in the conditions with altered reward
481 value or occurrence, as indicated by comparisons to the accuracy achieved by classifiers
482 trained with shuffled control datasets. Significant decoding of correct responses was hardly
483 possible *before* they actually occurred, possibly reflecting a certain lack of representation of
484 preparatory attentional or impulse control required for such responses (Figure 6E-F). When
485 comparing the accuracy achieved under baseline condition with that achieved under reward
486 devaluation (RM-ANOVA with pairwise Dunnett post-hoc tests), these two conditions rarely
487 differed, suggesting that the value of the reward has relatively little influence on the
488 discriminability of representations of rewarded vs. non-rewarded choices (Figure 6E-F). In
489 contrast, during both *extinction sessions*, the discrimination of correct responses vs. omissions
490 and, to a lesser extent, vs. premature responses, deteriorated from about 2 s after the choice
491 poke onwards and differed significantly from the accuracy achieved in the baseline condition,
492 in line with the lack of a reward (Figure 6E-F). This implies that beyond this time point, ACC
493 largely encodes the *actual* outcome, which is no longer distinct between the choice options in
494 the extinction sessions. Similarly, in the *second* extinction session, accuracy for decoding

495 correct responses vs. omissions differed already from 600 ms *before* the choice poke onwards
 496 and on most subsequent time points (Figure 6E). This supports the idea that ACC partly
 497 represents *expected* outcome, which has been learned to be identical between these two
 498 choice options by the time of the second extinction session.



499

Figure 7

500 **Figure 7. Encoding of reward and action in ACC depends on the relative value and**
 501 **presence of reward. (A)** Coefficient of partial determination (CPD) averaged across cells
 502 recorded in ACC in the four different reward-related conditions encoded by colour (number of
 503 animals): black, baseline (10); blue, devaluation (6); red, extinction session 1 (6); green,
 504 extinction session 2 (6). As in Figure 5, CPDs were determined for each predictor, stated on
 505 the left (rows) by subtracting the sum squared errors of the full linear regression model
 506 (incorporating every event type as predictor) from the sum squared error of the reduced
 507 regression model where one predictor was removed. To evaluate the encoding of spatial
 508 location as such, all four predictors reflecting spatial location were combined by removing
 509 them at once (see Figure 5A). Time bins where CPDs for a given event were significantly
 510 higher than zero after cross-validated linear regression are indicated with a dot at the top of
 511 each panel, colour-coded for the respective reward condition (one sample *t*-test with
 512 Benjamini-Hochberg post-hoc correction). Dots below represent Sidak tests comparing CPD

513 values from devaluation or extinction sessions (coded by colour) to those from the baseline
514 condition (conducted after significant effect of reward condition or time-reward interaction in
515 RM-ANOVA). Significance level is encoded by dot size. Triangles indicate average latency to
516 cue-presentation, reward receptacle entry and the start of the next ITI (the latter was often
517 close after reward receptacle entry in extinction conditions as no reward was provided). To
518 ensure, that CPDs in the baseline session are not simply higher because of a larger number
519 of trials or responses used for their calculation, the number of baseline-trials was down-
520 sampled to roughly match those of the reward condition it is compared to in each panel for
521 each poke-hole and response type.

522

523 To further elucidate the hypothesis that ACC encodes *expected* and, subsequently, *actual*
524 outcome, we performed an encoding analysis, as previously done for the varITI-challenge
525 (Figure 4), for the data obtained from these four conditions. Across time intervals, we
526 compared CPD values to the control value of 0% for all conditions (*t*-test with Benjamini-
527 Hochberg correction) and we compared the devaluation and extinction conditions to baseline
528 (RM-ANOVA followed by Sidak post-hoc test; Figure 7A). As expected, *extinction* – especially
529 when repeated – led to a reduction of the representation of reward. However, also *devaluation*
530 ensued a strong decrease of reward representation compared to baseline, suggesting that
531 reward determines ACC activity the stronger the higher its value is (a result that the prior
532 decoding analysis could not reveal, see Figure 6E-F). Strikingly, both devaluation and
533 extinction also led to a virtual loss of the spatial representation of the poke-hole: in all three
534 conditions, CPD values for the combined spatial predictors were significantly lower than those
535 in the baseline condition around the time of poking and, in contrast to the baseline, rarely
536 exceeded 0% (Figure 7A). This suggests that the representation of specific actions in ACC
537 depend on their expected (reward) value.

538

539 **Discussion**

540 Using miniscope recordings during the 5-CSRTT in mice, we here demonstrated that
541 excitatory neurons of the posterior ACC represent the choices available in this complex task,
542 both at the high level of response options that depend on *cognitive state* and at the low level

543 of the selected *action* (poke-hole identity). We found that coding at both levels happens
544 simultaneously, but that ACC activity is dominated by the selected action (graded spatial
545 representation of poke-hole location) during the execution of the choice poke, and by outcome
546 expectation from about 400 ms after the choice, followed by encoding of actual outcome, once
547 obtained. Whereas action representation faded, the high-level representation of the chosen
548 option (or its consequence) remained stable for at least 7 s after the choice, but deteriorated
549 much faster when reward was omitted during extinction sessions. When value or occurrence
550 of reward were changed, the share of cells selectively *activated* by reward and of cells
551 selectively *inhibited* during omissions – i.e., the share of cells representing those options
552 whose relative utility changed – increased. This suggests that ACC represents action-outcome
553 contingency rather than outcome itself. Likewise, the representation of rewarded (correct)
554 responses, but also of low-level spatial action parameters decreased when reward was either
555 devalued or omitted. This underlines that motor and reward representations in ACC depend
556 on the expectation and value of a choice's outcome.

557 Importantly, significant preparatory network activity during the ITI and before the cue which
558 could indicate a general level of high or low attention, impulsivity, or task engagement in a
559 given trial was either not found (attention-related activity discriminating correct from incorrect
560 choices) or was rather weak (discrimination between impulsive responses or omissions and
561 correct responses). In fact, all conducted analyses assessing response-type coding suggest
562 that choices are represented in ACC mainly *during* and *after* their time of occurrence, and that
563 *rewarded* responses are represented most reliably, as seen in other studies [12,13].

564 How can these observations be reconciled with earlier reports of signatures of attention and
565 impulsivity in ACC [9–11,22] and, most importantly, with the well-documented ability to
566 modulate these cognitive states by manipulation of ACC neurons [5,6,8,23]? Firstly, with
567 respect to pre-choice activity, the discrimination of correct responses from either premature or
568 omitted responses *was* actually possible, albeit at low (~60%) accuracy, already during at
569 least 4 s before the choice poke (but not at the beginning of the ITI; Figure 3, 6). Furthermore,

570 correct and premature responses were accompanied by time-locked activity of neurons in the
571 same pre-choice time period (Figure 2). This indicates a representation of the state of impulse
572 control that exists, albeit weakly, well before the choice, though not at the beginning of the ITI
573 – in line with ramping activity before premature responses observed in rats [10]. In contrast,
574 incorrect responses could only be distinguished from correct responses at time of occurrence
575 (Figure 3) and were not accompanied by time-locked activity before (Figure 2). This suggests
576 that the trial-to-trial variation of sustained attention is not represented in excitatory cells of
577 posterior ACC. The divergence between both phenomena is in line with the fact, that
578 impulsivity, but not accuracy, was modulated by chemogenetic manipulation of such neurons
579 in ITI-challenges [5]. Since modulation of ACC parvalbumin-interneurons (which we did not
580 record in this study) could alter both aspects of executive function [8], they could constitute a
581 locus of the attentional component in this task [14]. Thirdly, chemogenetic or pharmacological
582 modulation of ACC could exert its effect anatomically and temporally more globally, rather
583 than controlling behaviour directly and in individual trials. The ACC might be acting through
584 the demonstrated strong representation of obtained reward, errors, and action-outcome
585 contingencies, to shape tendencies of attentional and impulse control in upcoming trials
586 through other circuits, as shown for a subpopulation of ACC neurons already [6,7]. Fourthly,
587 and alternatively to the previous scenario, ACC modulation may influence the occurrence of
588 distinct response-types because of its role in action selection [21,24], as we suggested based
589 on our [5,8] and other [6] previous chemogenetic and optogenetic [7] data before [8]. In this
590 scenario, elevated activity of certain ACC pyramidal neurons triggers the response into a
591 certain poke-hole (in line with the early encoding of action found in this study). This could be
592 caused by the ACC due to reward expectation leading to strong excitatory AAC activity
593 entailing a correct response. But it could also be caused erroneously by higher order inputs
594 received by the ACC without the appropriate cue leading to incorrect or premature responses,
595 accompanied by somewhat less effective, less time-locked ACC excitation. Activation of PV-
596 interneurons [8] or direct partial inhibition of excitatory neurons in ACC [5] could increase the
597 threshold needed for such activation that ultimately triggers the selected response, so that the

598 strong activation triggering correct responses remains supra-threshold, whereas the weaker
599 activity which could otherwise cause erroneous responses falls sub-threshold. Further
600 physiological investigations are required to probe these mechanistic models.

601 Regarding mPFC, the number of mice we recorded in our study for comparative purposes is
602 too small to draw general conclusions regarding the coding in this region, which has also been
603 examined in other studies in the 5-CSRTT [10,14,26,27]. Nevertheless, the relatively minor
604 differences between representations in ACC and mPFC we observed are consistent with an
605 earlier finding of oscillatory coupling of both regions during choice events in the 5-CSRTT [22]
606 and, more generally, with a model of widely distributed encoding of behaviour across
607 neocortex [16]. Population activity in mPFC also represented choice options but less reliably
608 and partly shifted to later time points, i.e., after the choice, and again with a bias to encoding
609 rewarded choices more than non-rewarded ones. Also, a fine-grained encoding of spatial
610 aspects of a selected action was virtually absent. This is in line with previous
611 electrophysiological measurements in rats during this task, indicating that a majority of
612 responsive mPFC cells respond after the choice, representing trial outcome, and cells show
613 considerably stronger firing rate increases during correct than during premature responses
614 [10].

615 In conclusion, our temporally resolved analysis suggests that ACC excitatory neurons
616 represent a chosen action as it is made as well as its expected and actual outcome. Our results
617 uncover parallel encoding of fine-grained spatial parameters of selected actions - in
618 dependence on their outcome value - and of action-outcome contingency in ACC, and suggest
619 that trial-by-trial encoding of high-level cognitive states before the choice is either minimal (for
620 task engagement and impulsivity) or absent (for attention).

621

622

623 **Methods**

624 **Animals**

625 In total, 46 male C57BL/6J wildtype mice, were used for this study. Animals were group- or
626 single-housed in Type II-Long individually ventilated cages (Greenline, Tecniplast, G),
627 enriched with sawdust, sizzle-nestTM, and cardboard houses (Datesand, UK), and maintained
628 at a 13 h light / 11 h dark cycle. Water was available *ad libitum*. All experiments were
629 performed in accordance to the German Animal Rights Law (Tierschutzgesetz) 2013 and were
630 approved by the Federal Ethical Review Committee (Regierungspräsidium Tübingen) of
631 Baden-Württemberg, Germany (licence number TV1469).

632 **5-CSRTT training and testing with calcium imaging**

633 Mice started training in the 5-CSRTT at 3-5 mo of age and were kept under food-restriction at
634 85-95% of their prior average free-feeding weight which was measured over 3 days
635 immediately prior to the start of food restriction at the start of the behavioural training. Testing
636 was conducted in operant chambers placed individually in melamine-MDF sound-insulated
637 and ventilated outer boxes and fitted internally with an array of five nose-poke holes on one
638 wall and a reward receptacle on the opposite wall. All six apertures could be illuminated to
639 instruct the entry into them and were fitted with IR break-beams to detect entry and exist of
640 the animal's snout. All experiments were conducted in custom-made trapezoidal chambers
641 based on the pyControl system [19,28] (<https://pycontrol.readthedocs.io>).

642 The 5CSRTT training protocol was similar to what we previously described [5]. In brief, after
643 initiation of food-restriction, mice were accustomed to consume the reward (strawberry milk,
644 MüllermilchTM, G) first in their home cage, and then in the operant box (2-3 exposures each).
645 Subsequently, mice were trained in 2-13 sessions (30 min, once daily) of habituation training.
646 In each trial, all holes of the 5-poke wall were illuminated for an unlimited time and the mouse
647 could poke into any one of them to earn a 40 µl milk reward subsequently disposed from the
648 illuminated receptacle. If mice attained at least 30 rewards each in two consecutive sessions
649 or (in exceptional cases) had reached the 16th session of habituation training, they were moved

650 to the 5-CSRTT training, during which mice transitioned through five stages of increasing
651 difficulty, based on reaching certain performance criteria in each stage, as described
652 previously[5]. The difficulty was determined by the length of time the stimulus was presented
653 (stimulus duration, SD) and the length of waiting time between the end of the previous trial
654 and the stimulus presentation of the next trial (inter-trial-interval, ITI). In case a reward was
655 collected on the previous trial, the ITI was initiated by the removal of the snout of the animal
656 from the reward receptacle. In all 5-CSRTT protocols only one pseudo-randomly selected
657 aperture of the 5-choice wall was lit up after the ITI, indicating that this hole needs to be poked
658 into (*correct response*) in order to earn a 20 μ l milk reward (Figure 1A). Trials were not
659 rewarded but instead terminated immediately with a 5 s time-out period during which the
660 house light was turned off, if the animals either poked into any hole during the ITI (*premature*
661 *response*), poked into a non-illuminated hole (*incorrect response*) during the SD and limited-
662 hold time (LH, until 2 s after SD), or failed to poke throughout the trial (*omission*). The relative
663 numbers of such response types were used as performance indicators measuring premature
664 responding [%premature = $100 \times (\text{number of premature responses}) / (\text{number of trials})$],
665 sustained attention [accuracy = $100 \times (\text{number of correct responses}) / (\text{number of correct and}$
666 $\text{incorrect responses combined})$], and lack of participation [%omissions = $100 \times (\text{number of}$
667 $\text{omissions}) / (\text{number of trials})$]. A trial was considered to start at the beginning of the ITI, i.e.
668 included premature responses. Additionally, the time required to poke into the indicated hole
669 after it was illuminated (response latency) and the time from the exit from the correct hole until
670 the entry into the reward receptacle (reward latency) were measured, whereby the latter is
671 usually used as a compound indicator of motivation and locomotor drive[3]. In all stages,
672 sessions lasted 30 min and were performed once daily at the same time of day and in the
673 same box for each animal.

674 After surgery (see below), animals were trained until they had reached the final baseline stage
675 (BL; 2 s SD, 5 s ITI) obtaining an accuracy >80% and an omission rate <50% in two
676 consecutive sessions. For the last 5 d before the first imaging session, mice were accustomed

677 to be gently fixed in the experimenter's hand for the fixation of the miniscope before the
678 session and were trained with dummy 'miniscopes' that were equal in height and weight,
679 attached to the baseplate. Training with dummy miniscopes lasted until an accuracy >80%
680 and an omission rate <50% in two consecutive sessions was reached again. On subsequent
681 days, mice were trained for 3 d with an actual, tethered miniscope in distinct operant chambers
682 set up for simultaneous imaging, followed by the first imaging session conducted also in the
683 baseline stage. Subsequently, the imaging sessions (30 min) were repeated with different
684 challenge conditions in the same order for every mouse (see Figure 1E). Some imaging
685 sessions needed to be repeated due to technical failures. In between imaging sessions, mice
686 were trained in the same testing chambers in the baseline stage with the miniscope attached.
687 After imaging sessions with the challenge protocols were completed, some mice underwent
688 a separate set of pharmacological experiments in the 5-CSRTT (data not shown in this
689 manuscript), after which training in the baseline protocol and then further imaging sessions in
690 the combined 0.8s-SD/7s-ITI challenge protocol with concomitant manipulation of value or
691 occurrence of reward followed: Firstly, imaging was conducted under normal conditions of
692 food-restriction and reward-delivery (baseline), secondly imaging was conducted after
693 devaluation of reward by pre-feeding (providing 6 g of food overnight and 2 ml of milk reward
694 1 h before session start), thirdly, two sessions followed under conditions of normal food-
695 restriction but omission of reward (extinction). Training in the baseline protocol without imaging
696 was conducted before and after the devaluation session, but not between the extinction
697 sessions.

698 **Surgical procedures**

699 After the mice reached at least stage 4 (4 s SD, 5 s ITI) of the 5-CSRTT, animals were
700 anaesthetized using isoflurane (AbbVie, G), received subcutaneous injections of analgesics
701 (0.08 mg/kg buprenorphine, Bayer, G; 1 mg/kg meloxicam, Boehringer Ingelheim, G), and
702 local scalp anaesthesia (200 μ l of 0.025 % bupivacaine, AstraZeneca, UK) before placement
703 in a stereotaxic frame (Kopf, US; manual digital frame, World Precision Instruments, US) with

704 non-rupture mouse ear bars. The body temperature was stabilized using a feedback-
705 controlled heating blanket (Harvard Apparatus, US) and the anaesthesia was maintained with
706 1.5 % isoflurane. The following stereotaxic coordinates (from bregma) and volumes were used
707 for bilateral transfection of the stated areas; ACC: injection at AP +1.0, ML 0.4, DV 1.3 (300 nl)
708 and 1.7 (200 nl); ventral mPFC: injection at AP +2.2, ML 0.3, DV 2.2 (200 nl). An AAV5-
709 CamKII α -GCaMP6m vector suspension (6.2×10^{12} vg/ml; University of Zürich viral vector
710 facility; UZH-VVF, CH) was diluted down 1:1 to a final titre of 3.1×10^{12} vg/ml in 5 % sorbitol/PBS
711 (Sigma, G) and infused using a 10 μ l precision syringe (WPI, US) at an infusion rate of
712 100 nl/min. To minimize backflow of the virus, the needle was kept in place for 5 min at each
713 site after infusion, and additionally for another 5 min 0.1 mm above the last infusion site.
714 Subsequently, the wound was sutured, the mouse was allowed to recover in a temperature-
715 controlled chamber at 36°C, and provided with mesh-food, gel-food and daily post-operative
716 monitoring for 7 d, including application of meloxicam (Metacam, 1 mg/kg, Boehringer
717 Ingelheim, G) on the first 3 d. The mice were kept on *ad libitum* food.

718 Approximately, one week after the injection of the viral construct, a gradient refractory index
719 (GRIN) lens (Inscopix, CA, USA) was implanted. The surgery initially followed the steps
720 described above for the virus injection and was followed by two craniotomies into the occipital
721 and parietal bone where a screw (1 mm diameter, Precision Technologies, GB) was placed
722 into each hole for later implant stability. A craniotomy for the lens (1 or 0.5 mm in diameter for
723 ACC or mPFC, respectively) was made above the original infusion site. Before lowering the
724 lens into the brain tissue, the skull was dried for better glue attachment and the lens was
725 cleaned with 70% ethanol or 100% isopropanol. From the brain surface, the lens was held
726 vertically by a pipette tip with negative internal pressure created by a vacuum pump and
727 lowered by 20 μ m every 30 s into the brain tissue at the original infusion site until reaching a
728 depth of 1 mm (ACC). For mPFC, a custom-made GRIN-lens injector (“GRINjector”) was used
729 placing the lens at 2 mm from the brain surface. Super-glue (Loctite 401, Henkel, DE) was
730 applied to attach the lens to the skull, followed by light-curable dental adhesive (Breeze™,

731 Pentron, US). Dental cement was applied on the exposed skull with approx. 1 mm of the lens
732 protruding out from the skull. Kwik-Sil™ (WPI, US) was applied on the lens to protect it from
733 later mechanical damage. Post-operative care was applied as after the first surgery.

734 Between 2-6 wks after lens implantation, the GCaMP6m expression was checked. For this,
735 the mouse was anaesthetized and fixed to the stereotaxic frame as described above. After
736 removing the Kwik-Sil™ cone, the lens surface was wiped with lens tissue soaked in 70%
737 ethanol or 100% isopropanol. The baseplate was attached to the miniscope and to the
738 stereotactic frame using a clamp. Depending on the quality and quantity of GCaMP6m-
739 expressing cells, the baseplate was persistently fixated to the implant by applying light-curable
740 dental cement (Flow-It™, Pentron, US) around the lens, layer by layer leaving a small (approx.
741 1 mm) gap below the baseplate, which was filled with 2-component adhesive (Loctite 3090,
742 Henkel DE) for ultimate fixation. After drying, the miniscope was de-attached and a custom-
743 made protective cap was put on the baseplate and fixated by the baseplate screw.

744 **Calcium imaging**

745 Calcium imaging was done using UCLA miniscopes v3 or v4 [29], including their data
746 acquisition (DAQ) box (<https://open-ephys.org>) and acquisition software
747 (www.miniscope.org). The temporally aligned recording of behavioural events and imaging
748 frames was achieved through pyControl [19,28], connecting the miniscope DAQ-box
749 (<https://open-ephys.org>) via an input trigger GPIO SMA connector to the pyControl
750 microcontroller board through which the start and the end of image acquisition was controlled
751 by TTL-pulses sent to the miniscope DAQ-box. Prior to the session start, mice were equipped
752 with the miniscope and the optimal focus was set by adjusting the focus slider of the miniscope
753 manually (v3) or the focus electronically (v4). A thin and flexible coaxial cable (CW2040-
754 3650SR, Cooner Wire, US) connected the miniscope to the DAQ-box for power supply, LED
755 control, and CMOS data transmission. For some recordings a custom-made motorized
756 commutator [30] was used to eliminate the need to manually un-twist the cable. Images were

757 recorded at 20 fps, maximum gain, and with an excitation intensity that was adjusted for each
758 mouse individually.

759 **Histology**

760 Animals were given an over-dose of ketamine/medetomidine (≥ 200 mg/kg ketamine, Zoetis,
761 G; ≥ 2 mg/kg medetomidine, Pfizer, US) and perfused with 0.01 M phosphate-buffered saline
762 (PBS) followed by 4 % PFA/PBS. The baseplate and the implant were carefully detached from
763 the skull and the brain was removed and then stored in 4 % PFA/PBS overnight before
764 placement in 20 % sucrose for dehydration before sections were cut at 100 μ m thickness on
765 a vibratome (VT1000, Leica, DE). Every second section was stained with DAPI (10⁻⁴ % w/v)
766 for 30 min, washed with PBS twice and mounted on glass slides. A Leica DM6B
767 epifluorescence microscope (Leica, DE) was used to scan the slides with a 5x objective and
768 determine virus expression offline.

769 **Data analysis**

770 **Pre-processing of calcium traces**

771 Single-photon imaging data for each session were pre-processed using MATLAB as described
772 previously [17]. Each image frame was spatially down-sampled to a 400x400 pixel frame and
773 divided by its low-pass filtered version to remove wide-field fluctuations and brightness
774 gradients over the field of view. After band-pass filtering each frame to enhance structural
775 features of the image to facilitate the alignment of different frames, the TurboReg algorithm
776 [31] was used for motion correction. Each movie was temporally smoothed and temporally
777 down-sampled from 20 Hz to 5 Hz followed by signal normalization of each image frame in
778 units of relative changes in fluorescence, $\Delta F(t)/F_0 = (F(t) - F_0)/F_0$, where F_0 is the mean
779 image obtained by averaging the entire movie. For cell sorting, spatial filters corresponding to
780 individual neurons were identified using an automated cell sorting routine based on principal
781 and independent component analysis (PCA/ICA) [32]. Extracted spatial filters were verified as
782 neural cells upon visual inspection based on size, morphology and the activity trace.

783 **Temporal alignment of calcium traces to behavioural events**

784 Calcium traces of each cell were z-transformed using the Matlab function *zscore* to control for
785 variations between animals and sessions. Custom written Matlab scripts were used first to
786 align the behavioural and the imaging data determining the timestamps for the ITI start, cue
787 presentation, choice and outcome (referred to as epochs) of each trial in terms of frame
788 numbers, and for labelling the trial with the corresponding choice of the mouse, i.e. correct,
789 incorrect or premature response, or omission (referred to as response type). In each trial, the
790 calcium signals were extracted within defined time windows from 4 s before to 7 s after the
791 event timestamp of the onset of each epoch. The extracted traces were then averaged across
792 trials of the same response type, thereby forming population vectors that represented each
793 response type aligned to the onset of each epoch. Peri-event time histograms were created
794 by plotting heat maps of the population vectors for each response type aligned to the choice
795 onset (Figure 2). The cells were sorted based on their average peak latency. For cross-
796 validation, heat maps were created based on the population vectors created from averaging
797 only across even or odd trials, which were then correlated across cells within each time bin to
798 receive a measure for the reliability of the temporal activity pattern. K-means clustering was
799 applied using the Matlab function *kmeans* with a preset number of four clusters and the
800 distance metric set to 'cosine'. Thereby, the cells were grouped based on the similarity of their
801 calcium signals, which were extracted within defined time windows relative to the onset of the
802 respective choice (see above) and averaged across *odd* trials. Subsequently, the cells were
803 sorted according to their cluster assignment and peak latency applying the sorting order to the
804 calcium signals averaged across *even* trials and plotting them using peri-event time
805 histograms (Figure 2C-D). For each cluster, the single-trial calcium signal of a corresponding
806 exemplary cell was plotted for each response type using peri-event time histograms (Figure
807 2E-F).

808 **Decoding analysis of population activity**

809 Binary linear support vector machine (SVM) classifiers (Figure 3A-C and 6E-F) were trained
810 and tested on differentiating between trials with correct responses and those with either

811 omitted, premature or incorrect responses based on the amplitude of the calcium trace of all
812 neurons in a FOV combined in 200 ms time-bins in a time-window from 4 s before until 7 s
813 after the onset of each choice poke. For training the classifier, the Matlab function *fitcsvm* was
814 used with the kernel function, box constraint, kernel scale and standardization set to 'linear',
815 '1', 'auto' and 'false', respectively. The dataset of every session (with a minimum of 6 events
816 for each response type) was randomly partitioned into the training (each observation was
817 labelled with the response type) and test set (lacking label assignment) in a 80/20 ratio, while
818 ensuring the test set maintained balance, resulting in an imbalanced distribution within the
819 training set for most sessions. For sessions, where the number of trials varied extensively
820 across response types, trials from the response type with a higher number of events were
821 randomly removed until achieving balance in the test set. For training the SVM classifier, a
822 *balanced training set* is essential to prevent a classification bias towards the majority class
823 (i.e. the behavioural event class with the highest number of observations in the respective data
824 set) [35]. Using the *synthetic minority oversampling technique* (SMOTE) [36,37] on the training
825 set, the number of observations in each event class was equalized by artificially synthesizing
826 new samples in the minority classes (i.e. the behavioural event classes with a lower number
827 of observations than the majority class). This algorithm randomly selects an observation from
828 the underrepresented event class and identifies its four nearest neighbours, of which one is
829 randomly chosen. A value is randomly picked in the Euclidean distance between the
830 observation and the neighbour and is assigned to the new synthesized sample. The smote
831 approach requires the number of events in the minority class to be greater than the number of
832 set neighbours (i.e. four) and the ratio between the number of events in the majority and
833 minority class to be less than the set number of neighbours (i.e. four). In sessions where this
834 was not the case, events were up-sampled for the minority class using SMOTE with the
835 number of neighbours set to the number of events in the minority class and/or trials were
836 randomly removed from the majority class until the required conditions were met. The entire
837 procedure of random data set partitioning, SMOTE up-sampling of the training set, and the
838 subsequent training and testing of the decoder was repeated 100-times, thereby producing

839 100 averages and 100 s.e.m. values, which were averaged to yield a *grand mean* and *grand*
840 *s.e.m* value representing the decoding performance based on each session. As a control for
841 the test decoder, a second binary classifier model was established as described above, but
842 labels of the training set in each fold within each session were shuffled prior to classifier
843 training, thereby creating distributions that represented chance level (null distribution) [38].

844 Since mice mostly made an insufficient number of incorrect responses during the varITI
845 challenges, to allow this type of decoding analysis, the data from several sessions with at least
846 8 incorrect responses were grouped across challenge protocols (five combined sessions, two
847 varITI sessions, two fixed-ITI sessions) to perform a separate decoding analysis that included
848 incorrect responses (Figure 3C).

849 Multiclass decoding was performed to predict into which of the five poke holes the mouse
850 poked into during a given trial including all response types or correct responses only, using a
851 multi-class SVM classifier [39,40] with the same approach as described above (Figure 4D).
852 The linear SVM multi-classifier was trained using the matlab function *fitcecoc* with the kernel
853 function, box constraint, kernel scale and standardization set to 'linear', '1', 'auto' and 'off',
854 respectively. Additionally, the option coding was set to 'onevsall'; the one-vs-all strategy
855 performs a separate binary classification for each class in the dataset (i.e. in total four) treating
856 it as the positive class, whereas all other classes combined are treated as the negative class.
857 Testing is performed by independently applying every sample from the test data set on each
858 trained binary classifier yielding confidence values with the highest selecting the predicted
859 class for this sample.

860 **Encoding analysis of the modulation of neural activity by behavioural events**

861 Linear regression models were created to predict the calcium signal in 200 ms timebins in a
862 time window from 4 s before until 7 s after the onset of the choice epoch for each individual
863 neuron (Figure 5 and 7). Regularized linear regression was performed using the Matlab
864 function *lasso* applying L1 (lasso) regularization with 10-fold cross-validation to find the optimal
865 regularization strength λ that minimizes the loss. Binary predictors were used to code for the

866 presence of a poke (active poke vs. omission), the spatial poke identity (four one-hot predictors
867 corresponding to left-right directionality, right poke discrimination, left poke discrimination and
868 middle poke discrimination), and correct (rewarded) responses (vs. erroneous and omitted
869 responses combined) in each trial (see Figure 5A for predictor matrix). To test how much of
870 the variance of the activity of individual neurons at every time bin could be explained by each
871 predictor, the coefficient of partial determination (CPD) was calculated, measuring how much
872 further the predictor contributed to the explanation of the full regression model [21,33,34].
873 CPDs were determined for each predictor by subtracting the mean squared error of the full
874 linear regression model from the mean squared error of the reduced regression model, where
875 the predictor for the specific event type in question was removed. The CPD for predictor i is
876 defined as:

$$877 \quad CPD_i = \frac{MSE_{X-i} - MSE_X}{MSE_{X-i}}$$

878 where MSE_{X-i} is the mean squared error in a regression model that includes all of the relevant
879 predictor variables except i , and MSE_X is the mean squared error in a regression model that
880 includes all of the relevant predictor variables. To compute the CPD for spatial poke identity,
881 all of the four spatial one-hot predictors were removed together.

882 For the devaluation and extinction conditions (Figure 7), which led to reduced numbers of
883 correct responses, pseudo-randomly chosen events from the baseline (control) condition were
884 selected to roughly match the number of the respective experimental condition for each poke-
885 hole and response type, to ensure that CPDs in the baseline session are not simply higher
886 because of a larger number of trials or responses used for their calculation. This down-
887 sampling was repeated 100 times and the CPDs of the respective predictors averaged across
888 iterations before plotting.

889

890 **Statistics**

891 Behavioural data was analysed using Matlab (R2019a, The MathWorks Inc, USA) and only
892 two-sided tests were used. 5-CSRTT performance during calcium imaging sessions (Figure
893 1D-E, Figure 6A-D) was analysed using an ANOVA involving the task paradigm as between-
894 subject independent variable and one of the behavioural parameters as dependent variable.
895 In case of a significant effect of task paradigm, Dunnett's post-hoc tests were conducted
896 between the baseline and any other challenge. Decoding accuracies (Figure 3) were
897 statistically compared using repeated-measures ANOVA with the time-bin and epoch type
898 variable as within-subject factors. A Dunn-Sidak-test was used for post-hoc testing.
899 Comparisons against accuracies of control classifiers (trained with shuffled labels, performing
900 at chance level) in decoding analyses or against 0% CPD in encoding analyses have been
901 done with paired-sample or one-sample *t*-tests, respectively, with Benjamini-Hochberg
902 corrections for the repeated testing in each time interval. All applied statistical tests are stated
903 in the corresponding figure legends. All bar and line graphs display mean \pm s.e.m. or data from
904 individual mice, as indicated.

905

906 **Data availability**

907 All raw data can be obtained from the corresponding author upon reasonable request. Scripts
908 of all task files applied in custom-made operant boxes can be obtained from
909 <https://github.com/KaetzelLab/Operant-Box-Code> and design files for such operant boxes are
910 deposited at <https://github.com/KaetzelLab/Operant-Box-Design-Files>.

911

912 **Code availability**

913 Analysis scripts are available from GitHub at
914 https://github.com/martinjendryka/Jendryka_et_al_ACC_imaging_5CSRTT.git.

915

916 References

- 917 1. Dalley JW, Robbins TW. Fractionating impulsivity: neuropsychiatric implications. *Nat Rev*
918 *Neurosci.* 2017;18: 158–171. doi:10.1038/nrn.2017.8
- 919 2. Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, et al. Cognitive
920 dysfunction in psychiatric disorders: characteristics, causes and the quest for improved
921 therapy. *Nat Rev Drug Discov.* 2012;11: 141–168. doi:10.1038/nrd3628
- 922 3. Bari A, Dalley JW, Robbins TW. The application of the 5-choice serial reaction time task
923 for the assessment of visual attentional processes and impulse control in rats. *Nat Protoc.*
924 2008;3: 759–767.
- 925 4. Bari A, Robbins TW. Inhibition and impulsivity: Behavioral and neural basis of response
926 control. *Prog Neurobiol.* 2013;108: 44–79. doi:10.1016/j.pneurobio.2013.06.005
- 927 5. van der Veen B, Kapanaiiah SKT, Kilonzo K, Steele-Perkins P, Jendryka MM, Schulz S,
928 et al. Control of impulsivity by Gi-protein signalling in layer-5 pyramidal neurons of the
929 anterior cingulate cortex. *Commun Biol.* 2021;4: 1–16. doi:10.1038/s42003-021-02188-
930 w
- 931 6. Norman KJ, Koike H, McCraney SE, Garkun Y, Bateh J, Falk EN, et al. Chemogenetic
932 suppression of anterior cingulate cortical neurons projecting to the visual cortex disrupts
933 attentional behavior in mice. *Neuropsychopharmacol Rep.* 2021;41: 207–214.
934 doi:10.1002/npr2.12176
- 935 7. Norman KJ, Riceberg JS, Koike H, Bateh J, McCraney SE, Caro K, et al. Post-error
936 recruitment of frontal sensory cortical projections promotes attention in mice. *Neuron.*
937 2021 [cited 25 Feb 2021]. doi:10.1016/j.neuron.2021.02.001
- 938 8. Jendryka MM, Lewin U, van der Veen B, Kapanaiiah SKT, Prex V, Strahnen D, et al.
939 Control of sustained attention and impulsivity by Gq-protein signalling in parvalbumin
940 interneurons of the anterior cingulate cortex. *Transl Psychiatry.* 2023;13: 1–12.
941 doi:10.1038/s41398-023-02541-z
- 942 9. Totah NKB, Kim YB, Homayoun H, Moghaddam B. Anterior Cingulate Neurons
943 Represent Errors and Preparatory Attention within the Same Behavioral Sequence. *J*
944 *Neurosci.* 2009;29: 6418–6426. doi:10.1523/JNEUROSCI.1142-09.2009
- 945 10. Donnelly NA, Paulsen O, Robbins TW, Dalley JW. Ramping single unit activity in the
946 medial prefrontal cortex and ventral striatum reflects the onset of waiting but not imminent
947 impulsive actions. *Eur J Neurosci.* 2015; n/a-n/a. doi:10.1111/ejn.12895
- 948 11. Hyman JM, Whitman J, Emberly E, Woodward TS, Seamans JK. Action and Outcome
949 Activity State Patterns in the Anterior Cingulate Cortex. *Cereb Cortex.* 2013;23: 1257–
950 1268. doi:10.1093/cercor/bhs104
- 951 12. Broom E, Imbriotis V, Sengpiel F, Connelly WM, Ranson A. Recruitment of frontal
952 sensory circuits during visual discrimination. *Cell Rep.* 2022;39: 110932.
953 doi:10.1016/j.celrep.2022.110932
- 954 13. Wal A, Klein FJ, Born G, Busse L, Katzner S. Evaluating Visual Cues Modulates Their
955 Representation in Mouse Visual and Cingulate Cortex. *J Neurosci.* 2021;41: 3531–3544.
956 doi:10.1523/JNEUROSCI.1828-20.2021

- 957 14. Kim H, Åhrlund-Richter S, Wang X, Deisseroth K, Carlén M. Prefrontal Parvalbumin
958 Neurons in Control of Attention. *Cell*. 2016;164: 208–218. doi:10.1016/j.cell.2015.11.038
- 959 15. Hunt LT, Malalasekera WMN, Berker AO de, Miranda B, Farmer SF, Behrens TEJ, et al.
960 Triple dissociation of attention and decision computations across prefrontal cortex. *Nat*
961 *Neurosci*. 2018;21: 1471–1481. doi:10.1038/s41593-018-0239-5
- 962 16. Hunt LT, Hayden BY. A distributed, hierarchical and recurrent framework for reward-
963 based choice. *Nat Rev Neurosci*. 2017;18: 172–182. doi:10.1038/nrn.2017.7
- 964 17. Grewe BF, Gründemann J, Kitch LJ, Lecoq JA, Parker JG, Marshall JD, et al. Neural
965 ensemble dynamics underlying a long-term associative memory. *Nature*. 2017;543: 670–
966 675. doi:10.1038/nature21682
- 967 18. Lui JH, Nguyen ND, Grutzner SM, Darmanis S, Peixoto D, Wagner MJ, et al. Differential
968 encoding in prefrontal cortex projection neuron classes across cognitive tasks. *Cell*.
969 2021;184: 489-506.e26. doi:10.1016/j.cell.2020.11.046
- 970 19. Kapaniaiah SKT, van der Veen B, Strahnen D, Akam T, Kätzel D. A low-cost open-source
971 5-choice operant box system optimized for electrophysiology and optophysiology in mice.
972 *Sci Rep*. 2021;11: 22279. doi:10.1038/s41598-021-01717-1
- 973 20. Li Y, Mathis A, Grewe BF, Osterhout JA, Ahanonu B, Schnitzer MJ, et al. Neuronal
974 Representation of Social Information in the Medial Amygdala of Awake Behaving Mice.
975 *Cell*. 2017;171: 1176-1190.e17. doi:10.1016/j.cell.2017.10.015
- 976 21. Akam T, Rodrigues-Vaz I, Marcelo I, Zhang X, Pereira M, Oliveira RF, et al. The anterior
977 cingulate cortex predicts future states to mediate model-based action selection. *Neuron*.
978 2020;0. doi:10.1016/j.neuron.2020.10.013
- 979 22. Totah NKB, Jackson ME, Moghaddam B. Preparatory Attention Relies on Dynamic
980 Interactions between Prelimbic Cortex and Anterior Cingulate Cortex. *Cereb Cortex*.
981 2013;23: 729–738. doi:10.1093/cercor/bhs057
- 982 23. Koike H, Demars MP, Short JA, Nabel EM, Akbarian S, Baxter MG, et al. Chemogenetic
983 Inactivation of Dorsal Anterior Cingulate Cortex Neurons Disrupts Attentional Behavior
984 in Mouse. *Neuropsychopharmacology*. 2016;41: 1014–1023. doi:10.1038/npp.2015.229
- 985 24. Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct Roles of Rodent Orbitofrontal and
986 Medial Prefrontal Cortex in Decision Making. *Neuron*. 2010;66: 449–460.
987 doi:10.1016/j.neuron.2010.03.033
- 988 25. Sohal VS, Zhang F, Yizhar O, Deisseroth K. Parvalbumin neurons and gamma rhythms
989 enhance cortical circuit performance. *Nature*. 2009;459: 698–702.
990 doi:10.1038/nature07991
- 991 26. Terra H, Bruinsma B, de Kloet SF, van der Roest M, Pattij T, Mansvelder HD. Prefrontal
992 Cortical Projection Neurons Targeting Dorsomedial Striatum Control Behavioral
993 Inhibition. *Curr Biol*. 2020;30: 4188-4200.e5. doi:10.1016/j.cub.2020.08.031
- 994 27. de Kloet SF, Bruinsma B, Terra H, Heistek TS, Passchier EMJ, van den Berg AR, et al.
995 Bi-directional regulation of cognitive control by distinct prefrontal cortical output neurons
996 to thalamus and striatum. *Nat Commun*. 2021;12: 1994. doi:10.1038/s41467-021-22260-
997 7

- 998 28. Akam T, Lustig A, Rowland JM, Kapanaiiah SK, Esteve-Agraz J, Panniello M, et al. Open-
999 source, Python-based, hardware and software for controlling behavioural neuroscience
1000 experiments. Kemere C, Wassum KM, Kemere C, Siegle J, editors. *eLife*. 2022;11:
1001 e67846. doi:10.7554/eLife.67846
- 1002 29. Aharoni D, Hoogland TM. Circuit Investigations With Open-Source Miniaturized
1003 Microscopes: Past, Present and Future. *Front Cell Neurosci*. 2019;13.
1004 doi:10.3389/fncel.2019.00141
- 1005 30. Kapanaiiah SKT, Kätzel D. Open-MAC: A low-cost open-source motorized commutator
1006 for electro- and opto-physiological recordings in freely moving rodents. *HardwareX*.
1007 2023;14: e00429. doi:10.1016/j.ohx.2023.e00429
- 1008 31. Thevenaz P, Ruttimann UE, Unser M. A pyramid approach to subpixel registration based
1009 on intensity. *IEEE Trans Image Process*. 1998;7: 27–41. doi:10.1109/83.650848
- 1010 32. Mukamel EA, Nimmerjahn A, Schnitzer MJ. Automated Analysis of Cellular Signals from
1011 Large-Scale Calcium Imaging Data. *Neuron*. 2009;63: 747–760.
1012 doi:10.1016/j.neuron.2009.08.009
- 1013 33. Cai X, Kim S, Lee D. Heterogeneous Coding of Temporally Discounted Values in the
1014 Dorsal and Ventral Striatum during Intertemporal Choice. *Neuron*. 2011;69: 170–182.
1015 doi:10.1016/j.neuron.2010.11.041
- 1016 34. Chiang F-K, Wallis JD. Neuronal encoding in prefrontal cortex during hierarchical
1017 reinforcement learning. *J Cogn Neurosci*. 2018;30: 1197–1208.
1018 doi:10.1162/jocn_a_01272
- 1019 35. Krawczyk B. Learning from Imbalanced Data: Open Challenges and Future Directions.
1020 *Prog Artif Intell*. 2016;5: 221–232. doi:10.1007/s13748-016-0094-0
- 1021 36. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-
1022 sampling Technique. *J Artif Intell Res*. 2002;16: 321–357. doi:10.1613/jair.953
- 1023 37. Larsen BS. Synthetic Minority Over-Sampling Technique (SMOTE). 2020.
- 1024 38. Combrisson E, Jerbi K. Exceeding chance level by chance: The caveat of theoretical
1025 chance levels in brain signal classification and statistical assessment of decoding
1026 accuracy. *J Neurosci Methods*. 2015;250: 126–136. doi:10.1016/j.jneumeth.2015.01.010
- 1027 39. Boser BE, Guyon IM, Vapnik VN. A Training Algorithm for Optimal Margin Classifiers.
1028 *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. New
1029 York, NY, USA: Association for Computing Machinery; 1992. pp. 144–152.
1030 doi:10.1145/130385.130401
- 1031 40. Vapnik V. *The Nature of Statistical Learning Theory*. Springer Science & Business Media;
1032 2013.
- 1033
- 1034

1035 **Acknowledgments**

1036 We thank Stefanie Schulz (Ulm University) for assistance with histology. Funding: This work
1037 was funded by the Boehringer Ingelheim-Ulm University (BIU) Center (TPN010, to B.L., A.P.,
1038 D.K.), the Else-Kroener-Fresenius/German-Scholars-Organization Programme for excellent
1039 medical scientists from abroad (GSO/EKFS 12; to D.K.), the DFG (KA 4594/2-1; to D.K.), and
1040 the Alfred-Krupp Foundation (to B.L.).

1041

1042 **Author Contributions**

1043 M.M.J., B.L., A.P., T.A. and D.K. designed the study. M.J. and U.L. conducted behavioural
1044 experiments. M.M.J. conducted surgeries. S.K.T.K., T.A., and D.K. developed pyOS-5 operant
1045 box hardware and software; S.K.T.K. programmed operant box task protocols and integration
1046 of miniscope recordings. B.F.G. and H.D. provided assistance with manufacturing and usage
1047 of UCLA v3 miniscopes. B.F.G., B.L. and D.K. provided essential resources. M.M.J. analysed
1048 the data with advise from B.F.G., T.A. and D.K.. M.M.J. and D.K. wrote the manuscript, which
1049 was revised by all authors.

1050

1051 **Competing Interests statement**

1052 The authors declare no competing interest. A.P. is an employee of Boehringer Ingelheim.

1053