



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/182569/>

Version: Accepted Version

Article:

Delis, I, Ince, RAA, Sajda, P et al. (2022) Neural Encoding of Active Multi-Sensing Enhances Perceptual Decision-Making via a Synergistic Cross-Modal Interaction. *The Journal of Neuroscience*, 42 (11). pp. 2344-2355. ISSN: 0270-6474

<https://doi.org/10.1523/JNEUROSCI.0861-21.2022>

© 2022 the authors. This is an author produced version of an article published in *The Journal of Neuroscience*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

1 Neural encoding of active multi-sensing enhances perceptual decision-making
2 via a synergistic cross-modal interaction

3 Ioannis Delis^{1*}, Robin A.A. Ince², Paul Sajda^{3,4}, Qi Wang³

4

5 ¹School of Biomedical Sciences, University of Leeds, Leeds, LS2 9JT, UK

6 ²School of Psychology and Neuroscience, University of Glasgow, G12 8QQ, UK

7 ³Department of Biomedical Engineering, Columbia University, New York, NY,
8 10027, USA

9 ⁴Data Science Institute, Columbia University, New York, NY, 10027, USA

10

11 *Please address correspondence to i.delis@leeds.ac.uk

12

13 **Abbreviated title:** Neural mechanisms of active multisensory decisions

14

15 **Competing Interest Statement:** The authors declare no competing financial
16 interests.

17

18 **Acknowledgments**

19

20 This work was supported by the the European Commission (H2020-MSCA-IF-
21 2018/845884, “NeuCoDe” to I.D.), the Physiological Society (2018 Research
22 Grant Scheme to I.D.), the National Institutes of Health (R01-MH085092 to P.S.),
23 the U.S. Army Research Laboratory (W911NF-10-2-0022 to P.S.), the Wellcome
24 Trust ([214120/Z/18/Z] to R.A.A.I.) and the UK Economic and Social Research
25 Council (ES/L012995/1 to P.S.), and a NARSAD Young Investigator award to
26 Q.W.

27

28 No of pages: 32

29 No of figures: 4

30 Abstract: 250 words

31 Significance statement: 120 words

32 Discussion: 1331 words

33

34 **Abstract**

35 Most perceptual decisions rely on the active acquisition of evidence from the
36 environment involving stimulation from multiple senses. However, our
37 understanding of the neural mechanisms underlying this process is limited.
38 Crucially, it remains elusive how different sensory representations interact in the
39 formation of perceptual decisions. To answer these questions, we employed an
40 active sensing paradigm coupled with neuroimaging, multivariate analysis and
41 computational modeling to probe how the human brain processes multisensory
42 information to make perceptual judgments. Participants of both sexes actively
43 sensed to discriminate two texture stimuli using visual (V) or haptic (H)
44 information or the two sensory cues together (VH). Crucially, information
45 acquisition was under the participants' control, who could choose where to
46 sample information from and for how long on each trial. To understand the neural
47 underpinnings of this process, we first characterized where and when active
48 sensory experience (movement patterns) is encoded in human brain activity
49 (electroencephalography - EEG) in the three sensory conditions. Then, to offer a
50 neurocomputational account of active multisensory decision formation, we used
51 these neural representations of active sensing to inform a drift diffusion model of
52 decision-making behavior. This revealed a multisensory enhancement of the
53 neural representation of active sensing which led to faster and more accurate
54 multisensory decisions. We then dissected the interactions between the V, H and
55 VH representations using a novel information-theoretic methodology. Ultimately,
56 we identified a synergistic neural interaction between the two unisensory (V, H)
57 representations over contralateral somatosensory and motor locations that
58 predicted multisensory (VH) decision-making performance.

59

60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88

Significance Statement

In real-world settings, perceptual decisions are made during active behaviors, such as crossing the road on a rainy night, and include information from different senses (e.g. car lights, slippery ground). Critically, it remains largely unknown how sensory evidence is combined and translated into perceptual decisions in such active scenarios. Here we address this knowledge gap. First, we show that the simultaneous exploration of information across senses (multi-sensing) enhances the neural encoding of active sensing movements. Second, the neural representation of active sensing modulates the evidence available for decision and, importantly, multi-sensing yields faster evidence accumulation. Finally, we identify a crossmodal interaction in the human brain that correlates with multisensory performance, constituting a putative neural mechanism for forging active multisensory perception.

89 **Introduction**

90

91 In our daily lives, we make judgments based on noisy or incomplete
92 information that we gather from our environment (Heekeren et al., 2004;
93 Juavinett et al., 2018; Najafi and Churchland, 2018), usually including stimuli
94 from multiple senses (Angelaki et al., 2009; Chandrasekaran, 2017). The
95 acquired sensory information crucially depends on our actions - what we see,
96 hear and touch is influenced by our movements - a process known as active
97 sensing (Schroeder et al., 2010; Yang et al., 2016b). For example, imagine
98 attempting to cross the road on a rainy night. You need to interact with the
99 environment, i.e. turn your head and move your eyes, and process the incoming
100 stimuli (e.g. car lights, slippery ground) to decide whether and when it is safe to
101 do so. If you feel the road is slippery, you may need to monitor your steps and at
102 the same time you may have to walk faster or step back if a car is approaching.

103 This example indicates that in real-world settings most perceptual
104 decisions are made during active behaviors (Musall et al., 2019). The quality of
105 the acquired evidence is driven by such active behaviors, which, in turn, affect
106 the efficiency of the perceptual decisions that we make as a result of this active
107 sensing process (Yang et al., 2016a; Gottlieb and Oudeyer, 2018). A first crucial
108 element of fast and accurate perceptual decisions is the combination of evidence
109 from different sensory streams (e.g. sight and touch) to form a unified percept
110 and reduce uncertainty about the stimulus (Ernst and Banks, 2002). However,
111 while there is extensive evidence that the integration of information from different
112 sensory modalities improves perceptual choice accuracy (Lewis and Noppeney,
113 2010; Raposo et al., 2012) and response time (Drugowitsch et al., 2014),
114 multisensory information processing has not been studied in an active scenario,
115 where human participants are allowed to implement their own strategy for
116 gathering evidence - as is the case in real life settings.

117 Here we addressed this gap in the literature aiming to uncover the neural
118 mechanisms underlying the formation of perceptual decisions via the active

119 acquisition and processing of multisensory information. To achieve this, we
120 capitalised on our previous work probing the neural correlates of active tactile
121 decisions (Delis et al., 2018) and extended it to a multisensory setting that
122 includes visual and haptic information presented simultaneously or separately.
123 We hypothesised that the neural encoding of active sensory experience would be
124 enhanced when multisensory information was available and that this neural
125 multisensory gain would lead to improvements in decision-making performance.

126 An important aspect of our study is that the participants had full control of
127 the evolution and duration of each trial. In other words, they could choose how
128 much information to sample, where to sample this information from and for how
129 long. Thus, we first aimed to characterise cortical coupling to continuous active
130 sensing and then combined this with a popular sequential-sampling model of
131 decision-making, the drift diffusion model (DDM) (Ratcliff and McKoon, 2008), to
132 understand how the identified representations of active sensing behaviors
133 influence decisions in the human brain. Here, to bridge the gap between active
134 evidence acquisition and decision formation, we used the neural correlates of
135 active (multi-)sensing to constrain the DDM.

136 Finally, to quantify cross-modal interactions in the brain, we applied a
137 novel information-theoretic framework named Partial Information Decomposition
138 (PID) (Williams and Beer, 2010; Timme et al., 2014; Ince, 2017). PID quantifies
139 the contribution of a) each sensory modality and b) cross-modal representational
140 interactions (“redundant” or “synergistic”) to the multisensory neural
141 representation (Park et al., 2018). Redundancy measures the similarity of the
142 neural representation of the two modalities, while synergy indicates a better
143 prediction of the neural response from both modalities simultaneously. Ultimately,
144 this approach revealed the interactions between representations of different
145 sensing modalities in the brain and shed light onto their role in decision-making
146 behavior.

147
148

149 **Materials and Methods**

150

151 *Experimental design and paradigm.* Fourteen healthy right-handed participants
152 (8 female, aged 24 ± 2 years) performed a two-alternative forced choice (2AFC)
153 discrimination task during which they had to compare the amplitudes of two
154 sinusoidal stimuli of the same frequency. All experimental procedures have been
155 reviewed and approved by the Institutional Review Board (IRB) at Columbia
156 University.

157 To generate visual and tactile stimuli that can be actively sensed, we
158 employed a haptic device called a Pantograph (Campion et al., 2005), which can
159 be controlled to generate the sensation of exploring real surfaces (Fig. 1A). The
160 Pantograph is a 2-dimensional force-feedback device, that is, a) it produces a 2D
161 tactile output and b) it simultaneously measures 2D information about the finger
162 position and applied force. Here we used its first property to generate stimulation
163 and the second property to record the kinematics of the movements performed
164 by the participants while they actively explored the presented stimuli. In
165 particular, we split the workspace of the Pantograph (of dimensions 110 mm x 60
166 mm) into two subspaces (left - L and right - R, 55 mm x 60mm each) and
167 generated continuous sinusoidal stimuli of different amplitudes (but same
168 wavelength of 10 mm) in the two subspaces (Fig. 1B). Then, we instructed the
169 participants to discriminate the amplitude of the two subspaces as quickly and as
170 accurately as possible a) using only visual (V) information, b) using only haptic
171 (H) information and c) combining the two sensory cues (VH). Crucially for our
172 investigation here, participants were free to choose how to explore this virtual
173 texture, i.e., where and how fast to move their fingers and how long to explore
174 each one of the two sides for before making their perceptual choice. Participants
175 placed their right index finger on the interface plate of the Pantograph (see Fig.
176 1A) and moved it freely to explore the textures of both subspaces (Fig. 1C)
177 before reporting their choice (i.e., which amplitude is higher) by pressing one of
178 two buttons on a keyboard (left or right arrow) using their left hand.

179 Specifically, in the H condition, the Pantograph produced sinusoidal forces of
180 different intensity between L and R. When the participants placed their index
181 fingers on the plate (interface) of the Pantograph, these forces at the interface
182 had the effect of causing fingertip deformations and thus tactile sensations that
183 resembled exploring real surfaces. Thus, when moving their finger on the
184 Pantograph, participants had the sensation of touching a rough surface (with
185 different amplitudes between L and R – see Fig.1B middle panel). In the V
186 condition, stimuli matching the tactile stimuli were presented on a screen of the
187 same dimensions. More precisely, amplitudes of the sinusoidal virtual texture in
188 H were translated into contrast levels of sinusoidal gratings in V, i.e. the
189 participants were seeing black and white stripes of different intensity/contrast
190 between L and R. Presentation of visual stimuli was generated using
191 Psychtoolbox and visual contrast varied between 0.5 and 1.5 around the default
192 contrast value. The visual angle was $12^{\circ} \pm 6^{\circ}$. Stimulus presentation was
193 controlled by a real-time hardware system (Matlab XPCTarget) to minimize
194 asynchrony which was $<1\text{ms}$. Importantly, to match the sense of touch, only the
195 part of the workspace corresponding to the participant's finger location was
196 revealed on the screen (i.e. a moving dot following the participant's finger - see
197 Fig. 1B left panel). Thus, in the V condition, grayscale visual textures (of different
198 contrast between L and R) were shown wherever the participants moved their
199 fingers while no forces were applied to the participants' fingers (i.e. no H
200 stimulation). Hence, in both sensory domains, participants could only sense the
201 presented stimulus via active exploration, i.e. finger movements on the x-axis.
202 Accordingly, in the VH condition, both the visual and haptic textures were
203 congruently presented and sensed by the participants using finger movements
204 (Fig. 1B right panel). Overall, participants had to decide whether L or R had
205 higher amplitude based on their haptic (in H trials), visual (in V trials) or visuo-
206 haptic (in VH trials) perception of this virtual surface. Participants reported that
207 they perceived the V and H signals as one stimulus in the VH condition.

208 The amplitude difference between L and R (representing the difficulty of the task)
209 varied from trial to trial. On each trial, participants compared between the
210 reference amplitude 1 (presented either on the left or right subspace) and one of
211 six other amplitude levels (0.5, 0.75, 0.9, 1.1, 1.25, 1.5). Each trial was initiated
212 by the participant. Trial onset was considered the time point at which horizontal
213 finger velocity exceeded 0. Trial duration was determined by the participant and
214 lasted for the whole period during which the participant made exploratory
215 movements to sense the surface. The trial ended when the participant pressed
216 the < or > key on the keyboard with their left hand to indicate their L or R choice.
217 Each participant performed 20 trials for each amplitude level and for each
218 sensory condition (V, H, VH), resulting $K = 20$ trials x 6 amplitudes x 3 conditions
219 = 360 trials in total. One participant showed poor behavioral performance
220 (accuracy was not significantly different from chance level) and another
221 participant's EEG recordings were significantly contaminated with eye movement
222 artifacts, thus data from these two participants were removed from any
223 subsequent analyses. We report results from the remaining $N = 12$ participants.
224 We also discarded trials in which participants did not respond within 10secs from
225 trial onset or their RTs were shorter than 0.3 seconds. This resulted in the
226 rejection of 4.9% of the trials.

227 *Data recording and pre-processing.* During performance of the task, we
228 measured a) the choice accuracy and response time (RT) of participants'
229 responses, b) movement kinematics (x, y coordinates of finger position recorded
230 by the Pantograph) at a sampling frequency of 1000Hz and c) EEG signals at
231 2048 sampling frequency using a Biosemi EEG system (ActiveTwo AD-box, 64
232 Ag-AgCl active electrodes, 10-10 montage).

233 To compare accuracies and RTs across sensory conditions, we used two-way
234 ANOVAs with factors condition and stimulus difference followed by Bonferroni-
235 corrected post-hoc t-tests. We also fit psychometric curves to the accuracy data

236 of each participant using a cumulative Gaussian distribution and computed the
237 point of subjective equality (PSE) and slope of the curve at the PSE.

238 Single-trial movement velocity waveforms were computed using the derivatives
239 of the recorded position. EEG recordings were preprocessed using EEGLab
240 (Delorme and Makeig, 2004) as follows. EEG signals were first down-sampled to
241 1000Hz to match movement kinematics and dynamics. Then, they were
242 bandpass filtered to 1-50Hz using a Hamming windowed FIR filter. To isolate the
243 purely neural component of the EEG data, we used the following procedure: we
244 first reduced the dimensionality of the EEG data by reconstituting the data using
245 only the top 32 principal components derived from Principal Component Analysis
246 (PCA). Although we record from 64 channels, we expect our recordings to span a
247 considerably lower-dimensional space (as a result of correlations, crosstalk and
248 common sources), thus to enhance the ability of ICA to identify truly independent
249 components, we reduce the data dimensions to half using PCA. Thereafter, an
250 Independent Component Analysis (ICA) decomposition of the data was
251 performed using the Infomax algorithm (Bell and Sejnowski, 1995). We then used
252 an ICA-based artifact removal algorithm called MARA (Winkler et al., 2011) to
253 remove ICs attributed to blinks, horizontal eye movements (HEOG), muscular
254 activity (EMG), and any loose or highly noisy electrodes. MARA assigned each
255 IC a probability of being an artifact; we removed components with probabilities
256 above 0.5.

257 *Decoding finger kinematics from EEG signals.* To assess the neural encoding of the
258 participants' active sensory experience in the three sensory conditions, we used a
259 multivariate linear regression analysis introduced in (Di Liberto et al., 2015) and shown
260 in Equation 1 below. As in our previous work (Delis et al., 2018), we hypothesized that
261 the sensorimotor strategy employed by the participant can be represented by the
262 velocity profiles of the participant's exploratory movements which capture changes of
263 movement direction as well as speed changes. Thus, as kinematic feature representing
264 the active sensing behavior, we used 1-d finger velocity on the x-axis (capturing L-R
265 finger movements) but also finger position (on the x-axis) yielded qualitatively very

266 similar results. Finger movement in the y-axis (which did not provide any sensory
 267 information) did not show any significant correlation with the EEG signals and was not
 268 considered further. We thus performed a multivariate ridge regression (Crosse et
 269 al., 2016) predicting the 1-d finger velocity (on the x-axis) from the EEG data.
 270 Specifically, our decoding analysis aimed to reconstruct the movement velocity
 271 from a linear combination of the EEG recordings with time lags ranging between
 272 [-200ms, 400ms] with respect to the instantaneous velocity values. Specifically,
 273 we aimed to decode the velocity profile $s(t)$ of the participants' scanning
 274 movements from the simultaneously recorded EEG signals $m(i, t)$, as
 275 follows:

$$\hat{s}(t) \cong \sum_{\tau} \sum_i g(\tau, i) m(t + \tau, i) \quad (1)$$

276 where $\hat{s}(t)$ is the reconstructed finger velocity and $g(i, \tau)$ is a filter that integrates
 277 information spatially across EEG channels i and temporally across time lags τ to
 278 decode the velocity profile from the EEG recordings. Here we used $\tau \in$
 279 $[-200ms, 400ms]$, i.e. we examined the EEG information about the finger velocity
 280 at time t from $t - 200ms$ (200ms earlier) up to $t + 400ms$ (400ms later). Varying
 281 these lags did not improve reconstruction performance and yielded qualitatively
 282 similar results with the main effects always in the $[-200ms, 400ms]$ temporal
 283 window, so we used this window for all our further analyses. To learn the
 284 decoding filters and compute the velocity approximation accuracy (r^2) between
 285 the original and the reconstructed velocity profiles, we used the multivariate
 286 temporal response function (mTRF) Matlab Toolbox implementing regularised
 287 linear (ridge) regression (Crosse et al., 2016). In all our filter estimations, we
 288 used a cross-validation procedure. We first randomly split our data into two sets,
 289 a training set (80% of the trials) to learn the filters and a test set (the remaining
 290 20% of the trials) to apply the filters to and compute the reported r^2 values. In the
 291 training set, we performed 5-fold cross-validation to identify the optimal value of
 292 the ridge parameter λ (varying $\lambda = 2^0, \dots, 2^{20}$) that maximizes r^2 between the

293 estimated and the measured velocity. These investigations revealed that values
294 of λ between 2^0 and 2^4 yielded almost identical r^2 across all models, thus we
295 used $\lambda = 2^2$ for all models for consistency.

296 Since the weights of the decoding filters are not interpretable in terms of the
297 neural origins of the underlying processes (Haufe et al., 2014), we transformed
298 them into encoding filters $f(\tau, i)$ using the “forward model” formalism (Parra et
299 al., 2002; Haufe et al., 2014), as follows:

$$f(\tau, i) = \frac{m(t, i)^T m(t, i) g(i, \tau)}{\hat{s}(t)^T \hat{s}(t)} \quad (2)$$

300 We then plotted the weights of the forward models $f(\tau, i)$ at specific time lags τ
301 as scalp maps to visualise the relationship between sensorimotor behavior and
302 neural activity in each one of the three sensory conditions (V, H, VH). *Statistical*
303 *analysis of EEG-behavior couplings*. To determine statistical significance of the
304 learned EEG-velocity mappings, we randomized the phase spectrum of the EEG
305 signals, which disrupted the temporal relationship between the EEG activity and
306 the kinematics while preserving the autocorrelation structure of the signals
307 (Theiler et al., 1992). We generated 1000 phase-randomized surrogates of the
308 EEG data and computed correlations with the kinematics to define the null
309 distribution from which we estimated p-values. This phase-randomization
310 procedure maintains the magnitude spectrum of the EEG signals, thus
311 conserving their autocorrelation structure, which is a fundamental feature of the
312 original signals when the significance of cross-correlation is assessed. Hence,
313 using this procedure, the obtained surrogates that define the null distribution are
314 a more plausible comparison (resulting in a stricter statistical test) than randomly
315 shuffled surrogates.

316 *Informed modeling of decision-making performance*. Having characterised
317 the cortical coupling to the sensorimotor strategies in the three sensory
318 conditions, we then probed the relationship between the identified EEG-velocity

319 couplings and decision-making performance. To provide this missing link
320 between active sensing and decision formation, we implemented a Hierarchical
321 Drift Diffusion Model (HDDM), a well-known cognitive model of decision-making
322 behaviour, and informed it with the results of our previous decoding analysis.

323 We fit the participants' decision-making performance, i.e. accuracy and
324 response time (RT), with a hierarchical drift diffusion model (HDDM) (Wabersich
325 and Vandekerckhove, 2014) which assumes a stochastic accumulation of
326 sensory evidence over time, toward one of two decision boundaries
327 corresponding to correct and incorrect choices (Ratcliff and McKoon, 2008). The
328 model returns estimates of internal components of processing such as the rate of
329 evidence accumulation (drift rate), the distance between decision boundaries
330 controlling the amount of evidence required for a decision (decision boundary), a
331 possible bias towards one of the two choices (starting point) and the duration of
332 non-decision processes (non-decision time), which include stimulus encoding
333 and response production. As per common practice, we assumed that stimulus
334 differences affected the drift rate (Palmer et al., 2005).

335 In short, the model iteratively adjusts the above parameters to maximize
336 the summed log likelihood of the predicted mean response time (RT) and
337 accuracy. The DDM parameters were estimated in a hierarchical Bayesian
338 framework, in which prior distributions of the model parameters were updated on
339 the basis of the likelihood of the data given the model, to yield posterior
340 distributions (Wiecki et al., 2013; Wabersich and Vandekerckhove, 2014). The
341 use of Bayesian analysis, and specifically the hierarchical drift diffusion model
342 has several benefits relative to traditional DDM analysis. First, this framework
343 supports the use of other variables as regressors of the model parameters to
344 assess relations of the model parameters with other physiological or behavioral
345 data (Frank et al., 2015; Turner et al., 2015; Nunez et al., 2017). This regression
346 model, which is included in HDDM, allows estimation of trial-by-trial influences of
347 a covariate (e.g. a brain measure) onto DDM parameters. In other words, trial-by-

348 trial fluctuations of the estimated HDDM parameters can be approximated as a
349 linear combination of other trial-by-trial measures of cognitive function (Wiecki et
350 al., 2013; Forstmann et al., 2016). This property of the HDDM enabled us to
351 establish the link between the results of the EEG-velocity coupling analysis and
352 the decision parameters of the model, by using the EEG-velocity couplings as
353 predictors of the HDDM parameters, as explained below (also see Eq. 3 for an
354 example of such a linear regression of the drift rate parameter). Second, the
355 model estimates posterior distributions of the main parameters (instead of
356 deterministic values), which directly convey the uncertainty associated with
357 parameter estimates (Kruschke, 2010). Third, as a result of the above, the
358 hierarchical structure of the model allows estimation of the HDDM parameters
359 across participants and conditions, thus yielding distributions at different levels of
360 the model hierarchy (e.g. the population level and the participant level
361 respectively). In this way, the HDDM capitalizes on the statistical power offered
362 by pooling data across participants (population-level parameters) but at the same
363 time accounts for differences across participants (represented by the variance of
364 the population-level distribution and the individual participant-level estimates).
365 Fourth, the Bayesian hierarchical framework has been shown to be especially
366 effective when the number of observations is low (Ratcliff and Childers, 2015).

367 To implement the hierarchical DDM, we used the JAGS Wiener module
368 (Wabersich and Vandekerckhove, 2014) in JAGS (Plummer, 2003), via the
369 Matjags interface in Matlab to estimate posterior distributions. For each trial, the
370 likelihood of accuracy and RT was assessed by providing the Wiener first-
371 passage time (WFPT) distribution with the four model parameters (boundary
372 separation, starting point, non-decision time, and drift rate). Capitalizing on the
373 advantages of HDDM, we ran the model pooling data across all participants and
374 conditions and estimated both population-level and participant-level distributions.
375 Parameters were drawn from uniformly distributed priors and were estimated with
376 non-informative mean and standard deviation group priors. As per standard
377 practice for accuracy-coded data, the starting point was set as the midpoint

378 between the two decision boundaries as participants could not develop a bias
379 towards correct or incorrect choices.. For each model, we ran 3 separate Markov
380 chains with 5500 samples of the posterior parameters each; the first 500 were
381 discarded (as “burn-in”) and the rest were subsampled (“thinned”) by a factor of
382 50 following the conventional approach to MCMC sampling whereby initial
383 samples are likely to be unreliable due to the selection of a random starting point
384 and neighboring samples are likely to be highly correlated (Wabersich and
385 Vandekerckhove, 2014). The remaining samples constituted the probability
386 distributions of each estimated parameter. To ensure convergence of the chains,
387 we computed the Gelman-Rubin R^2 statistic (which compares within-chain and
388 between-chain variance) and verified that all group-level parameters had an R^2
389 close to 1 and always lower than 1.01.

390 Here, to obtain a mechanistic account of the effect of EEG-velocity
391 coupling on decision-making behaviour, we incorporated the single-trial
392 measures of these couplings (r^2 values defined above) into the HDDM
393 parameter estimation (Fig. 3B). Specifically, as part of the model fitting within the
394 HDDM framework, we used the single-trial velocity reconstruction accuracies r^2
395 as regressors of the decision parameters to assess the relationship between trial-
396 to-trial variations in EEG-velocity couplings and each model parameter.
397 Furthermore, to characterise the effect of active sensing movements on decision
398 formation, we also incorporated movement parameters in the HDDM framework.
399 Specifically, we computed the following movement parameters: a) the average
400 finger velocity (v_m) on each trial, b) the number of crossings (n_{cr}) between L and
401 R which is an indicator of the time it took participants to switch between the two
402 stimuli and c) the time participants spent exploring one of the two stimuli (here
403 we arbitrarily selected the low-amplitude stimulus on each trial, t_{low}) as an
404 indicator of exploration time. To understand how these movement parameters
405 affect the decision-making process and specifically whether they relate to a)
406 sensory processing and movement planning/execution (i.e. non-decision

407 processes) and/or b) evidence accumulation (i.e. decision processes) and/or c)
 408 the speed-accuracy trade-off adopted by the participants, we used these
 409 parameters as regressors for non-decision time, drift rate and decision boundary,
 410 as follows:

411

$$\tau = \beta_0 + \beta_1 * r^2 + \beta_v * v_m + \beta_{sw} * n_{cr} + \beta_{exp} * t_{low} \quad (3)$$

$$\delta = \gamma_0 + \gamma_1 * r^2 * s + \gamma_v * v_m + \gamma_{sw} * n_{cr} + \gamma_{exp} * t_{low} \quad (4)$$

$$\alpha = \theta_0 + \vartheta_1 * r^2 + \vartheta_v * v_m + \vartheta_{sw} * n_{cr} + \vartheta_{exp} * t_{low} \quad (5)$$

412 where τ, δ, α represent the single-trial non-decision time, drift rate and decision
 413 boundary respectively., Velocity reconstruction accuracy r^2 , mean finger velocity
 414 v_m , number of crossings n_{cr} and time spent exploring the lower amplitude
 415 stimulus t_{low} are the single-trial predictor variables with regression coefficients
 416 $\beta_i, \gamma_i, \delta_i$ respectively and $s = 0.1, 0.25, 0.5$ is the stimulus difference on each trial
 417 $k = 1, \dots, K$ of each participant $n = 1, \dots, N$. As per common practice, we modelled
 418 a linear relationship between drift rates and stimulus differences reflecting the
 419 dependence of the speed of information integration on the amount of evidence
 420 available (Palmer et al., 2005; Ratcliff and McKoon, 2008).

421 By using the above regression approach we were able to test the influence of
 422 the above EEG and movement parameters on each of the HDDM parameters.
 423 Thus, we tested different models in which the single-trial values of the above
 424 parameters were used as predictors for all combinations of the HDDM
 425 parameters (drift rate, non-decision time and decision boundary). To select the
 426 best-fitting model, we used the Deviance Information Criterion (DIC), a measure
 427 widely used for fit assessment and comparison of hierarchical models
 428 (Spiegelhalter et al., 2002). DIC selects the model that achieves the best trade-
 429 off between goodness-of-fit and model complexity. Lower DIC values favor
 430 models with the highest likelihood and least degrees of freedom.

431 *Statistical analysis of modeling results.* Posterior probability densities of each
432 regression coefficient were estimated using the sampling procedure described
433 above. Significantly positive (negative) effects were determined when >95% of
434 the posterior density was higher (lower) than 0. To take into account the
435 hierarchical structure of the model which estimated both population-level
436 distributions and participant-level distributions of the parameters, all statistical
437 tests at the population level were performed by contrasting the group-level
438 distributions (not the individual participant means) across sensory conditions.
439 This hierarchical statistical testing has been shown to reduce biases and actually
440 yield conservative effect sizes (Boehm et al., 2018).

441

442 *Partial Information Decomposition.* We then aimed to uncover whether the visual
443 (V) and haptic (H) neural representations of active sensing contained the same
444 information (redundancy) that is present in the multisensory (VH) representation
445 or to what extent their contributions are distinct (unique information) or
446 complementary (synergy). To achieve this, we used the Partial Information
447 Decomposition (PID) (Williams and Beer, 2010; Timme et al., 2014) applied to
448 the predictions of the finger velocity encoding models learned in the different
449 experimental conditions. PID provides an information theoretic approach to
450 compare the outputs of different predictive models that goes beyond simply
451 comparing accuracy to determine whether the different models share or convey
452 unique predictive information content (Daube et al., 2019b). PID extends the
453 concept of co-information (McGill, 1954), which is defined as follows:

454
$$I(VH;V;H) = I(VH;V) + I(VH;H) - I(VH;[V,H]) \quad (6)$$

455 where $I(X;Y)$ denotes the mutual information (MI) between variables X and Y.
456 MI is a nonparametric measure of dependence between two variables which has
457 the unique property that its effect size is additive (Shannon, 1948). Hence, co-
458 information (also called interaction information when defined with opposite sign)

459 quantifies the difference between the sum of the MI when each modality is
460 considered alone and the MI when the two modalities are observed together
461 (Park et al., 2018).

462 Positive values of this difference indicate that some information about the
463 predictions of the multisensory VH model is shared between the predictions
464 obtained from the models trained in the unisensory V and H conditions. I.e. there
465 are common or redundant representations of finger velocity in both V and H
466 conditions. Negative values of the interaction information indicate a super-
467 additive or synergistic interaction between the predictions of the V and H models,
468 i.e. the two models provide more information about the multisensory (VH)
469 prediction when observed together than would be expected from observing each
470 individually. However, interaction information measures the net difference
471 between synergy and redundancy in the system, thus it is possible to have zero
472 interaction information, even in the presence of redundant and synergistic
473 interactions that cancel out in the net value (Williams and Beer, 2010; Ince,
474 2017). This occurs because classic Shannon quantities cannot separate
475 redundant and synergistic contributions, which has led to a growing field
476 developing Partial Information Decomposition measures to address this
477 shortcoming.

478 To give a simple example of such a case, let us consider 3 variables, each
479 consisting of two bits (i.e. binary (0/1) variables with $p(0)=p(1)=0.5$). Let also
480 assume that the first bit is shared between all 3 variables and the second bit
481 follows the XOR distribution across the three variables. In this case, there is clear
482 redundancy and synergistic structure, but co-information / interaction information
483 is zero (Griffith and Koch, 2014).

484 More precisely, PID addresses this methodological problem by decomposing MI
485 into unique redundant and synergistic components, as follows:

$$486 \quad I(VH;V;H) = I_{uni}(VH;V) + I_{uni}(VH;H) + I_{red}(VH;V,H) + I_{syn}(VH;V,H) \quad (7)$$

487 where $I_{uni}(VH;V)$ is the part of the VH model predictions that can be explained
488 only from the V model predictions, $I_{uni}(VH;H)$ is the part of the VH model
489 predictions that can be explained only from the H model predictions,
490 $I_{red}(VH;V,H)$ is the part of the VH model predictions that is common (redundant)
491 to both the V and H model predictions and $I_{syn}(VH;V,H)$ is the extra (synergistic)
492 information about the VH model predictions that arises when both V and H
493 predictions are considered together. PID decomposes the joint mutual
494 information between two predictor signals (here the EEG activity predicted from
495 an encoding model trained in the unisensory V, H conditions) and a target signal
496 (here the EEG activity predicted from an encoding model trained in the
497 multisensory VH condition) into four terms: redundancy, the unique information in
498 each predictor, and synergy. Redundancy quantifies the information in the target
499 signal that is shared between the two predictor signals. Synergy quantifies
500 improvement in prediction of the target when both predictors are observed
501 together and represents information about the target signal which cannot be
502 obtained from the individual predictors separately.

503 To perform PID here, we used a recent implementation based on common
504 change in surprisal for Gaussian variables (Ince, 2017) which has been shown to
505 be effective when applied to neuroimaging data (Park et al., 2018; Daube et al.,
506 2019a).

507 To implement the above approach on our data, we used the recordings of the VH
508 condition where the two unisensory representations of active sensory experience
509 could be directly compared with the multisensory representation. We took the
510 velocity-encoding models obtained in each condition (V, H, VH) and applied them
511 to the VH data (see Eq. 3) to obtain the V, H and VH predictions of each EEG
512 sensor activity for all VH trials. Since the unisensory models (V, H) were fit in the
513 corresponding unisensory condition, they could only have learned a unisensory
514 representation, whereas the VH model learned a multisensory representation of

515 active sensing velocity. Thus, we applied PID for each participant separately to
516 predict the VH model predictions from the two unisensory V and H model
517 predictions, which enabled us to quantify the cross-modal interactions between
518 the two unisensory representations across all EEG sensors.

519 *Statistical analysis of PID results.* We performed this decomposition
520 independently for each EEG channel and obtained scalp maps for the four PID
521 terms (redundant information, unique information of A, unique information of V,
522 synergistic information) for each participant. To avoid over-fitting, we
523 implemented a 5-fold cross-validation procedure. We randomly split the VH data
524 into 5 subsets used 4 of them to learn the VH, V and H models and the held-out
525 set to perform the PID on. We repeated this process 5 times to obtain PID values
526 for all the VH data. To assess statistical significance of the obtained values, we
527 performed a permutation test. Specifically, we shuffled the target signal, i.e. the
528 VH model of active sensing, 1000 times while keeping the two predictor signals
529 (V and H models respectively) unchanged and applied PID to predict the VH
530 model surrogate data. Output values of the original PID decomposition were
531 considered significant if they exceeded the 99-th percentile of the distribution of
532 the surrogate data. Multiple comparisons were corrected for using FDR
533 (Genovese et al., 2002).

534
535
536
537
538
539
540
541
542
543
544

545 **Results**

546

547 We collected behavioral and EEG data while 14 participants actively
548 interrogated a 2-dimensional texture stimulus that differed in its amplitude in one
549 dimension (left - L vs right – R). Participants used visual information (V), haptic
550 information (H) or both types of sensory information simultaneously (VH) to make
551 a 2-alternative forced perceptual choice (2AFC), i.e., report (via a key press) as
552 quickly and as accurately as possible on which side (L or R) the texture stimulus
553 had higher amplitude (Fig. 1B). To sample information from both sides,
554 participants performed finger movements scanning the workspace of the
555 Pantograph before reaching a decision (Fig. 1C).

556 In the H condition, the Pantograph (see Materials and Methods for more
557 details on the device used to generate the stimuli) was programmed to produce
558 sinusoidal forces, which yielded the sensation of exploring a rough texture
559 surface (with different amplitudes between L and R (when participants moved
560 their index finger on the workspace of the Pantograph (see Fig.1B middle panel).
561 In the visual domain, participants were moving their fingers to reveal greyscale
562 stripes of different intensity/contrast between L and R (see Fig. 1B left panel). In
563 the VH condition, both the visual and haptic textures were congruently presented
564 wherever the participants moved their fingers (Fig. 1B right panel). Overall,
565 participants had to decide whether L or R had higher amplitude based on their
566 haptic (in H trials), visual (in V trials) or visuo-haptic (in VH trials) perception of
567 this virtual surface.

568

569 *Multisensory gain in behavioral performance*

570 Multisensory stimulation resulted in significantly higher discrimination
571 accuracy (91.5%±2.1% in VH vs 85.8%±2.2% in V and 86.3%±2.2% in H, two-
572 way ANOVA with factors condition and stimulus difference, $F(2,99)=5.64$,
573 $p<0.005$, see also slopes in the corresponding psychometric curves in Fig. 1D,

574 $PSE_v = 0.034 \pm 0.013, PSE_h = -0.001 \pm 0.009, PSE_{vh} = -0.019 \pm 0.007, slope_v =$
575 $2.397 \pm 0.2964, slope_h = 1.826 \pm 0.147, slope_{vh} = 3.001 \pm 0.2514$) compared to
576 the unisensory conditions (post-hoc t-tests, Bonferroni corrected, $p=0.009$ for V-
577 VH and $p=0.019$ for H-VH). Response times also reduced in VH ($4.11 \pm 0.30s$ vs
578 $4.41 \pm 0.31s$ in V and $4.25 \pm 0.29s$ in H, two-way ANOVA with factors condition and
579 stimulus difference, $F(2,99)=3.19, p=0.045$, see also corresponding cumulative
580 distribution functions in the three conditions, Fig. 1E). This result was significant
581 at the population level for VH vs V differences (post-hoc t-test, $p=0.021$,
582 Bonferroni corrected) but not VH vs H differences (post-hoc t-test, $p=0.066$,
583 Bonferroni corrected) in response times. As expected, we also found a main
584 effect of stimulus differences, with accuracy increasing ($F(2) = 91.82, p < 0.0001$)
585 and reaction times decreasing ($F(2) = 4.56, p < 0.02$) with larger stimulus
586 differences, respectively. There was no interaction between the sensory
587 condition and stimulus difference on either measure (accuracy: $F(4) = 0.66, p =$
588 0.62 ; reaction times: $F(4) = 0.05, p = 0.99$). Taken together, these results indicate
589 that multisensory information increased decision-making performance.

590

591 *Reconstruction of active sensing velocity from EEG recordings*

592 We then aimed to establish a relationship between brain activity and the
593 active sensory experience of the participants in each one of the three sensory
594 conditions. To this end, we performed a multivariate ridge regression (Crosse et
595 al., 2016) between the EEG data and the 1-d finger velocity data (on the x-axis)
596 to quantify neural encoding of sensorimotor behavior.

597 This analysis yielded the optimal linear combination of EEG channel
598 activations with time lags ranging between $[-200ms, 400ms]$ that approximated
599 the measured movement velocities. We found that reconstruction accuracy
600 r^2 was above chance level in all sensory conditions (all $p-values < 0.01$, Fig. 2B).
601 To obtain interpretable topographies of the neural activity underlying these EEG-
602 velocity couplings, we inverted the obtained velocity-decoding (backward)

603 models into velocity-encoding (forward) models (Parra et al., 2005; Haufe et al.,
604 2014). This revealed that centro-frontal locations (with positive weights) and
605 occipital locations (with negative weights) contributed most to velocity
606 reconstruction in the three sensory conditions with time lags ranging from 20 to
607 160 ms - see Fig. 2A showing the scalp topographies of the forward models and
608 Fig. 2C-D showing the corresponding temporal response functions (averaged
609 across frontal and occipital channels respectively) in the three sensory
610 conditions.

611

612

613 *Impact of active multi-sensing on the quality of perceptual evidence*

614 To characterise the relationship between the identified EEG-velocity
615 couplings and decision-making performance,, we employed a Hierarchical Drift
616 Diffusion Model (HDDM). In brief, the HDDM decomposes task performance (i.e.
617 accuracy and RT), into internal components of processing representing the rate
618 of evidence integration (drift rate, δ), the amount of evidence required to make a
619 choice (decision boundary separation, α), and the duration of other processes,
620 such as stimulus encoding and response production (non-decision time, τ).
621 Ultimately, by comparing the obtained values of all three core HDDM parameters
622 across the V, H and VH trials, we could associate any behavioral differences
623 resulting from the deployment of multisensory information (more accurate and
624 faster perceptual choices as in Fig. 1) to the constituent internal process reflected
625 by each model parameter.

626 Here, to obtain a mechanistic account of the formation of perceptual
627 decisions via the active sampling of (multi-)sensory information, we incorporated
628 the single-trial measures of brain-sensing- couplings (r^2 values) into the HDDM
629 parameter estimation (Fig. 3B). Specifically, we applied the obtained decoding
630 filters to the single-trial EEG data and computed velocity reconstruction
631 accuracies for each trial of each sensory condition (using a nested cross-

632 validation process, see Materials and Methods for more details). Then, as part of
633 the HDDM fitting process, we integrated these single-trial r^2 values in the
634 HDDM framework by using them as regressors of the three core HDDM
635 parameters (drift rate, non-decision time and decision boundary – see Materials
636 and Methods). The corresponding regression coefficients were estimated
637 together with the HDDM parameters thus enabling the assessment of the
638 relationship between trial-to-trial variations in EEG-velocity couplings and each
639 model parameter. We also used as regressors three movement parameters
640 (average velocity v_m , number of crossings between L and R n_{cr} and time spent
641 on the lower amplitude stimulus t_{low}) which served to dissociate the effect of the
642 exploratory movements (captured by these parameters) on decision formation
643 from the effect of the neural encoding of these active sensing movements
644 (captured by r^2).

645 We found that the best-fitting model (achieving the best complexity-
646 approximation trade-off as evaluated by the Deviance Information Criterion, Fig.
647 3A) was the one using r^2 as regressor of the drift rate only and n_{cr} , t_{low} as
648 regressors of non-decision time only (Figure 3B shows a graphical illustration of
649 the best-fitting model and Fig. 3C shows the model fitting of the accuracy and RT
650 data where bars represent actual data and lines represent model fits). The
651 means and confidence intervals of the estimated values of the three core HDDM
652 parameters are reported in Table 1. Crucially for our investigation here, the EEG-
653 velocity couplings r^2 were predictive of drift rates in single trials (regression
654 coefficients β_1 were larger than zero for all three sensory conditions,
655 $Prob(\gamma_1(V) > 0) > 0.97, Prob(\gamma_1(H) > 0) > 0.99, Prob(\gamma_1(VH) > 0) > 0.999$; Fig.
656 3D). Furthermore, the contribution of r^2 to drift rate was higher in VH trials
657 compared to V and H trials ($Prob(\gamma_1(VH) > \gamma_1(V)) > 0.95$ and $Prob(\gamma_1(VH) >$
658 $\gamma_1(H)) > 0.99$; Fig. 3D) indicating a multisensory enhancement of evidence
659 accumulation rates via an increased weighting of the EEG-velocity couplings in
660 the VH condition.

661 We then examined whether this multisensory gain could explain the
662 observed improvements in behavioral performance when multisensory
663 information is available. Indeed, this enhanced contribution of r^2 to drift rate was
664 predictive of multisensory improvements in behavioral performance. Specifically,
665 cross-participant differences in β_1 's across conditions correlated with the reported
666 increases in accuracy ($r = 0.58$, $p = 0.049$ for VH vs V and $r = 0.75$, $p = 0.005$
667 for VH vs H; Fig. 3F), suggesting that differences in accuracies across
668 participants were accounted for by the contributions of EEG-velocity couplings to
669 evidence accumulation. Thus, participants with greater drift rate amplification
670 achieved stronger enhancements in their behavioral performance as a result of
671 multisensory information available.

672 We also found that both switching time between the two stimuli as
673 captured by n_{cr} and exploration time spent on one of the two stimuli as captured
674 by t_{low} were predictive of non-decision time ($Prob(\beta_{sw} > 0) > 0.999$, $Prob(\beta_{exp} >$
675 $0) > 0.999$ for all V,H,VH; Fig. 3G-H) in single trials indicating that non-decision
676 processes (i.e. related to sensory processing and movement planning/execution)
677 are dependent on switching and exploration times. There was a positive cross-
678 participant correlation ($r = 0.695$, $p = 0.0121$) between β_{exp} and RT (averaged
679 across trials and sensory conditions) suggesting that participants with larger
680 contributions of exploration time to their non-decision times took longer to
681 respond (Fig 3I). However, we found no reliable difference in the corresponding
682 regression coefficients (β_{sw}, β_{exp}) between the three sensory conditions
683 ($Prob(\beta_{sw}(VH) > \beta_{sw}(V)) = 0.632$, $Prob(\beta_{sw}(VH) > \beta_{sw}(H)) = 0.843$,
684 $Prob(\beta_{exp}(VH) > \beta_{exp}(V)) = 0.107$, $Prob(\beta_{exp}(VH) > \beta_{exp}(H)) = 0.210$; Fig.
685 3G-H). There was also no difference in the decision boundaries in the three
686 sensory conditions ($Prob(\alpha(VH) > \alpha(V)) = 0.731$, $Prob(\alpha(VH) > \alpha(H)) > 0.804$;
687 Fig. 3E). These results indicate that neither the switching and exploration times

688 nor the amount of evidence required to make a decision were dependent on the
689 sensory condition.

690

691 *Quantifying multisensory interactions*

692 Having established that the neural encoding of the behavioral kinematics
693 is related to the multisensory gain in decision evidence, we then aimed to assess
694 how the neural representations of the two unisensory stimuli (V, H) interact to
695 form a multisensory representation. To this end, we employed Partial Information
696 Decomposition, which enables the quantification of cross-modal representational
697 interactions in the human brain (see Methods for details). Specifically, the PID
698 information theoretic framework quantifies the degree to which a) each
699 unisensory (V,H) representation contributes uniquely to the encoding of active
700 sensing behavior (unique V or H information), b) the two unisensory (V,H)
701 representations share information about active sensing (redundancy) and c) the
702 two unisensory (V,H) representations convey more information when observed
703 simultaneously (synergy). Here, we used PID to predict the forward (velocity-
704 encoding) VH model (target signal) from the two unisensory forward models V
705 and H (predictor signals). The decomposition revealed that the V model provided
706 unique information in right parieto-temporal locations whereas the H model
707 contributed uniquely in left prefrontal and parieto-occipital locations (Fig. 4A, all
708 p -values<0.01, FDR corrected). Crucially, we also found multisensory
709 interactions in the form of a) redundant effects in left prefrontal and parieto-
710 occipital electrodes and b) synergistic effects over left centro-parietal scalp (Fig.
711 4A, all ps <0.01, FDR corrected). Here, a redundant interaction means that the
712 representation of velocity is common to both the V and H modalities (Ince et al.,
713 2017; Park et al., 2018). A synergistic interaction means a better prediction of
714 the modelled multisensory response can be made when considering both the V
715 and the H representations together (rather than independently). That is,

716 knowledge of the simultaneous combination of the EEG signal predicted by V
717 and H models gives more information about the VH EEG signal.

718

719 *Multisensory accuracy scales with synergistic interactions*

720 Next we investigated the behavioral relevance of the identified cross-
721 modal interactions. In particular, we asked whether the identified synergistic
722 representation of the two modalities was predictive of behavioral performance
723 across participants. Indeed, we found a significant positive correlation (Pearson's
724 $R = 0.75$ and 0.72 , all $p < 0.01$) between synergy in both significant channels
725 (CP3 and C5) and accuracy in VH, suggesting that participants with more
726 synergistic representations at left centro-parietal electrodes achieved better
727 multisensory performance (Fig. 4B). This result suggests that synergy in
728 contralateral centro-parietal EEG signals modulates multisensory decision-
729 making behavior. Due to small sample size we cannot be sure this finding will
730 generalise, but nonetheless report it as an interesting exploratory finding.

731

732

733

734

735

736

737

738

739

740

741

742 **Discussion**

743

744 In this work, we coupled neural decoding of continuous sensorimotor behavior
745 with modeling of decision-making performance and a quantitative assessment of
746 crossmodal neural interactions to understand how the human brain forms
747 perceptual decisions via the active acquisition of multisensory evidence. We
748 showed that the neural encoding of active sensing modulates the decision

749 evidence regardless of the sensing modality. We further demonstrated that the
750 simultaneous sensing of different modalities enhances this neural coupling and
751 this enhancement drives the dynamics of active multisensory decisions. We
752 finally dissected the neural information conveyed by cross-modal interactions and
753 identified a potential neural mechanism supporting multisensory decisions.

754 Recent research on active sensing uncovered the strategies implemented by
755 humans to sample sensory information (Yang et al., 2016b). Here we
756 investigated this active sensing approach in a decision-making task using a
757 computational approach which decodes the neural activity that encodes
758 movement kinematics. Crucially, we made a first step in broadening this line of
759 research to a) include sensory information from multiple modalities and b) reveal
760 its neural underpinnings. These two developments enabled us to uncover the
761 different sensory representations of active sampling behavior in the human brain.

762 To achieve this, we implemented an informed cognitive modeling approach that
763 linked the neural correlates and the movement characteristics of active sensing
764 behavior with the cognitive processes involved in decision-making. Specifically,
765 we asked if decision-making depends on the neural representations of active
766 (multi-)sensing. To answer this question, we used a single-trial measure of the
767 neural encoding of active sensing behavior as predictor of decision-making
768 performance and found that, indeed, trial-to-trial fluctuations of the neural
769 representations of active sensing are predictive of the rate of evidence
770 accumulation for all three sensory conditions (V, H, VH). Crucially, we showed
771 that the multisensory (VH) representation of active sensing was a stronger
772 predictor of drift rate (Figure 3D) thus offering a neural link between active multi-
773 sensing and perceptual decision-making. We also split the motion profile into its
774 two main components, i.e. a) switching between the two alternative stimuli and b)
775 exploration within one particular stimulus and demonstrated that both
776 components were predictive of the duration of non-decision processes (Figure
777 3G-H), thus simply reflecting the time spent for movement planning and

778 execution and the consequent acquisition and encoding of sensory information.
779 These novel findings were only made possible by the use of an active multi-
780 sensing paradigm in a decision-making task and the joint cognitive modeling of
781 behavioral, neural and sensorimotor signals.

782 We then capitalized on the identified neural representations of active (multi-
783 sensing), to dissect cross-modal interactions in the human brain. To this end, we
784 employed PID, a recently developed rigorous methodology for the quantification
785 of information conveyed uniquely or jointly by different neural representations
786 (Williams and Beer, 2010; Timme et al., 2014; Ince, 2017). PID further
787 distinguishes between two types of interactions between the neural
788 representations of the two sensory modalities (V, H). A synergistic interaction
789 indicates that a better prediction of the multisensory neural response can be
790 made when the predicted values of the unimodal forward models for V and H are
791 considered jointly rather than independently. Our results suggest that this
792 synergistic interaction of the two neural representations correlates with
793 multisensory behavioural performance (Figure 4B). Instead, a redundant
794 interaction indicates that the two unimodal models provide the same information
795 about the multisensory condition, thus the multisensory response there is
796 common to both modalities (Park et al., 2018; Daube et al., 2019a). This
797 suggests that the underlying neural signals reflect a modality-invariant
798 representation.

799 As a result of this analysis, we were able to identify neural signals representing
800 these two types of interactions. Specifically, we found that EEG channels in
801 (parieto-)occipital and prefrontal areas carried redundant representations of the
802 two sensory streams, perhaps reflecting supramodal coding mechanisms of
803 active sensing (Figure 4A, redundancy). This finding is in line with previous
804 research assigning a multimodal role to occipital cortex (Lacey et al., 2007;
805 Murray et al., 2016) and suggesting that multisensory enhancements originate
806 from the sensory cortices (Kayser and Logothetis, 2007; Lakatos et al., 2007;

807 Lewis and Noppeney, 2010). Specifically, recent research involved the visual
808 cortex in audio-visual interactions (Mishra et al., 2007; Cao et al., 2019; Rohe et
809 al., 2019) as well as tactile perception and visuo-haptic interactions (Lucan et al.,
810 2010; Sathian, 2016; Gaglianese et al., 2020). In agreement with the above, here
811 we also found unique H information in parieto-occipital electrodes. Concerning
812 the prefrontal cortex (PFC), recent evidence assigned to it a modality-general
813 role in arbitrating between segregation or fusion of sensory evidence from
814 different modalities (Cao et al., 2019). Thus, the involvement of the PFC in the
815 regulation of adaptive multisensory behaviors in general (Koechlin and
816 Summerfield, 2007; Donoso et al., 2014; Tomov et al., 2018) and perceptual
817 decisions in particular (Heekeren et al., 2006; Philiastides et al., 2011; Rahnev et
818 al., 2016; Sterzer, 2016) makes it a likely contributor to the formation of the most
819 appropriate sensory representation that drives decision-making behavior. In other
820 words, the PFC may support a mechanism gauging candidate (multisensory or
821 unisensory) representations for selecting among multiple strategies to solve the
822 task at hand (Calvert, 2001; Hein et al., 2007; Noppeney et al., 2010; Cao et al.,
823 2019). Our active multi-sensing task requires participants to continuously weigh
824 different sensing strategies and refine their scanning patterns to maximize
825 information gain. Hence, the PFC may capitalise on multisensory information
826 (when of benefit) to support such flexible behavior striking a balance between
827 sampling more evidence and committing to a choice.

828 The above findings are consistent with our previous study focusing on the tactile
829 modality, which attributed a sensory processing function to occipital cortex
830 (specifically localized to the lateral occipital complex) and a decision formation
831 function to right prefrontal cortex (middle frontal gyrus) (Delis et al., 2018). Taken
832 together with the current results, our findings suggest these two brain areas may
833 play a crossmodal role in supporting active perception and decision-making.
834 Overall, our work adds to the existing literature on multi-sensory interactions by
835 quantifying how sensory representations interact to encode active sensing
836 behaviors.

837 More importantly, here we revealed a novel functional role for contralateral
838 centro-parietal signals in active visuo-haptic decisions. We found that brain
839 signals over left centro-parietal scalp locations showed stronger encoding of
840 active sensing when the two sensory streams were available (Figure 4A,
841 synergy), thus possibly representing a neural mechanism of multisensory
842 integration. In line with the ongoing debate on the multisensory nature of primary
843 sensory cortices (Ghazanfar and Schroeder, 2006; Liang et al., 2013), cross-
844 modal visuo-haptic interactions leading to enhanced neural representations have
845 been found in the primary somatosensory cortex (S1) (Zhou and Fuster, 2000;
846 Dionne et al., 2010). Here we further characterised these interactions as carrying
847 super-additive/synergistic representations of the active multi-sensory experience
848 and demonstrated that they are related to the accuracy of active multisensory
849 judgments.

850 It is also worth noting that our results do not rule out the possibility that
851 other brain areas – not directly related to active sensing - may contribute to
852 regulating the speed and accuracy of active multisensory decisions. In fact,
853 recent research breakthroughs have explained the development of multisensory
854 representations from different sensory streams in the human brain (Aller and
855 Noppeney, 2019; Cao et al., 2019; Rohe et al., 2019). Furthermore, recent
856 studies have started to investigate how the interactions between sensory
857 representations shape decision formation (Bizley et al., 2016; Franzen et al.,
858 2020; Mercier and Cappe, 2020).

859 Our primary aim here was to provide the missing link between the active
860 acquisition of multisensory evidence and its transformation to choice. Overall, our
861 findings validated the hypotheses that a) active sensing guides decision
862 formation via evidence sampling and accumulation and b) multisensory
863 information spurs perceptual decisions by enhancing the neural encoding of
864 active behaviors. Our information-theoretic analysis also revealed the neural
865 substrates of multisensory interactions in the human brain that support active
866 multisensory perception. Ultimately, we identified and characterised a set of

867 human brain signals that underpin multisensory judgements by subserving an
868 enhancement of the neural encoding of active perception when multisensory
869 information is available.

870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898 **References**
899

900 Aller M, Noppeney U (2019) To integrate or not to integrate: Temporal dynamics of hierarchical
901 Bayesian causal inference. *PLoS Biol* 17:e3000210.
902 Angelaki DE, Gu Y, DeAngelis GC (2009) Multisensory integration: psychophysics,
903 neurophysiology, and computation. *Curr Opin Neurobiol* 19:452-458.
904 Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind
905 deconvolution. *Neural Comput* 7:1129-1159.
906 Bizley JK, Jones GP, Town SM (2016) Where are multisensory signals combined for perceptual
907 decision-making? *Curr Opin Neurobiol* 40:31-37.

908 Boehm U, Marsman M, Matzke D, Wagenmakers EJ (2018) On the importance of avoiding
909 shortcuts in applying cognitive models to hierarchical data. *Behav Res Methods*
910 50:1614-1631.

911 Calvert GA (2001) Crossmodal processing in the human brain: insights from functional
912 neuroimaging studies. *Cereb Cortex* 11:1110-1123.

913 Campion G, Wang Q, Hayward V (2005) The Pantograph Mk-II: A haptic instrument. 2005
914 IEEE/Rsj International Conference on Intelligent Robots and Systems, Vols 1-4:723-728.

915 Cao Y, Summerfield C, Park H, Giordano BL, Kayser C (2019) Causal Inference in the Multisensory
916 Brain. *Neuron* 102:1076-1087 e1078.

917 Chandrasekaran C (2017) Computational principles and models of multisensory integration. *Curr*
918 *Opin Neurobiol* 43:25-34.

919 Crosse MJ, Di Liberto GM, Bednar A, Lalor EC (2016) The Multivariate Temporal Response
920 Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous
921 Stimuli. *Frontiers in Human Neuroscience* 10.

922 Daube C, Ince RAA, Gross J (2019a) Simple Acoustic Features Can Explain Phoneme-Based
923 Predictions of Cortical Responses to Speech. *Curr Biol* 29:1924-1937 e1929.

924 Daube C, Giordano BL, Ince RAA, Gross J (2019b) Quantitatively Comparing Predictive Models
925 with the Partial Information Decomposition. In: 2019 Conference on Cognitive
926 Computational Neuroscience. Berlin, Germany.

927 Delis I, Dmochowski JP, Sajda P, Wang Q (2018) Correlation of neural activity with behavioral
928 kinematics reveals distinct sensory encoding and evidence accumulation processes
929 during active tactile sensing. *Neuroimage* 175:12-21.

930 Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG
931 dynamics including independent component analysis. *Journal of Neuroscience Methods*
932 134:9-21.

933 Di Liberto GM, O'Sullivan JA, Lalor EC (2015) Low-Frequency Cortical Entrainment to Speech
934 Reflects Phoneme-Level Processing. *Current Biology* 25:2457-2465.

935 Dionne JK, Meehan SK, Legon W, Staines WR (2010) Crossmodal influences in somatosensory
936 cortex: Interaction of vision and touch. *Hum Brain Mapp* 31:14-25.

937 Donoso M, Collins AG, Koechlin E (2014) Human cognition. *Foundations of human reasoning in*
938 *the prefrontal cortex. Science* 344:1481-1486.

939 Drugowitsch J, DeAngelis GC, Klier EM, Angelaki DE, Pouget A (2014) Optimal multisensory
940 decision-making in a reaction-time task. *Elife* 3.

941 Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically
942 optimal fashion. *Nature* 415:429-433.

943 Forstmann BU, Ratcliff R, Wagenmakers EJ (2016) Sequential Sampling Models in Cognitive
944 Neuroscience: Advantages, Applications, and Extensions. *Annu Rev Psychol* 67:641-666.

945 Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, Badre D (2015) fMRI and EEG
946 predictors of dynamic decision parameters during human reinforcement learning. *J*
947 *Neurosci* 35:485-494.

948 Franzen L, Delis I, De Sousa G, Kayser C, Philiastides MG (2020) Auditory information enhances
949 post-sensory visual evidence during rapid multisensory decision-making. *Nat Commun*
950 11:5440.

951 Gaglianese A, Branco MP, Groen IIA, Benson NC, Vansteensel MJ, Murray MM, Petridou N,
952 Ramsey NF (2020) Electrocorticography Evidence of Tactile Responses in Visual Cortices.
953 *Brain Topogr* 33:559-570.

954 Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional
955 neuroimaging using the false discovery rate. *Neuroimage* 15:870-878.

956 Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? *Trends Cogn Sci*
957 10:278-285.

958 Gottlieb J, Oudeyer PY (2018) Towards a neuroscience of active sampling and curiosity. *Nat Rev*
959 *Neurosci* 19:758-770.

960 Griffith V, Koch C (2014) Quantifying Synergistic Mutual Information. In: *Guided Self-*
961 *Organization: Inception. Emergence, Complexity and Computation* (M. P, ed). Berlin,
962 Heidelberg: Springer.

963 Haufe S, Meinecke F, Gorgen K, Dahne S, Haynes JD, Blankertz B, Biessmann F (2014) On the
964 interpretation of weight vectors of linear models in multivariate neuroimaging.
965 *Neuroimage* 87:96-110.

966 Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for
967 perceptual decision-making in the human brain. *Nature* 431:859-862.

968 Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG (2006) Involvement of human
969 left dorsolateral prefrontal cortex in perceptual decision making is independent of
970 response modality. *Proceedings of the National Academy of Sciences of the United*
971 *States of America* 103:10023-10028.

972 Hein G, Doehrmann O, Muller NG, Kaiser J, Muckli L, Naumer MJ (2007) Object familiarity and
973 semantic congruency modulate responses in cortical audiovisual integration areas. *J*
974 *Neurosci* 27:7881-7887.

975 Ince RA, Giordano BL, Kayser C, Rousset GA, Gross J, Schyns PG (2017) A statistical framework
976 for neuroimaging data analysis based on mutual information estimated via a gaussian
977 copula. *Hum Brain Mapp* 38:1541-1573.

978 Ince RAA (2017) Measuring multivariate redundant information with pointwise common change
979 in surprisal. *Entropy* 19:318.

980 Juavinett AL, Erlich JC, Churchland AK (2018) Decision-making behaviors: weighing ethology,
981 complexity, and sensorimotor compatibility. *Curr Opin Neurobiol* 49:42-50.

982 Kayser C, Logothetis NK (2007) Do early sensory cortices integrate cross-modal information?
983 *Brain Struct Funct* 212:121-132.

984 Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive
985 function. *Trends Cogn Sci* 11:229-235.

986 Kruschke JK (2010) What to believe: Bayesian methods for data analysis. *Trends in Cognitive*
987 *Sciences* 14:293-300.

988 Lacey S, Campbell C, Sathian K (2007) Vision and touch: multiple or multisensory representations
989 of objects? *Perception* 36:1513-1521.

990 Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and
991 multisensory interaction in primary auditory cortex. *Neuron* 53:279-292.

992 Lewis R, Noppeney U (2010) Audiovisual synchrony improves motion discrimination via
993 enhanced connectivity between early visual and auditory areas. *J Neurosci* 30:12329-
994 12339.

995 Liang M, Mouraux A, Hu L, Iannetti GD (2013) Primary sensory cortices contain distinguishable
996 spatial patterns of activity for each sense. *Nat Commun* 4:1979.

997 Lucan JN, Foxe JJ, Gomez-Ramirez M, Sathian K, Molholm S (2010) Tactile shape discrimination
998 recruits human lateral occipital complex during early perceptual processing. *Hum Brain*
999 *Mapp* 31:1813-1821.

1000 McGill WJ (1954) Multivariate information transmission. *Psychometrika*:97-11.
1001 Mercier MR, Cappe C (2020) The interplay between multisensory integration and perceptual
1002 decision making. *Neuroimage* 222:116970.
1003 Mishra J, Martinez A, Sejnowski TJ, Hillyard SA (2007) Early cross-modal interactions in auditory
1004 and visual cortex underlie a sound-induced visual illusion. *J Neurosci* 27:4120-4131.
1005 Murray MM, Thelen A, Thut G, Romei V, Martuzzi R, Matusz PJ (2016) The multisensory
1006 function of the human primary visual cortex. *Neuropsychologia* 83:161-169.
1007 Musall S, Urai AE, Sussillo D, Churchland AK (2019) Harnessing behavioral diversity to
1008 understand neural computations for cognition. *Curr Opin Neurobiol* 58:229-238.
1009 Najafi F, Churchland AK (2018) Perceptual Decision-Making: A Field in the Midst of a
1010 Transformation. *Neuron* 100:453-462.
1011 Noppeney U, Ostwald D, Werner S (2010) Perceptual decisions formed by accumulation of
1012 audiovisual evidence in prefrontal cortex. *J Neurosci* 30:7434-7446.
1013 Nunez MD, Vandekerckhove J, Srinivasan R (2017) How attention influences perceptual decision
1014 making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of*
1015 *Mathematical Psychology* 76:117-130.
1016 Palmer J, Huk AC, Shadlen MN (2005) The effect of stimulus strength on the speed and accuracy
1017 of a perceptual decision. *J Vis* 5:376-404.
1018 Park H, Ince RAA, Schyns PG, Thut G, Gross J (2018) Representational interactions during
1019 audiovisual speech entrainment: Redundancy in left posterior superior temporal gyrus
1020 and synergy in left motor cortex. *PLoS Biol* 16:e2006558.
1021 Parra L, Alvino C, Tang A, Pearlmutter B, Yeung N, Osman A, Sajda P (2002) Linear spatial
1022 integration for single-trial detection in encephalography. *Neuroimage* 17:223-230.
1023 Parra LC, Spence CD, Gerson AD, Sajda P (2005) Recipes for the linear analysis of EEG.
1024 *Neuroimage* 28:326-341.
1025 Philiastides MG, Auksztulewicz R, Heekeren HR, Blankenburg F (2011) Causal role of dorsolateral
1026 prefrontal cortex in human perceptual decision making. *Curr Biol* 21:980-983.
1027 Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs
1028 sampling In: 3rd International Workshop on Distributed Statistical Computing.
1029 Rahnev D, Nee DE, Riddle J, Larson AS, D'Esposito M (2016) Causal evidence for frontal cortex
1030 organization for perceptual decision making. *Proc Natl Acad Sci U S A* 113:6059-6064.
1031 Raposo D, Sheppard JP, Schrater PR, Churchland AK (2012) Multisensory decision-making in rats
1032 and humans. *J Neurosci* 32:3726-3735.
1033 Ratcliff R, McKoon G (2008) The diffusion decision model: Theory and data for two-choice
1034 decision tasks. *Neural Computation* 20:873-922.
1035 Ratcliff R, Childers R (2015) Individual Differences and Fitting Methods for the Two-Choice
1036 Diffusion Model of Decision Making. *Decision (Wash D C)* 2015.
1037 Rohe T, Ehlis AC, Noppeney U (2019) The neural dynamics of hierarchical Bayesian causal
1038 inference in multisensory perception. *Nat Commun* 10:1907.
1039 Sathian K (2016) Analysis of haptic information in the cerebral cortex. *J Neurophysiol* 116:1795-
1040 1806.
1041 Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of Active Sensing
1042 and perceptual selection. *Curr Opin Neurobiol* 20:172-176.
1043 Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379-423 &
1044 623-656.

1045 Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model
1046 complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical*
1047 *Methodology* 64:583-616.

1048 Sterzer P (2016) Moving forward in perceptual decision making. *Proc Natl Acad Sci U S A*
1049 113:5771-5773.

1050 Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD (1992) Testing for Nonlinearity in Time-
1051 Series - the Method of Surrogate Data. *Physica D* 58:77-94.

1052 Timme N, Alford W, Flecker B, Beggs JM (2014) Synergy, redundancy, and multivariate
1053 information measures: an experimentalist's perspective. *J Comput Neurosci* 36:119-140.

1054 Tomov MS, Dorfman HM, Gershman SJ (2018) Neural Computations Underlying Causal Structure
1055 Learning. *J Neurosci* 38:7143-7157.

1056 Turner BM, van Maanen L, Forstmann BU (2015) Informing Cognitive Abstractions Through
1057 Neuroimaging: The Neural Drift Diffusion Model. *Psychological Review* 122:312-336.

1058 Wabersich D, Vandekerckhove J (2014) Extending JAGS: A tutorial on adding custom
1059 distributions to JAGS (with a diffusion model example). *Behavior Research Methods*
1060 46:15-28.

1061 Wiecki TV, Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-
1062 Diffusion Model in Python. *Front Neuroinform* 7:14.

1063 Williams PL, Beer RD (2010) Nonnegative Decomposition of Multivariate Information.
1064 arXiv:10042515v1.

1065 Winkler I, Haufe S, Tangermann M (2011) Automatic classification of artifactual ICA-components
1066 for artifact removal in EEG signals. *Behav Brain Funct* 7:30.

1067 Yang SC, Lengyel M, Wolpert DM (2016a) Active sensing in the categorization of visual patterns.
1068 *Elife* 5.

1069 Yang SCH, Wolpert DM, Lengyel M (2016b) Theoretical perspectives on active sensing. *Current*
1070 *Opinion in Behavioral Sciences* 11:100-108.

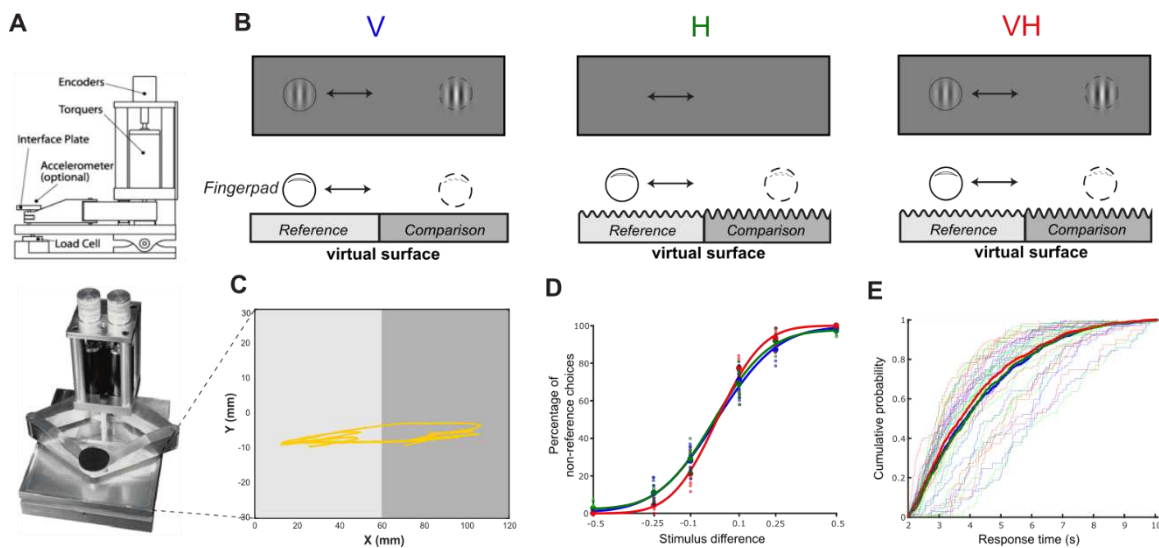
1071 Zhou YD, Fuster JM (2000) Visuo-tactile cross-modal associations in cortical somatosensory cells.
1072 *Proc Natl Acad Sci U S A* 97:9777-9782.

1073

1074

1075
1076

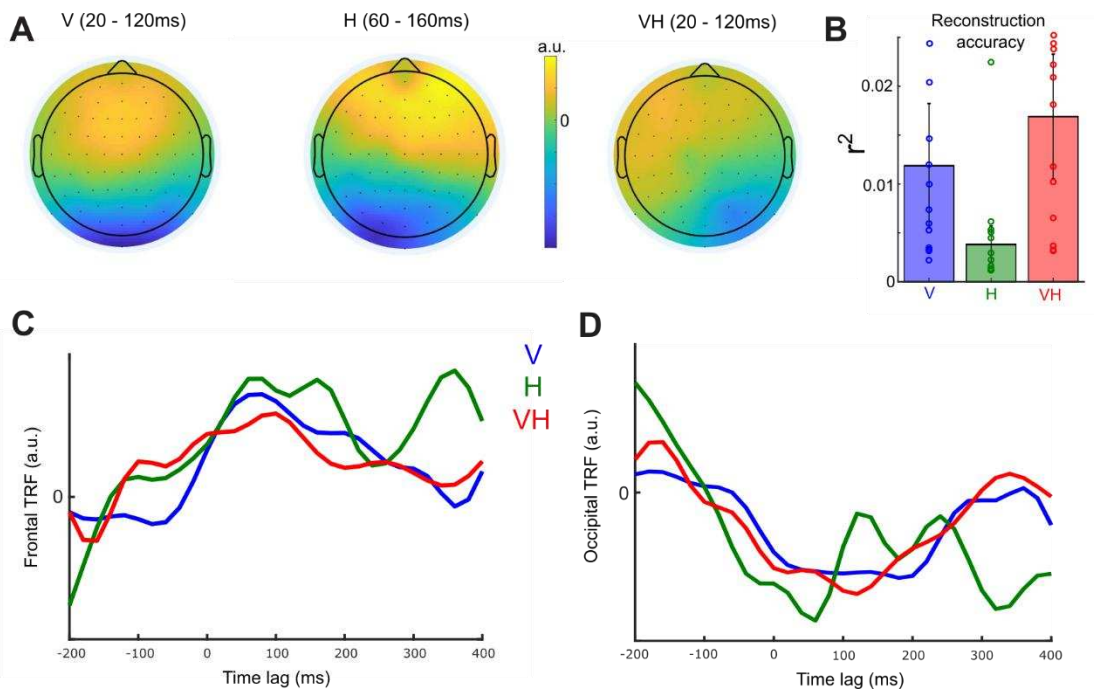
Figures and Tables



1077

1078 **Figure 1. Experimental design and behavioral results.** A. The Pantograph is a
1079 haptic device used to render virtual surfaces that can be actively sensed. Top:
1080 the parts of the Pantograph shown from a lateral view. Participants placed their
1081 index finger on the interface plate. Bottom: The Pantograph device used in this
1082 experiment. B. The stimulus in the three sensory conditions. We programmed the
1083 Pantograph to generate a virtual grating texture. The workspace was split into
1084 two subspaces (left - L and right - R) that differed in the amplitude of the virtual
1085 surface that the participants actively sensed. One of the two sides (randomly
1086 assigned) had the reference amplitude (equal to 1) and the other had the
1087 comparison amplitude that varied on each trial taking one of the values: 0.5,
1088 0.75, 0.9, 1.1, 1.25, and 1.5. Participants performed the task using visual
1089 information only (V), haptic information only (H) or the two sensory streams
1090 together (VH). Amplitude of the stimulus in the haptic domain (H) was translated
1091 as contrast in the visual domain (V). Crucially, to match the H condition, only a
1092 moving dot following the participant's finger was revealed on the screen in V. C.
1093 Index finger trajectory indicating the scanning pattern of the virtual texture in one
1094 trial. On this trial, the participant actively sensed the left subspace first, then
1095 moved to the right subspace and explored it before coming back to the left
1096 subspace again and reporting their choice. D. Psychometric curves indicating the
1097 percentage of non-reference choices for all three sensory conditions (V in blue, H
1098 in green, VH in red) and for all stimulus differences. Large dots represent
1099 average percentage of choices across participants and smaller dots represent
1100 individual participant means. Data are fit using cumulative Gaussian functions. E.
1101 Cumulative distributions (CDF) of response times for all three sensory conditions
1102 V in blue, H in green, VH in red) across all trials of all participants. Thick lines
1103 indicate CDFs across all participant data and thin lines indicate individual
1104 participant CDFs for each sensory condition.

1105
1106
1107
1108
1109
1110
1111
1112
1113
1114



1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132

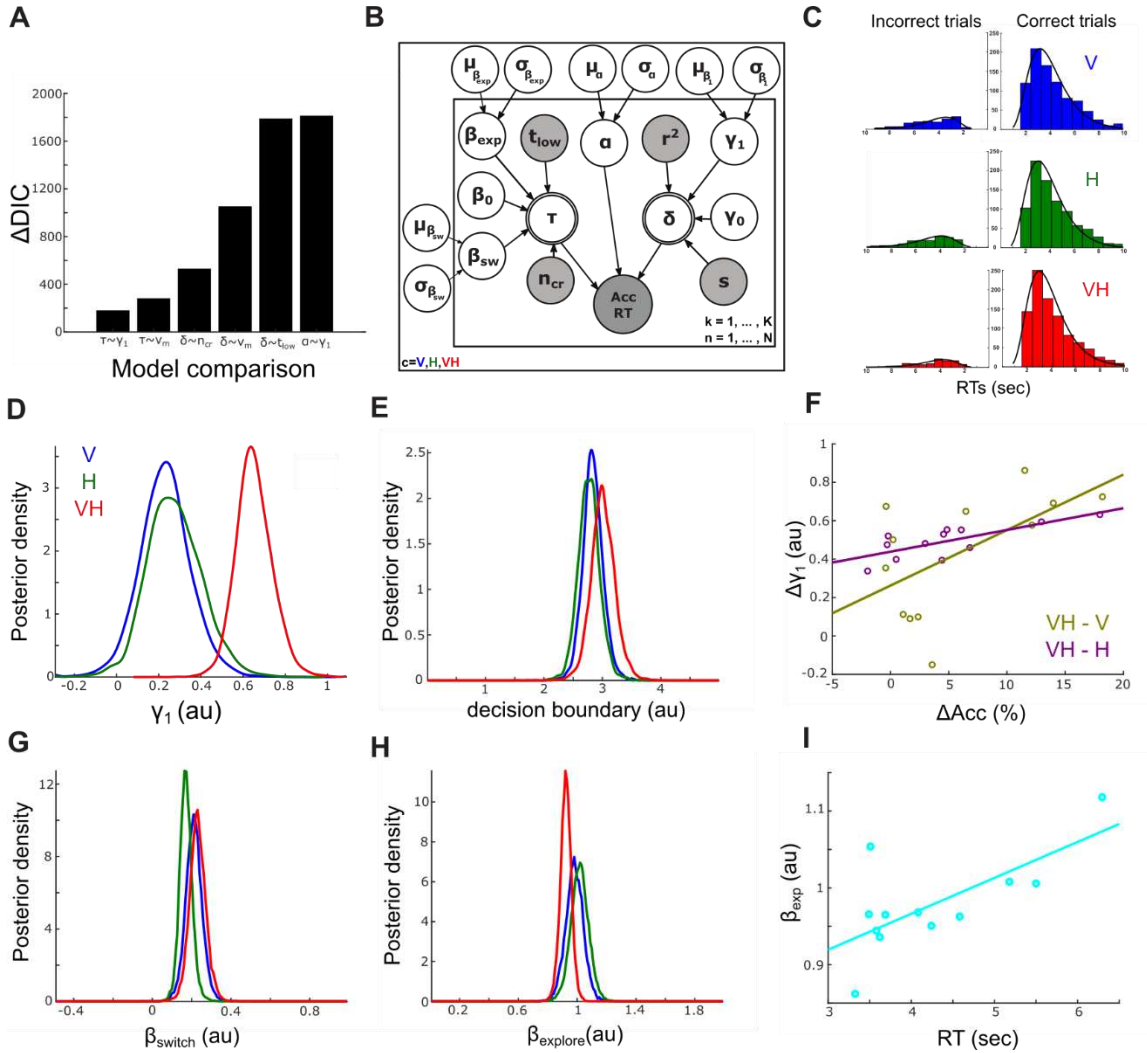
Figure 2. Results of velocity reconstruction analysis using EEG signals. A. Scalp topographies of the forward models representing neural encoding of instantaneous finger velocity for the three sensory conditions. The presented scalp maps show velocity-encoding EEG signals averaged over the following time windows: [20,120]ms lags between velocity and EEG for V and VH and [60,160]ms lags for H. B. Accuracy of the velocity reconstruction from the EEG signals measured using the squared correlation coefficient (r^2) between the original and the approximated velocity profile in the three sensory conditions (V in blue, H in green, VH in red). Bars represent means across participants and errorbars represent standard errors (sem). Dots represent individual participant data. C-D. Temporal response functions (TRFs) of the velocity-encoding EEG activity in the three sensory conditions (V in blue, H in green, VH in red) averaged over frontal electrodes (in C) and over occipital electrodes (in D).

1133
1134

Table1: Estimated values of the three core HDDM parameters for the best-fitting model

Parameter	Mean	Confidence Interval (5%)	Confidence Interval (95%)
Drift rate (δ)	0.897	0.628	1.162
Non-decision time (τ)	2.897	2.710	3.045
Decision boundary (α)	2.853	2.501	3.256

1135
1136
1137
1138
1139
1140
1141
1142
1143
1144

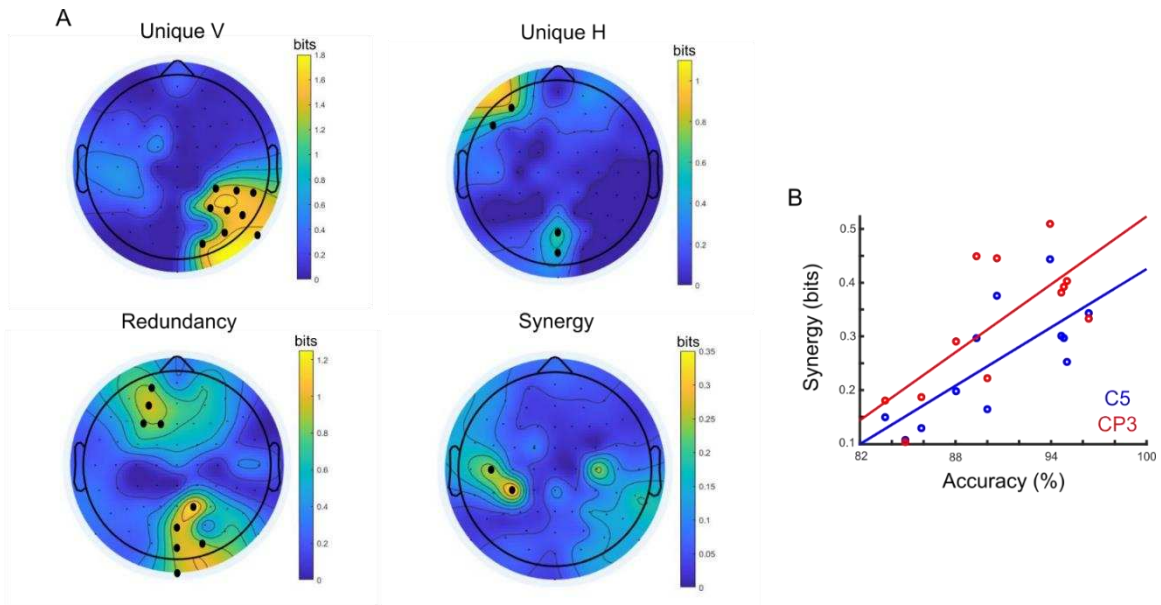


1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160

Figure 3. Informed modeling of decision-making behavior. A. Comparison of the best-fitting model (with r^2 as a regressor of drift rate δ only and n_{cr} , t_{low} as regressors of non-decision time τ only) with alternate models using the Deviance Information Criterion (DIC). Positive ΔDIC ($DIC_{\text{model}} - DIC_{\text{optimal}}$) values for all six models indicate that the model of choice achieved a better trade-off between goodness-of-fit and number of free parameters. B. Graphical representation showing hierarchical estimation of HDDM parameters. Round nodes represent continuous random variables and double-bordered nodes represent variables defined in terms of other variables. Shaded nodes represent recorded or computed signals, i.e. single-trial behavioral data (accuracy, RT and stimulus differences s), EEG-velocity couplings (r^2) and kinematic parameters (n_{cr} , t_{low}). Parameters are modelled as Gaussian random variables with inferred means μ and variances σ^2 . Plates denote that multiple random variables share the same

1161 parents and children. The outer plate is over sensory conditions (V,H, VH) and
1162 the inner plate is over all trials (K) and participants (N). C. Behavioral RT
1163 distributions are shown as histograms for each sensory condition (V in blue, H in
1164 green, VH in red) for correct (right) and incorrect (left) trials together with the
1165 HDDM fits (black lines). Higher histogram values on the right indicate higher
1166 proportion of correct choices. D. Posterior distributions of regression coefficients
1167 (γ_1) of the EEG-velocity couplings (r^2), as predictors of the drift rate (δ) of the
1168 HDDM shown in A. The three coloured curves indicate posterior distributions for
1169 the three sensory conditions (blue – V, green – H, red – VH). E. Posterior
1170 distributions of decision boundaries for the three sensory conditions (blue – V,
1171 green – H, red – VH). F. Cross-participant correlation of differences in choice
1172 accuracy (ΔAcc - x-axis) and differences in β_1 ($\Delta\beta_1$ – y-axis) between the
1173 multisensory (VH) and the two unisensory (V,H) conditions (VH-V in yellow, VH-
1174 H in purple). G. Posterior distributions of regression coefficients (β_{sw}) of the
1175 number of crossings between L and R (n_{cr}), as predictor of non-decision time (τ)
1176 of the HDDM shown in A. H. Posterior distributions of regression coefficients
1177 (β_{exp}) of the time spent on the low-amplitude stimulus (t_{low}), as predictor of non-
1178 decision time (τ) of the HDDM shown in A. I. Cross-participant correlation of
1179 average response times across trials and sensory conditions (x-axis) and β_{exp} (y-
1180 axis).

1181
1182
1183
1184
1185
1186
1187
1188



1189
 1190 **Figure 4. Neural representations and cross-modal interactions.** A. Results of
 1191 PID applied to predict the multisensory (VH) model of active sensing from the two
 1192 unisensory (V and H) models. Dots on the scalp topographies indicate the EEG
 1193 channels that provide significant ($p < 0.01$, FDR corrected) visual unique (top left),
 1194 haptic unique (top right), redundant (bottom left) and synergistic (bottom right)
 1195 neural information respectively. B. Across-subject correlation between synergy
 1196 in the two significant EEG channels (CP3 in red and C5 in blue) and choice
 1197 accuracy in the VH condition.
 1198

1199

1200 **Author Contributions:** Conceptualization: ID, PS, QW
 1201 Methodology: ID, RAAI, PS, QW
 1202 Investigation: ID
 1203 Supervision: PS, QW
 1204 Writing—original draft: ID
 1205 Writing—review & editing: ID, RAAI, PS, QW
 1206

1207

1208