Neural population dynamics underlying expected value computation

Hiroshi Yamada*[123], Yuri Imaizumi[4], Masayuki Matsumoto[123]

[1]Division of Biomedical Science, Faculty of Medicine, University of Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8577, Japan.

[2]Graduate School of Comprehensive Human Sciences, University of Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8577, Japan

[3]Transborder Medical Research Center, University of Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8577, Japan

[4]Medical Sciences, University of Tsukuba, 1-1-1 Tenno-dai, Tsukuba, Ibaraki 305-8577, Japan.

*Address correspondence to

Hiroshi Yamada, Ph.D.

Division of Biomedical Science, Faculty of Medicine, University of Tsukuba

1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577 Japan.

TEL/FAX: 81-29-853-6013, e-mail: h-yamada@md.tsukuba.ac.jp

Abstract

Computation of expected values, i.e., probability times magnitude, seems to be a dynamic integrative process performed in the brain for efficient economic behavior. However, neural dynamics underlying this computation remain largely unknown. We examined (1) whether four core reward-related regions detect and integrate the probability and magnitude cued by numerical symbols and (2) whether these regions have different dynamics in the integrative process. Extractions of mechanistic structure of neural population signal demonstrated that expected-value signals simultaneously arose in central part of orbitofrontal cortex (cOFC, area 13m) and ventral striatum (VS). These expected-value signals were incredibly stable in contrast to weak and unstable signals in dorsal striatum and medial OFC. Notably, temporal dynamics of these stable expected-value signals were unambiguously distinct: sharp and gradual signal evolutions in cOFC and VS, respectively. These intimate dynamics suggest that cOFC and VS compute the expected-values with unique time constants, as distinct, partially overlapping processes.

Impact Statement

Our empirical study on neural population dynamics suggests that the orbitofrontal cortex and ventral striatum co-operate on expected-value computation with unique time constants, as distinct, partially overlapping processes.

Introduction

Economic behavior requires a reliable perception of the world for maximizing benefit (Houthakker, 1950; Samuelson, 1950; Savage, 1954; Von Neumann & Morgenstern, 1944). Such maximization is achieved by computation of expected values (i.e. probability times magnitude) (Glimcher et al., 2008), which seems to be a dynamic process to detect and integrate the probability and magnitude to yield expected value signals in the brain. Indeed, humans and animals behave as if they compute the expected values (Glimcher et al., 2008; Kahneman & Tversky, 1979; Stephens & Krebs, 1986). One salient example, discovered over a century ago and repeatedly measured, is human economic behavior, in which a series of models originating from the standard theory of economics (Von Neumann & Morgenstern, 1944) has been developed to describe efficient economic behavior. Despite the ubiquity of this phenomenon, neural dynamics underlying the expected-value computation has not been identified.

Substantial previous researches have shown that many regions of the brain process rewards in terms of signaling the probability and/or magnitude mostly during economic decision makings (Barraclough et al., 2004; Eshel et al., 2016; Hirokawa et al., 2019; Lopatina et al., 2016; Ma & Jazayeri, 2014; Platt & Glimcher, 1999; Roesch et al., 2009; Rudebeck & Murray, 2014; Tobler et al., 2005; Xie & Padoa-Schioppa, 2016; Yamada et al., 2018; Zhou et al., 2019). In these previous literature, the computation of expected values are hypothesized to be processed by these regions without their neural dynamics, in line with an expected value theory shared across multiple disciplines (Glimcher et al., 2008; Stephens & Krebs,

1986; Sutton & Barto, 1998; Von Neumann & Morgenstern, 1944). Neuroimaging studies in humans and non-human primates also have suggested that multiple brain regions in the reward circuitry (Haber & Knutson, 2010) are involved in this computational process (Fouragnan et al., 2019; Howard et al., 2015; Howard & Kahnt, 2017; Hsu et al., 2009; Levy & Glimcher, 2012; O'Doherty et al., 2004; Papageorgiou et al., 2017; Tom et al., 2007), although the underlying neural mechanism has not been elucidated because of the limited time resolution of current neuroimaging techniques (Goense & Logothetis, 2008; Milham et al., 2018). These brain regions may employ the expected-value computation, however, none have captured and compared the temporal aspects of neural activities in the multiple brain regions. These observations have led us to test the hypothesis that neural population dynamics within subsecond-order time resolutions (Chen & Stuphorn, 2015; Churchland et al., 2012; Mante et al., 2013; Murray et al., 2017; Takei et al., 2017) play key roles in the expected-value computation, that is the detection and integration of the probability and magnitude on multiple neural population ensembles.

We targeted reward-related cortical and subcortical structures of non-human primates (Haber & Knutson, 2010) including the central orbitofrontal cortex (cOFC, area 13M), the medial orbitofrontal cortex, (mOFC, area 14O), the dorsal striatum (DS), and ventral striatum (VS). We used a recently developing mathematical approach, called state space analysis (Chen & Stuphorn, 2015; Churchland et al., 2012; Mante et al., 2013; Murray et al., 2017), to test how the expected-values computation is processed within each of the four neural population ensembles in

89 an order of 10^{-2}-second time resolution. We found that the cOFC and VS neural
90 populations coincidently detect and integrate the probability and magnitude
91 information to yield the expected value signals on neural population assembles.
92 These signals are maintained in a stable state, though the cOFC signal sharply
93 develops in contrast to gradual evolution of the VS signal. Contrastingly, the
94 signals in the mOFC and DS were unstable and fluctuated. Thus, we conclude that
95 the cOFC and VS are responsible for computing the expected values from
96 probability and magnitude.
97

98 **Results**

99 **Task and behavior in monkeys**

100 We designed a behavioral task in which monkeys estimated the expected values
101 of rewards from numerical symbols, mimicking events performed by humans. The
102 task involved a visual pie-chart that included two numerical symbols associated
103 with the probability and magnitude of fluid rewards with great precision. After
104 monkeys fixated the central grey target, a visual pie-chart comprising green and
105 blue pies was presented (Fig. 1A). The number of the green pies indicated the
106 magnitude of fluid rewards in 0.1 mL increments (0.1-1.0 mL). Simultaneously, the
107 number of the blue pies indicated the probability of reward in 0.1 increments (0.1-
108 1 where 1 indicates a 100% chance). After a 2.5 second delay, the visual pie-chart
109 disappeared, and a reward outcome was provided to the monkeys with the
110 indicated amount and probability of reward, unless no reward was given. Under

111 this experimental condition, the expected values of rewards are defined as the
112 probability multiplied by the magnitude cued by the numerical symbols (Table. S1).
113 To examine whether the monkeys accurately perceived the expected values
114 from the numerical symbols for probability and magnitude, we applied a choice
115 task to the monkeys (Fig. 1B). The two monkeys exhibited a near-efficient
116 performance in selecting a larger expected value option among two alternatives
117 during choice trials (Fig. 1C). We examined which of the following three behavioral
118 models best described the monkey's behavior: model 1 (M1), monkeys make
119 choices based on the number of pies; model 2 (M2), monkeys make choices based
120 on the probability and magnitude; and model 3 (M3), monkeys make choices based
121 on the expected values (Table. S1). Comparisons of the model performances
122 based on Akaike's Information Criterion (AIC) and Bayesian Information Criterion
123 (BIC) (Burnham & Anderson, 2004) revealed that the model 3 best explained the
124 monkey's behavior as indicated by the smallest AIC and BIC values (Fig. S1). The
125 model 3 consistently showed the highest pseudo r-squared values in each monkey
126 (Fig. 1D). These results indicated that the monkeys utilize the expected values
127 estimated from the numerical symbols for probability and magnitude.
128

129 **Neural population data**

130 We constructed four pseudo-populations of neurons by recording single-neuron
131 activity during the single cue task (Fig. 1A) from the DS (194 neurons), VS (144
132 neurons), cOFC (190 neurons) and mOFC (158 neurons) (Fig. 1E). The four
133 constructed neural populations exhibited changes in their activities at different

134 times in the task trials (Fig. S2A). Approximately 40 to 50 % of neurons in the four
135 neural populations demonstrated peak activity during the cue presentation (Fig.
136 S2B, Chi-squared test, n = 686, P = 0.32, X^2 = 3.55, df = 3), with several basic
137 firing properties (Fig. S2C-F). Strong peak activities with short-latency were
138 observed in the cOFC (Kruskal-Wallis test, Latency: Fig. S2C, n = 314, P = 0.013,
139 H = 10.9, df = 3, Peak firing rate: Fig. S2D, n = 314, P < 0.001, H = 32.1, df = 3).
140 Activity changes were slow in the mOFC (Fig. S2E, Kruskal-Wallis test, n = 314, P
141 = 0.003, H = 13.4, df = 3). Baseline firing rates were the highest in the cOFC (Fig.
142 S2F, Kruskal-Wallis test, n = 686, P < 0.001, H = 60.3, df = 3). In short, the strong
143 activity with short-latency frequently occurred in the cOFC in contrast to the phasic
144 activity at various latencies in the DS and VS and relatively tonic and gradual
145 activity changes in the mOFC.
146
147 **Conventional analyses for detecting the expected-value signals**
148 We first applied common conventional analyses (linear regression, model selection
149 based on AIC, and model selection based on BIC) to the four neural populations
150 to examine neural modulations by the probability, magnitude, and expected values
151 at a single neuron level (see Methods). During a fixed one second time window
152 after the cue onset, these analyses showed that neurons in all the four brain
153 regions signal the probability, magnitude, and expected values to some extent (Fig.
154 2). For example, neurons signaling the expected values were found in each brain
155 region (Fig. 2A-H). In addition, neurons signaling probability or magnitude were
156 also observed in each brain region (Fig. 2I-L, blue and green). Moreover, a subset

of neurons in the cOFC and VS signaled high-risk high-return or low-risk low-return. These neurons were characterized as a strong activity elicited when the cue indicated a low probability and large magnitude (hence, high-risk high-return, Fig. S3, Fig. 2J and K, brown). Indeed, each neural population was composed of a mixture of these signals (Fig. 2I-L), indicating that signals for the expected values and its components (i.e., probability and magnitude) appeared in each neural population during one second after the cue onset. If we analyzed these neural modulation patterns through a task trial (Fig. 2M-P), there was no significant difference in the proportions of neural modulation types, except the VS (chi-squared test: DS, n = 104, df = 75, X^2 = 91.4, P = 0.096; VS, n = 104, df = 75, X^2 = 98.2, P = 0.037; cOFC, n = 104, df = 75, X^2 = 83.2, P = 0.242; mOFC, n = 104, df = 75, X^2 = 79.0, P = 0.353), indicating that these conventional analyses did not clearly provide evidence of temporal dynamics in the modulation patterns of neural populations. Thus, we developed an analytic tool to examine how the detection and integration of probability and magnitude are processed within these neural population ensembles.

State space analysis for detecting neural population dynamics

The state space analysis can provide temporal dynamics of neural population signal related to cognitive and motor performances (Churchland et al., 2012; Mante et al., 2013). In our lottery task, such population dynamics can describe how the expected values are evolved within the neural population ensembles. To describe how each neural population detects and integrates the probability and magnitude

into the expected values, we represented each neural population signal as a time-series of vectors in the space of probability and magnitude in two steps. First, we used a linear regression to project a time series of each neural activity into the regression subspace composed of the probability and magnitude in each neural population. This step captures the across-trial variance caused by probability and magnitude moment-by-moment at a population level. Second, we applied a principal component analysis (PCA) to the time series of neural activities in the regression subspace in each neural population. This step determines the main feature of the neural population signal moment-by-moment in the space of probability and magnitude. Because activations are dynamic and changing over time, the analysis identified whether and how signal transformations occurred to convert the probability and magnitude into the expected values as a time-series of eigenvectors (Fig. 3A). The directions of these eigenvectors capture the expected values as an angle moment-by-moment at a population level (Fig. 3B).

We evaluated properties of the eigenvectors for the first and second principal components (PC1 and PC2) in each neural population in terms of the vector angle, size, and deviance (Fig. 3C). A stable population signal is described as a small variation in the eigenvector properties through a trial, whereas an unstable population signal is described as a large variation in the eigenvector's properties. It must be noted that our procedure is a variant of the state space analysis in line with the use of linear regression to identify dimensions of a neural population signal (Chen & Stuphorn, 2015; Mante et al., 2013). It was not, however, aimed at projecting the population activity as trajectories in multi-dimensional space.

203

204 **Stable and unstable neural population signals**

205 The eigenvector analyses yielded clear differences in neural population signals
206 among the four populations (Fig. 4A-D). We first confirmed adequate performance
207 of the state space analysis indicated by the percentages of variance explained in
208 each population (Fig. 4A). The VS population exhibited the highest performance
209 among the four neural populations, which was followed by the cOFC and DS
210 populations, with the lowest performance exhibited by the mOFC population. Thus,
211 performances to process the probability and magnitude information was distinct
212 among the four neural populations.

213 To characterize the whole structure of each neural population signal, we
214 analyzed the aggregated properties of the eigenvectors without a temporal order
215 of the vectors through a task trial. We first examined the properties of the
216 eigenvectors for PC1. The aggregated eigenvectors revealed both stable and
217 unstable neural population signals during the cue presentation (Fig. 4B, green).
218 The VS population exhibited the highest performance (37%) and the eigenvectors
219 for PC1 were stable throughout cue presentation, with directions close to 45º, i.e.,
220 the expected value (Fig. 4B, VS, vector angle, PC1, mean ± s.e.m., 37.5º ± 0.98,
221 7.5º difference from 45º). The cOFC population also exhibited a stable expected-
222 value signal with the second-best performance (31%), but they were deviated more
223 from the ideal expected-value signal (Fig. 4B, cOFC, vector angle, PC1, mean ±
224 s.e.m., 59.4º ± 1.16, 14.4º difference from 45º, Wilcoxon rank-sum test, n = 52, W
225 = 122, $P < 0.001$). Stability of the vectors was the best in the VS and cOFC, as

226 indicated by the smallest deviation from its mean vector among the four neural
227 populations (Fig. 4C, left, PC1). Thus, the VS and cOFC populations signaled the
228 expected values in a stable manner.

229 In contrast, unstable population signals were observed in the DS and mOFC
230 (Fig. 4B, green). The DS population showed a considerable variability in their
231 eigenvectors (Fig. 4C, left, PC1) compared to those in the VS and cOFC neural
232 populations. The signal carried by the DS neural population was close to 0º, i.e.,
233 the probability (Fig. 4B, DS, vector angle, PC1, mean ± s.e.m., DS, 11.4º ± 1.72)
234 with a performance closer to that of the cOFC (29%). The mOFC population
235 exhibited a large variability in the eigenvectors (Fig. 4B, mOFC, PC1, vector angle,
236 mean ± s.e.m., 38.1º ± 5.80, Fig. 4C, left, PC1) due to the poorest performance of
237 PCA (14%), indicating a weak and fluctuating population signal. Thus, neural
238 populations in the DS and mOFC did not signal the expected values due to the
239 dynamic changes and weakness of the signals, respectively.

240 Second, we examined the properties of the eigenvectors for PC2. The
241 eigenvectors for PC2 revealed another feature of the neural population signals,
242 that reflected risk-return in the VS and cOFC (Fig. S3, Fig. 4B, blue, vector angle,
243 PC2, mean ± s.e.m., VS, 306.7º ± 1.07, 8.3º difference from 315º, cOFC, 322.4º ±
244 1.94, 7.4º difference from 315º). The deviations from the ideal risk-return signal
245 were not significantly different between the VS and cOFC populations (Wilcoxon
246 rank-sum test, n = 52, W = 319, P = 0.737). These signals were equally stable in
247 the VS and cOFC (Fig. 4C, right, PC2). In clear contrast, the DS and mOFC signals
248 were unstable and fluctuated more (Fig. 4C, right, vector angle, PC2, mean ±

249 s.e.m., DS, 64.8 ± 19.0, mOFC, 320.2 ± 8.77), similar to those observed for PC1
250 (Fig. 4C, left, PC1). Thus, the VS and cOFC were key brain regions to signal risk-
251 return as well as the expected values within their neural population ensembles,
252 suggesting that integrated information of the probability and magnitude could be
253 signaled in these neural populations.

254 To further examine the significance of these findings, we used a shuffle control
255 procedure in two ways (see Methods). First, we randomly shuffled the allocation of
256 probability and magnitude conditions to neural activity in each trial in each neuron
257 (shuffled condition 1). If we shuffled the linear projection of neural activity into the
258 regression subspace by this way, the neural population structure disappeared in
259 all the four brain regions (Fig. 4F). The performance of PCA for PC1 and PC2 were
260 all below 20% (Fig. 4E) and were significantly reduced from the observed data in
261 all the four brain regions, even in the mOFC (Fig. S4A, variance explained, $P <$
262 0.001 for all populations in PC1 and PC2). The deviance of eigenvectors under the
263 shuffle control clearly increased in the cOFC and VS, which resulted in no
264 significant differences among the four neural populations (Fig. 4G, Kruskal-Wallis
265 test, PC1, n = 104, df = 3, H = 4.99, $P = 0.18$, PC2, n = 104, df = 3, H = 2.78, $P =$
266 0.43). We also tested another shuffle control, in which the trial conditions were
267 shuffled in each analysis window throughout a trial (shuffled condition 2). Under
268 this full shuffled control, the PCA performances decreased further, albeit slightly
269 (Fig. 4I and Fig. S4B), without significant differences among the four populations
270 (Fig. 4J-K, Deviance, Kruskal-Wallis test, PC1, n = 104, df = 3, H = 1.38, $P = 0.71$,

271 PC2, n = 104, df = 3, H = 0.53, *P* = 0.91). Thus, these shuffle procedures
272 appropriately evaluated the significance of our population findings.

273 Next, we examined whether the size of eigenvectors is different among the
274 four neural populations, which represents the extent of neural modulation due to
275 the probability and magnitude in each neural population as an arbitrary unit. The
276 size of the eigenvectors was not significantly different (Fig. 4D, left, PC1, Kruskal-
277 Wallis test, n = 104, df = 3, H = 2.62, *P* = 0.45, right, PC2, n = 104, df = 3, H = 4.76,
278 *P* = 0.19), but it strongly depended on a temporal resolution for the analysis (Fig.
279 S5). Size of the eigenvector decreased as the analysis window size decreased
280 (Fig. S5B, E, and F), although all results and conclusions described above were
281 maintained across the window's sizes (Fig. S5A-D). The decrease in eigenvector
282 size would be because signal-to-noise ratios generally decrease if the window size
283 decreases. These effects were observed as the decrease in PCA performance (Fig.
284 S5A) and percentages of neural modulations in the conventional analyses (Fig.
285 S6).

286 Collectively, these observations suggested that the probability and magnitude
287 of rewards could be detected and integrated within the activity of cOFC and VS
288 neural populations as the expected value and risk-return signals in a stable state.

289

290 **Temporal structure of the neural population signals**

291 To characterize temporal aspects of the VS and cOFC neural populations that yield
292 the expected values signals, we compared temporal dynamics of all the four neural
293 population signals. Specifically, we compared temporal patterns of the vector

change exhibited by each neural population (Fig. 5). At the time point after the cue onset when the monkeys initiated the expected-value computation, all the four neural populations developed eigenvectors (Fig. 5A). Size of the eigenvectors increased and then decreased within a second, however temporal patterns of this change in size were different among the four neural populations. The onset latencies, detected by comparing to the vector size during the baseline period, seemed to be coincident for the cOFC, VS, and DS population, which was followed by a late noisy signal in the mOFC (Fig. 5B). In contrast, the detected peak of vector size for each neural population seemed to appear at different times. To statistically examine these temporal dynamics at the population level, we used a bootstrap re-sampling technique (see Methods).

The analysis revealed no significant difference in the onset latencies among cOFC, VS, and DS populations (Fig. 5C, bootstrap re-sampling, onset latency, mean ± s.d., cOFC, 107.1 ± 26.0 ms, VS, 138,7 ± 61.3 ms, DS, 155.0 ± 52.4 ms) while these signals were followed by a late noisy signal in the mOFC (mOFC, 287 ± 98.8 ms). In contrast, if we compared the peak latencies (Fig. 5D), the cOFC exhibited the earliest peak (292 ± 37.5 ms), followed by the DS (371 ± 43.0 ms), mOFC (444 ± 113.5 ms), and VS (508 ± 76.7), which exhibited the latest peak. Thus, the expected-value signal sharply developed in the cOFC in contrast to the gradual development in the VS. Note that mOFC signals were very noisy as indicated by the large variation in vector size during the baseline period (Fig. 5B, bottom, see horizontal line).

316 Lastly, we examined temporal changes in vector angle, which indicate how
317 fast the stable expected value signals evoked in the cOFC and VS (Fig. 5E). As
318 seen in the time series of the vector angle after detected onsets, the signals carried
319 by the VS and cOFC neural populations at the early time period were almost 45º
320 (i.e., expected values), indicating that these two neural populations integrate the
321 probability and magnitude information into the expected values just after the
322 numerical symbol appeared (see intercepts of the regression lines). Moreover,
323 these two expected value signals were not the same, but rather idiosyncratic in
324 each neural population: a gradual and slight shift of the vector angle directed to
325 90º (i.e., magnitude, cOFC, Fig. 5E, regression coefficient, $r = 5.31$, $n = 129$, $t =$
326 6.04, $df = 126$, $P < 0.0001$) or 0º (i.e., probability, Fig. 5E, VS, $r = -3.91$, $n = 127$, t
327 $= -4.16$, $df = 124$, $P < 0.0001$) was observed toward the end of cue presentation.
328 Similar to the VS, the DS population showed the same direction of the angle shift
329 (Fig. 5E, DS, $r = -5.38$, $n = 127$, $t = -3.31$, $df = 124$, $P = 0.001$). In contrast, such
330 significant shift of the vector angle was not observed in the mOFC population ($r =$
331 -4.30, $n = 120$, $t = -0.94$, $df = 117$, $P = 0.351$). These results indicated that the
332 neural populations in both the VS and cOFC integrate the probability and
333 magnitude information into the expected values immediately after the cue
334 presentation, whereas their temporal dynamics were idiosyncratic in each of the
335 two stable population signals.
336
337 **Discussion**

Extraction of neural population dynamics is a recently developing approach for understanding computational processes implemented in the domain of cognitive and motor processing (Chen & Stuphorn, 2015; Churchland et al., 2012; Mante et al., 2013; Murray et al., 2017; Takei et al., 2017). This approach provides a mechanistic structure of neural population signals regarding temporal aspects, such as oscillatory activities during reaching (Churchland et al., 2012), co-activation patterns of spinal neurons and muscles (Takei et al., 2017), and a dynamic unfolding of task-related activity for perceptual decisions (Mante et al., 2013). Here, we found that the VS and cOFC neural populations maintain the stable expected-value signals at a population level (Fig. 4). This is the first mechanistic demonstration of expected value signals embedded in multiple neural populations when monkeys compute the expected values from the numerical symbols cueing the probability and magnitude of rewards. The temporal dynamics of these two stable neural populations are unique in the aspect of time constants (Fig. 5B-D) and gradual shifts of their structures (Fig. 5E). These results indicate that the cOFC and VS compute the expected values as distinct, partially overlapping processes. These expected-value signals could be followed by the transition of neural population signals from evaluation to choices during economic decision makings (Chen & Stuphorn, 2015; Gardner et al., 2019; Yoo & Hayden, 2020).

Two idiosyncratic expected-value signals in the cOFC and VS

State space analysis can detect both stable (Murray et al., 2017) and flexible (Mante et al., 2013) neural signals at a population level. In the present study, expected-value signals observed in the VS and cOFC were similarly stable in terms of the fluctuation of vector angles but were significantly different in terms of the temporal aspects (Fig. 5). These signal properties indicate that information processing in these two brain regions were not same. For example, the fast cOFC signal may reflect the calculation of expected values from the probability and magnitude symbols like mental arithmetic, while the slow VS signal may reflect the secondary process to maintain the calculated expected-value information. It is known that the fronto-striatal projection plays a large role in a variety of cognitive functions (Haber & Knutson, 2010). Since the cOFC project to the VS anatomically, these two processes must act cooperatively through the cortico-basal ganglia loop. Indeed, both population signals were similar in terms of the heterogeneous signals carried by each individual neuron (Fig. 2J and K) throughout the task trial (Fig. 2N and O). However, these two expected-value signals were unambiguously distinctive in terms of their time course (Fig. 5B-D) and gradual shift (Fig. 5E). Therefore, the cOFC and VS may compute the expected values within each of these cortical and striatal local circuits in a co-operative manner.

Our results were consistent with human imaging studies, in which activity in the VS and cOFC represents value-related signals (Noonan et al., 2017; O'Doherty et al., 2004; Yan et al., 2016), but were not consistent with the evidences that value signals exist in human ventromedial prefrontal cortex (vmPFC) (Levy & Glimcher, 2012; Tom et al., 2007), which includes mOFC. The reasons for why the mOFC

383 showed very weak signals related to all aspects of the expected values in the
384 present study are unclear (Fig. 2L and 4B). One possibility of this inconsistency
385 may come from the species differences between human and non-human primates
386 in the orbitofrontal network (Wallis, 2011). The mOFC is a part of the vmPFC, but
387 the comparability between human and macaque monkeys remains elusive.
388 Another possibility is that the vmPFC is not involved in simple information
389 processing, such as association between cues and outcomes, but is involved in
390 more complicated behavioral context for making economic decisions (Yamada et
391 al., 2018) and setting of mood (Ongur & Price, 2000).
392

393 **Fluctuating signals in the DS and mOFC**

394 Fluctuating signals were observed in the DS and mOFC because of the instability
395 or weakness of the signals (Fig. 4). The mOFC signal would not be completely
396 meaningless, since the PCA performance in mOFC population was better than that
397 in shuffle controls (Fig. S4). However, the signal carried by mOFC population was
398 weak (Fig. 2L), indicating that the eigenvector fluctuation in mOFC population
399 reflects weak signal modulations by the probability and magnitude. In contrast,
400 PCA performance in the fluctuated DS population was equivalent to that in the
401 cOFC population (Fig. 4A), where the stable expected-value signal appeared.
402 Moreover, considerable modulation of the neural activity in DS was observed in
403 the conventional analyses (Fig. 2I and M). Thus, the fluctuating DS signal must
404 reflect some functional role employed by the DS neural population in detecting and
405 integrating the probability and magnitude information. The DS signal fluctuated

with a significant shift from relatively close to the expected values to probability signal (Fig. 5E, top). The consistent direction of the shift between VS and DS populations implies that striatal neural populations may prefer probabilistic phenomena (Ma & Jazayeri, 2014; Pouget et al., 2013), whereas cOFC neural population may prefer magnitude, that is continuous variable.

Significance of the population signals revealed by our state space analysis

State space analysis reveals temporal structures of neural population in multi-dimensional space for both cognitive (Murray et al., 2017) and motor tasks (Churchland et al., 2012; Takei et al., 2017). However, the interpretation of the extracted population structure depends on the way to use this procedure (Elsayed & Cunningham, 2017). In the present study, we did not seek to determine the population structure as a trajectory in neural state space, as was performed in previous studies. Instead, we aimed to detect the main features underscoring the population structure in the space of the probability and magnitude that compose the expected values. For this purpose, stability of the regression subspace is critical. We elaborately projected the neural firing rates into the regression subspace by preparing a completely orthogonal data matrix in our task design. Thus, our state space analysis is informative on whether and how the expected-value signals are composed of the probability and magnitude moment-by-moment as the series of eigenvectors.

The existence of the neural population signals for expected-values is consistent with the expected values theory, whereas the co-existence of risk-return signals may reflect a behavioral bias for risk-preferences, a phenomenon observed across species (Stephens & Krebs, 1986; Yamada, Tymula, et al., 2013). The sharp and slow evolutions of expected-value signals in the cOFC and VS, respectively, indicate that each brain region has a unique time constant in the expected-value computation. When the monkeys perceive the probability and magnitude from numerical symbols, learned expected-values may be computed and recalled through the OFC-striatum circuit (Hirokawa et al., 2019). Our results indicate that the expected-value signals as observed in population ensemble activities are compatible with the framework of dynamical systems (Churchland et al., 2012; Mante et al., 2013).

Materials and Methods

Subjects and experimental procedures

Two rhesus monkeys were used (*Macaca mulatta,* SUN, 7.1 kg, male; *Macaca fuscata*, FU, 6.7 kg, female). All experimental procedures were approved by the Animal Care and Use Committee of University of Tsukuba (protocol#. H30.336) and performed in compliance with the US Public Health Service's Guide for the Care and Use of Laboratory Animals. Each animal was implanted with a head-restraint prosthesis. Eye movements were measured using a video camera system at 120 Hz. Visual stimuli were generated by a liquid-crystal display at 60 Hz that was placed 38 cm from the monkey's face when seated. The subjects performed the cued lottery task five days a week. The subjects practiced the cued lottery task for 10 months, after which they became proficient at making choices of lottery options.

Behavioral task

Cued lottery tasks. Animals performed one of two visually cued lottery tasks: *single cue task* or *choice task*. Activity of neurons were recorded only during the single cue task.

Single cue task: At the start of each trial, the monkeys had 2 s to align their gaze to within 3° of a 1° diameter grey central fixation target. After fixating for 1 s, an 8° pie-chart providing information about the probability and magnitude of rewards was presented for 2.5 s at the same location as the central fixation target. The pie-chart was then removed and 0.2 s later a 1 kHz and 0.1 kHz tone (0.15 s duration)

indicated reward and no-reward outcomes, respectively. The high tone preceded a reward by 0.2 s. The low tone indicated that no reward would be delivered. The animals received a fluid reward, for which magnitude and probability were indicated by the green and blue pie-charts, respectively; otherwise, no-reward was delivered. An inter-trial interval of 4 to 6 s followed.

Choice task: At the start of each trial, the monkeys had 2 s to align their gaze to within 3° of a 1° diameter grey central fixation target. After fixating for 1 s, two peripheral 8° pie-charts providing information about the probability and magnitude of rewards for each of the two target options were presented for 2.5 s, 8° to the left and right of the central fixation location. Grey 1° choice targets appeared at these same locations. After a 0.5 s delay, the fixation target disappeared, cueing saccade initiation. The animals were free to choose for 2 s by shifting their gaze to either target, within 3° of the choice target. A 1 kHz and 0.1 kHz tone of 0.15 s duration indicated reward and no-reward outcomes, respectively. The animals received the fluid reward indicated by the green pie-chart of the chosen target, with the probability indicated by the blue pie-chart of the chosen target; otherwise, no-reward was delivered. An inter-trial interval of 4 to 6 s followed.

Pay-off and block structure. Green and blue pie-charts indicated reward magnitudes from 0.1 to 1.0 mL in 0.1 mL increments and reward probabilities from 0.1 to 1.0 in 0.1 increments, respectively (Table S1). A total of 100 pie-charts was used. In the single cue task, each of the 100 pie-charts was presented once in random order. In the choice task, two out of the 100 pie-charts were randomly

allocated to the two options. During one session of electrophysiological recordings, 20 to 50 trial blocks of the choice task were interleaved with 100 to 120 trial blocks of the single cue task.

Calibration of reward supply system. The precise amount of liquid reward was controlled and delivered to the monkeys with a solenoid valve. An 18-gauge tube (0.9 mm inner diameter) was attached to the tip of the delivery tube to reduce the trial-by-trial variability of reward supply. The amount of reward in each payoff condition was calibrated by measuring the weight of water with 0.002 g precision (hence, 2μL) on a single trial basis. This calibration method was the same as that in a previous study (Yamada et al., 2018).

Electrophysiological recordings

We used conventional techniques for recording single neuron activity from the DS, VS, cOFC, and mOFC. The monkeys were implanted with recording chambers (28 × 32 mm) targeting the OFC and striatum, centered 28 mm anterior to the stereotaxic coordinates. The locations of the chambers were verified using anatomical magnetic resonance imaging. In each recording session, a stainless-steel guide tube was placed within a 1-mm spacing grid, and a tungsten microelectrode (1-3 MΩ, FHC) was passed through the guide tube. For recording of neurons in the mOFC and cOFC, the electrode was lowered until it approximated the bottom of the brain after passing through the cingulate cortex or dorsolateral prefrontal cortex or between them, respectively. For recording of neurons in the

DS, the electrode was lowered until low spontaneous activity was observed after passing through the cortex and white matter. For recording of neurons in the VS, the electrode was lowered further until it passed through the internal capsule. At the end of the VS recording session, the electrode was occasionally lowered close to the bottom of the brain to confirm the recording depth relative to the bottom. Electrophysiological signals were amplified, band pass filtered, and monitored. Single neuron activity was isolated on the basis of spike waveforms. We recorded from the four brain regions of single hemisphere of the two monkeys: 194 DS neurons (98 and 96 from monkeys SUN and FU, respectively), 144 VS neurons (89, SUN and 55, FU), 190 cOFC neurons (98, SUN and 92, FU), and 158 mOFC neurons (64, SUN and 94, FU). The activity of all single neurons was sampled when the activity of an isolated neuron demonstrated a good signal-to-noise ratio (>2.5). Blinding was not performed. The sample sizes required to detect effect sizes (number of recorded neurons, number of recorded trials in a single neuron, and number of monkeys) were estimated with reference to previous studies (Chen & Stuphorn, 2015; Yamada, Inokawa, et al., 2013; Yamada et al., 2018). Neural activity was recorded during 100-120 trials of the single cue task. Note that presumed projection neurons (PANs, phasically active neurons) (Yamada et al., 2016) were recorded from the DS and VS.

Statistical analysis

For statistical analysis, we used the statistical software package R (http://www.r-project.org/). All statistical tests for behavioral and neural analyses were two-tailed.

Effects of units on statistical analysis. In the present study, we used two variables for analyses; probability and magnitude. We defined the probability of reward from 0.1 to 1.0, and the magnitude of reward from 0.1 to 1.0 mL. Under this definition of units, the effects of probability and magnitude on the data were equivalent. Thus, the data were not standardized in the analyses.

Behavioral analysis

We examined whether the monkey's choice behavior depended on the expected values of the two options located left and right side to the center. We pooled choice data across all recording sessions, yielding 44,883 and 19,292 choice trials for monkeys SUN and FU, respectively. The percentage of right target choices was estimated for all combinations of the expected values of left and right target options.

Model Fitting. The percentage of choosing the right-side option was analyzed using a general linear model with binominal distribution,

$$Pchooses_R = 1 / (1 + e^{-Z}) \qquad (1)$$

where the relation between $Pchooses_R$ and Z was given by the logistic function in each of the following three models: number of pies displayed (M1), probability and magnitude (M2), and the expected values (M3).

The first model, M1, assumed that the monkeys chose a target by comparing the number of pies for two targets,

$$Z = b_0 + b_1 Npie_L + b_2 Npie_R \qquad (2)$$

where b_0 is the intercept, and $Npie_L$ and $Npie_R$ are the number of pies contained in left and right pie-chart stimuli, respectively. Values of b_0 to b_2 were free parameters and estimated by maximizing log likelihood.

The second model, M2, assumed that the monkeys chose a target by comparing the probability and magnitude of two targets,

$$Z = b_0 + b_1P_L + b_2P_R + b_3M_L + b_4M_R \qquad (3)$$

where b_0 is the intercept; P_L and P_R are the probability of rewards for left and right pie-chart stimuli, respectively; and M_L and M_R are the magnitude of rewards for left and right pie-chart stimuli, respectively. Values of b_0 to b_4 were free parameters and estimated by maximizing log likelihood.

The third model, M3, assumed that the monkeys chose a target by comparing the expected values of rewards for two targets,

$$Z = b_0 + b_1EV_L + b_2EV_R \qquad (4)$$

where b_0 is the intercept, EV_L and EV_R are the expected values of rewards as probability times magnitude for left and right pie-chart stimuli, respectively. Values of b_0 to b_2 were free parameters and estimated by maximizing log likelihood.

Model comparisons. To identify the best structural model to describe the monkeys' behavior, we compared the three models described above. In each model, we estimated a combination of the best-fit parameters to explain monkeys' choice behavior. We compared their goodness-of-fits based on Akaike's Information Criterions (AIC) and Bayesian Information Criterion (BIC) (Burnham & Anderson, 2004),

$$\text{AIC (Model)} = -2L + 2k \qquad (5)$$

$$\text{BIC (Model)} = -2L + k \log n \qquad (6)$$

where L is the maximum log-likelihood of the model, k is the number of free parameters, and n is the sample size. After estimating the best-fit parameters in each model, we selected one model that exhibited the smallest AIC and BIC. To evaluate the model fits, we estimated a McFadden's pseudo r-squared statistic using the following equation:

$$\textit{Pseudo}\ \text{r-squared} = (L_0 - L_{Model}) / L_0 \qquad (7)$$

where L_{Model} is the maximum log likelihood for the model given the data, and L_0 is the log likelihood under the assumption that all free parameters were zero in the model.

Neural analysis

Basic firing properties. Peri-stimulus time histograms (PSTH) were drawn for each single neuron activity aligned at the visual cue onset. To display the histogram of neurons, the peak activity (maximum firing rate in each histogram) was detected for each neuron. The average activity curves were smoothed using a 50 ms Gaussian kernel (σ = 50 ms) and normalized by the peak firing rates. The percentage of neurons that showed an activity peak during cue presentation was compared among the four brain regions using a chi-squared test at $P < 0.05$. Basic firing properties, such as peak firing rates, latency of the peak, and duration of the peak activity (half peak width) were compared among the four brain regions using parametric and non-parametric tests, respectively, with a statistical significance

level of $P < 0.05$. Baseline firing rates during 1 s before the appearance of central fixation targets were also compared with a statistical significance level of $P < 0.05$.

Estimation of neural firing rates through task trials. We analyzed neural activity during a 2.7 s time period from the onset of pie-chart stimuli to the onset of outcome feedbacks during the single cue task. To obtain a time series of neural firing rates through a trial, we estimated the firing rates of each neuron for every 0.1, 0.05, or 0.02 s time window (without overlap) during the 2.7 s period. No gaussian kernel was used.

Estimation of neural firing rates in a fixed time window. We analyzed neural activity during a 1 s time window after the onset of pie-chart stimuli during the single cue task. The 1 s activity was used for the conventional analyses below. No gaussian kernel was used.

Conventional analyses to detect neural modulations in each individual neuron

Linear regression and model selection. For conventional and standard analyses of neural modulation by the values of probability and magnitude indicated by pie-chart stimuli, we used *linear regression* and *model selection analyses*. As mentioned above, we estimated the firing rate of each neuron during the 1 s time period after the onset of pie-chart stimulus during the single cue task. No gaussian kernel was used.

Linear regression. Neural discharge rates (F) were fitted by a linear combination of the following variables:

$$F = b_0 + b_p \text{Probability} + b_m \text{Magnitude} \quad (8)$$

where Probability and Magnitude are the probability and magnitude of rewards indicated by the pie-chart, respectively. b_0 is the intercept. If b_p and b_m were not 0 at $P < 0.05$, the discharge rates were regarded as being significantly modulated by that variable.

On the basis of the linear regression, activity modulation patterns were categorized into several types: “Probability” type with a significant b_p and without a significant b_m; “Magnitude” type without a significant b_p and with a significant b_m; “Expected value” type with significant b_p and b_m with both the same signs (i.e., positive b_p and positive b_m or negative b_p and negative b_m); “Risk-Return” type with significant b_p and b_m with both having opposite signs (i.e., negative b_p and positive b_m or positive b_p and negative b_m) and “non-modulated” type without significant b_p and b_m. Note that risk-return types reflect high-risk high-return (prefer low probability and large magnitude) or low-risk low-return (prefer high probability and low magnitude).

Model selection. Neural discharge rates (F) were fitted by the following five models:

$$\text{M1: } F = b_0 \quad (9)$$

$$\text{M2: } F = b_0 + b_p \text{Probability} \quad (10)$$

$$\text{M3: } F = b_0 + b_m \text{Magnitude} \quad (11)$$

M4: $F = b_0 + b_p \text{Probability} + b_m \text{Magnitude}$ (12)

M5: $F = b_0 + b_{ev} \text{Expected value}$ (13)

where Expected value is the expected value indicated by the visual pie-chart as the probability multiplied by magnitude (Table S1). b_0 is the intercept. Probability and Magnitude are the probability and magnitude of reward indicated by the pie-chart, respectively. Among the five models, we selected one model that exhibited the smallest AIC or BIC.

If the selected model was M1, neurons were defined as the “non-modulated” type. If the selected model was M2, neurons were defined as the “Probability” type. If the selected model was M3, neurons were defined as the “Magnitude” type. If the selected model was M4 with the same signs of b_p and b_m, neurons were defined as the “Expected value” type. If the selected model was M4 with opposite signs of b_p and b_m, neurons were defined as the “Risk-Return” type. If the selected model was M5, neurons were defined as the “Expected value” type.

Application of conventional analyses to neural activity through task trials. We applied the three conventional analyses above (linear regression, AIC-based model selection, and BIC-based model selection) for the activity of neurons estimated at every time window in the four brain regions. As mentioned above, we estimated the firing rate of each neuron for every 0.1, 0.05, or 0.02 s time window (without overlap) during the 2.7 s period. No gaussian kernel was used. The activity modulation type was defined in each time window during the 2.7 s period. The

analyses described percentages of the neural modulation types throughout cue presentation (Fig. S6).

Population dynamics using principal component analysis

Estimation of neuron firing rates through task trials. As mentioned above, we estimated the firing rate in each neuron every 0.1, 0.05, or 0.02 s time window (without overlap) during the 2.7 s period. No gaussian kernel was used.

Regression subspace. We used a linear regression to determine how the probability and magnitudes of rewards affected the activities of each neuron in the four neural populations. Each neural population was composed of all the recorded neurons in each brain region. We first prepared the probability and magnitude as 0.1 to 1.0 and 0.1 to 1.0 mL, respectively. We then described average firing rates of the neuron *i* at time *t* as a linear combination of the probability and magnitude in each neural population:

$$F_{(i,t,k)} = b_{0(i,t)} + b_{1(i,t)}\text{Probability}_{(k)} + b_{2(i,t)}\text{Magnitude}_{(k)} \qquad (14)$$

where $F_{(i,t,k)}$ is the average firing rates of neuron *i* at time t on trial *k,* $\text{Probability}_{(k)}$ is the probability of rewards cued to the monkeys on trial *k;* $\text{Magnitude}_{(k)}$ *is* the magnitude of rewards cued to the monkeys on trial *k*. The regression coefficients $b_{0(i,t)}$ to $b_{2(i,t)}$ describe the degree to which the firing rates of neuron *i* depend on the mean firing rates (hence, firing rates independent of task variables), the probability of rewards, and the magnitude of rewards, respectively, at a given time *t* during the trials.

We used the regression coefficients described in Eq. 14, above, to identify how the dimensions of neural population signals were composed from the probability and magnitudes as aggregated properties of individual neural activity. This step corresponds to the fundamental conceptual step of viewing the regression coefficients as a temporal structure of neural modulations by the probability and magnitude at the population level. Our procedures are analogous to the state-space analysis performed by Mante et al (Mante et al., 2013), in which the regression coefficients were used to provide an axis (or dimension) of the interested variables in multi-dimensional state space obtained by principal component analysis (PCA). In the present study, our orthogonalized task design allowed us to project the neural firing rates into the regression subspace reliably. It must be noted that our analyses were not aimed at describing population dynamics of the neural signals as a trajectory in the multi-dimensional task space, which is the standard goal of state space analysis.

Principal component analysis. We used a PCA to identify dimensions of the neural population signal in the orthogonal spaces composed of the probability and magnitude of rewards in each of the four neural populations. In each neural population, we first prepared a two dimensional data matrix X of size $N_{(neuron)} \times N_{(C \times T)}$; the regression coefficient vectors, $b_{1(i,t)}$ and $b_{2(i,t)}$ in Eq.14, whose rows correspond to the total number of neurons in each neural population and whose columns correspond to C, the total number of conditions (i.e., two: probability and magnitude); and T, the total number of analysis windows (i.e., 2.7 s divided by the

window size). The PCs of this data matrix are vectors $v_{(a)}$ of length $N_{(neuron)}$, total number of recorded neurons; if $N_{(C\times T)}$ is larger than $N_{(neuron)}$, otherwise the length is $N_{(C\times T)}$. The PCs were indexed from the principal component explaining the most variance to the least variance. The eigenvectors were obtained by prcomp () function in the R software. It must be noted that we did not perform de-noising in the PCA (Mante et al., 2013), since we did not aim to project firing rates into state space. Instead, we intended to use the principal components to identify main features of the neural modulation signals at the population level through task trials.

Eigenvectors. When we applied the PCA to the data matrix X, we could deconstruct the set into eigenvectors and eigenvalues. The eigenvectors and eigenvalues exist as pairs with every eigenvector having a corresponding eigenvalue. In our analysis, the eigenvectors at time t represent a vector in the space of probability and magnitude. The eigenvalues at time t for the probability and magnitude are scalars, indicating how much variance is in the data in that vector. The first principal component is thus the eigenvectors with the highest eigenvalues. We analyzed eigenvectors for the first (PC1) and second principal components (PC2) in the following analyses. It must be noted here that we applied the PCA once in each neural population and thus the total variances contained in the data were different among the four populations.

Analysis of eigenvectors. We evaluated characteristics of the eigenvectors for PC1 and PC2 in each of the four neural populations in terms of the vector angle, size,

and deviance in the space of probability and magnitude (Fig. 3). The angle is the vector angle from the horizontal axis from 0º to 360º. The size is the length of the eigenvector. The deviance is the difference between vectors, and we estimated the deviance from the mean vector in each neural population. These three characteristics of the eigenvectors were compared among the four neural populations at $P < 0.05$ using the Kruskal-Wallis test and Wilcoxon rank-sum test with Bonferroni correction for multiple comparisons. The vector during the first 0.1 s was extracted from these analyses.

Shuffle control for PCA. To examine the significance of the population structure described by the PCA, we performed two shuffle controls. When we projected neural activity into the regression subspace, we randomized the data by shuffling in two ways. In the shuffled condition 1, $b_{1(i,t)}$ and $b_{2(i,t)}$ in Eq.14 were estimated with the randomly shuffled allocation of trial number k to the Probability$_{(k)}$ and Magnitude$_{(k)}$ only once for all time t in each neuron. This shuffle provided a data matrix X of size $N_{(neuron)} \times N_{(C \times T)}$, eliminating the modulation of probability and magnitude observed in the condition C, but retaining the temporal structure of these modulations across time. In the shuffled condition 2, $b_{1(i,t)}$ and $b_{2(i,t)}$ in Eq.14 were estimated with the randomly shuffled allocation of trial number k to the Probability$_{(k)}$ and Magnitude$_{(k)}$ at each time t in each neuron. This shuffle provided a data matrix X of size $N_{(neuron)} \times N_{(C \times T)}$, eliminating the structure across conditions and times. In these two shuffle controls, matrix X was estimated 1,000 times. The PCA performance was evaluated by constructing distributions of the explained

variances for PC1 to PC4 (Fig. S4). The statistical significance of the explained variances by PC1 and PC2 was estimated based on the bootstrap standard errors (i.e., standard deviation of the reconstructed distribution).

Bootstrap re-samplings for onset and peak latencies of the neural population signals. To detect onset and peak latencies of the population signal, we analyzed dynamic changes in population structure with the size of the eigenvectors in each of the four populations. We used a time-series of eigenvectors in 0.02 s analysis windows and estimated sizes of the time-series of vectors for PC1. To obtain smooth changes in vector size, a cubic spline function was applied with a resolution of 0.005 s (Fig. 5B). Vector sizes during a baseline period were obtained by applying PCA to the matrix data X with time t from 0.3 s before cue onset to the onset of feedback (i.e., 3.0 s time period). A standard deviation of the vector sizes during the baseline period was obtained in each of the four neural populations. An onset latency of the population signal was defined as a time when the spline curve was above a value of more than three s.d. during the baseline period. Peak latency of the population signal was defined as the time from cue onset to the time when the maximum size of the vectors was obtained.

We estimated mean latencies of the onset and peak using a parametric bootstrap re-sampling method. In each neural population, the neurons were randomly re-sampled with a duplicate, and a data matrix X of size $N_{(neuron)} \times N_{(C \times T)}$ was obtained. The PCA was applied to the data matrix X. The time-series of the eigenvectors was obtained, and their sizes were estimated. The onset and peak

latencies were estimated as above. This resampling was conducted 1,000 times, and distributions of the onset and peak latencies were obtained (Fig. 5C and D). The statistical significance of the onset and peak latencies was estimated based on the bootstrap standard errors (i.e., standard deviation of the reconstructed distribution) (Efron. & Tibshirani., 1993).

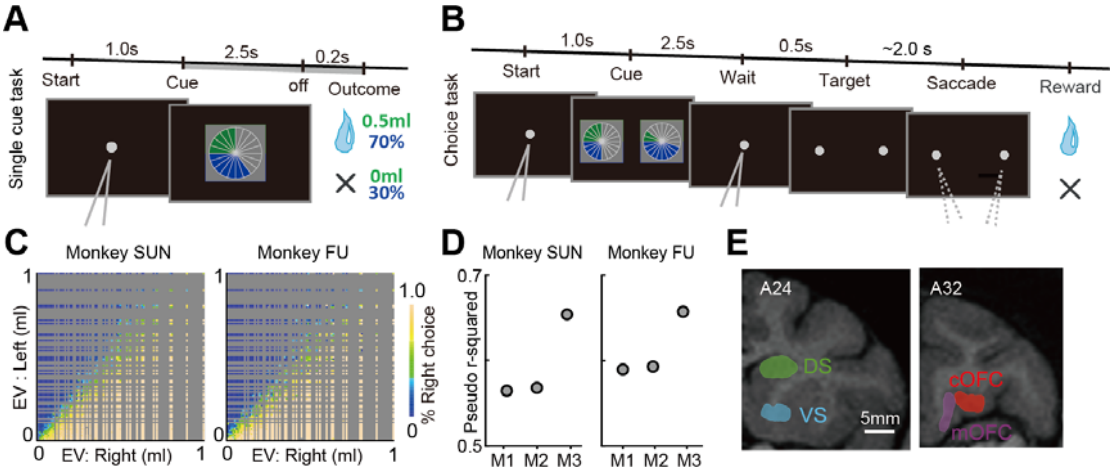
Acknowledgments

The authors appreciate the valuable comments of Tomomichi Oya, Tomohiko Takei, Tsuyoshi Setogawa, Jun Kunimatsu, Masafumi Nejime, Narihisa Matsumoto, Hiroshi Abe, and Takashi Kawai. In addition, the authors thank Takashi Kawai, Ryo Tajiri, Yoshiko Yabana, and Yuki Suwa for their technical assistance. Monkey FU was provided by NBRP "Japanese Monkeys" through the National Bio Resource Project of the MEXT, Japan.

Funding: This research was supported by JSPS KAKENHI Grant Number JP:15H05374, 18K19492, 19H05007, Takeda Science Foundation, Council for Addiction Behavior Studies, Narishige Neuroscience Research Foundation, The Ichiro Kanehara Foundation (H.Y.); JSPS KAKENHI Grant Number JP:26710001, MEXT KAKENHI Grant Number JP:16H06567 (M.M.).

Competing interests: The authors declare no competing of interests.

807 **Figures and legends**



808 Fig. 1. Task and behavior for recording neural activity.

809 (**A**) Sequence of events during the single cue task. A single visual pie-chart comprising
810 green and blue pies was presented to the monkeys. (**B**) Choice task. Two visually
811 displayed pie-charts were presented to the monkeys left and right to the center. After visual
812 fixation on the re-appeared central target, the central fixation target disappeared, and
813 monkeys chose either one of the targets by fixating on the target. A block of choice trials
814 was interleaved between single cue trial blocks. During the choice trials, neural activity
815 was not recorded. (**C**) Percentages of right target choice during the choice task plotted
816 against the expected values of left and right target options. (**D**) Pseudo r-squared
817 estimated in the three behavioral models. M1: number of pies. M2: probability and
818 magnitude. M3: the expected values. (**E**) An illustration of neural recording areas based
819 on sagittal magnetic resonance images. Neurons were recorded from medial (mOFC, 14O,
820 orbital part of area 14) and central parts of the orbitofrontal cortex (cOFC, 13M, medial
821 part of area 13) at A31-A34 anterior-posterior (A-P) level. Neurons were also recorded
822 from the dorsal (DS) and ventral striatum (VS) at A21-A27. The white scale bar indicates
823 5 mm.

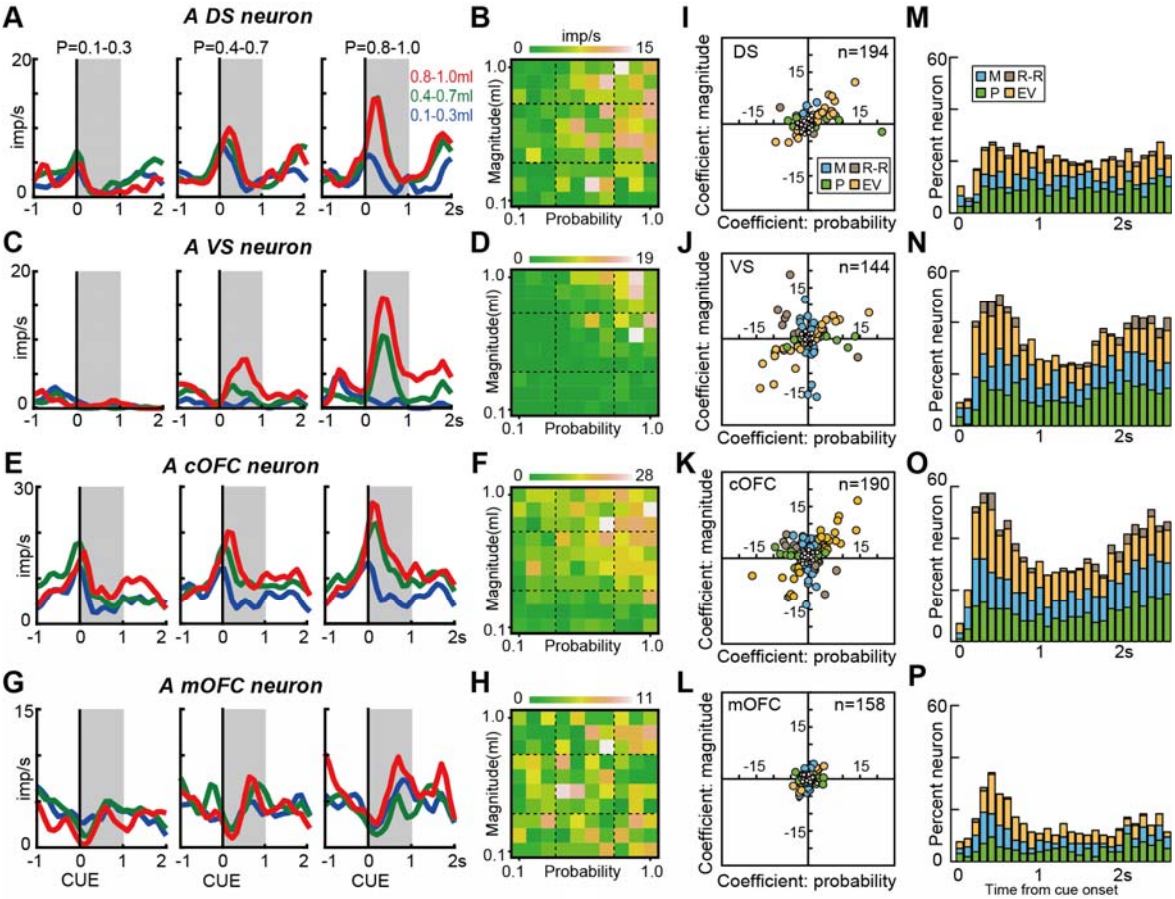


Fig. 2. Expected-value signals detected by conventional analyses.

(**A**) An example activity histogram of a DS neuron modulated by the expected values during the single cue task. The activity aligned to the cue onset is represented for three different levels of probability (0.1-0.3, 0.4-0.7, 0.8-1.0) and magnitude (0.1-0.3 mL, 0.4-0.7 mL, 0.8-1.0 mL) of rewards. Gray hatched time windows indicate a 1 s time window used to estimate the neural firing rates shown in **B**. The neural modulation pattern was defined as the Expected value type based on all three analyses (linear regression, AIC-based model selection, and BIC-based model selection). Regression coefficients were 6.17 ($P < 0.001$) and 2.54 ($P = 0.007$) for the probability and magnitude, respectively. (**B**) An activity plot of the DS neuron during the 1 s time window shown in **A** against the probability and magnitude of rewards. (**C-D**) same as **A-B**, but for a VS neuron defined as the Expected

835 value type based on all the three analyses. Regression coefficients were 7.14 (P < 0.001)
836 and 6.71 (P < 0.001) for the probability and magnitude, respectively. (**E-F**) same as **A-B**,
837 but for a cOFC neuron defined as the Expected value type based on all the three analyses.
838 Regression coefficients were 8.55 (*P* < 0.001) and 11.1 (*P* < 0.001) for the probability and
839 magnitude, respectively. (**G-H**) same as **A-B**, but for a mOFC neuron. The neural
840 modulation pattern was defined as the Expected value type based on the AIC-based
841 model selection, as the Probability type based on the linear regression, and as the non-
842 modulated type based on the BIC-based model selection. Regression coefficients were
843 1.76 (*P* = 0.032) and 0.50 (*P* = 0.54) for the probability and magnitude, respectively. (**I-L**)
844 Plots of the regression coefficients for the probability and magnitude of rewards estimated
845 for all neurons in the DS (**I**), VS (**J**), cOFC (**K**), and mOFC (**L**). Filled colors indicate the
846 neural modulation pattern classified by the BIC-based model selection. P: Probability type,
847 M: Magnitude type, EV: Expected value type, and R-R: Risk-Return type. The non-
848 modulated type is indicated by the small open circle. (**M-P**) Percentages of neural
849 modulation types defined based on the BIC-based model selection through the cue
850 presentation in the DS (**M**), VS (**N**), cOFC (**O**), and mOFC (**P**). The results using 0.1 s time
851 window (without overlap) were shown. The classification of neural modulation types was
852 dependent on the methods, although an overall tendency for the differences in neural
853 modulations among the neural populations was consistent among the three analyses.
854

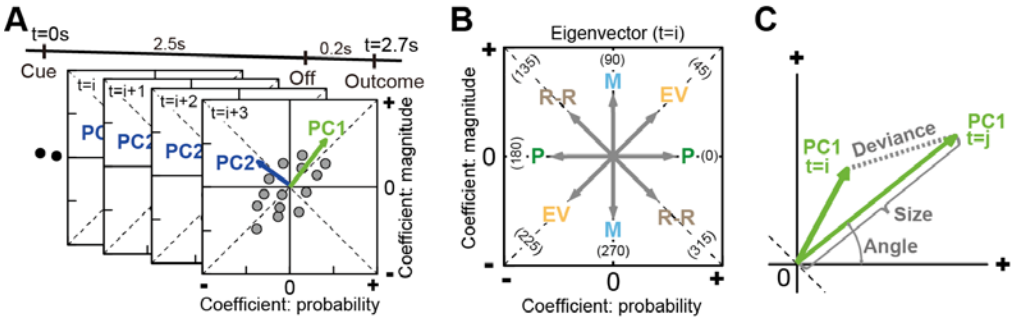
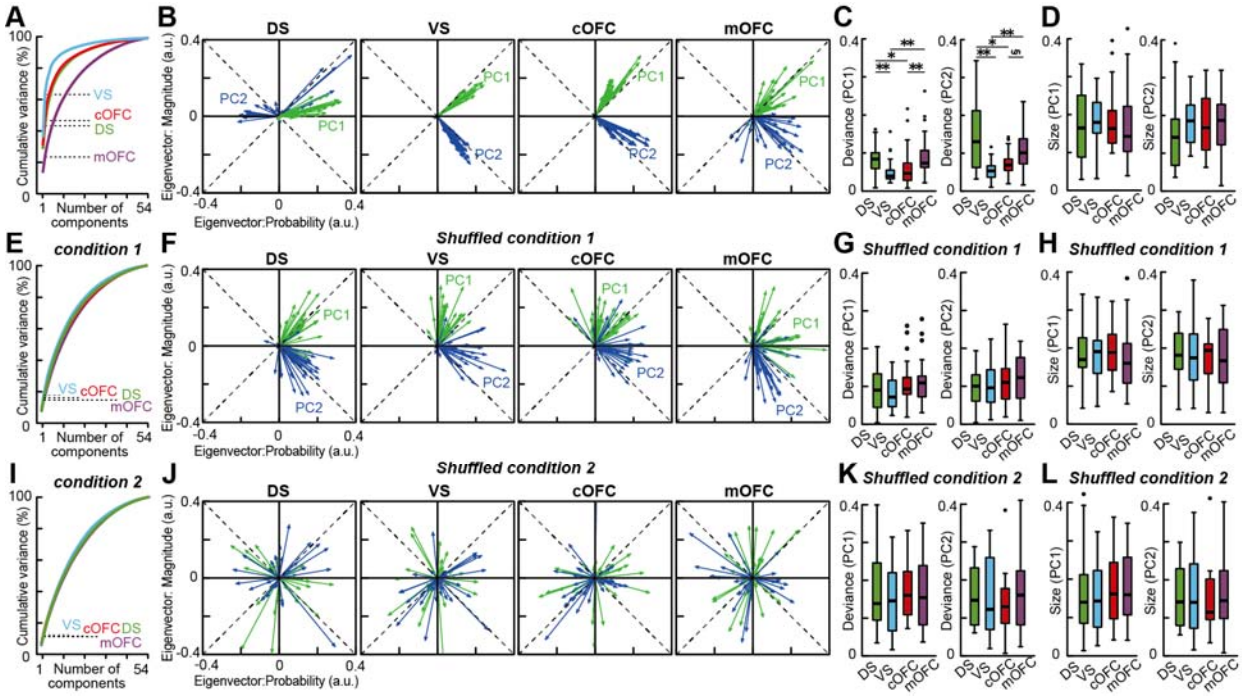


Fig. 3. Schematic depictions for the analysis of neural population dynamics using PCA. (**A**) Time series of a neural population activity projected into a regression subspace composed of the probability and magnitude. A series of eigenvectors was obtained by applying PCA once to each of the four neural populations. PC1 and PC2 indicate the first and second principal components, respectively. The number of eigenvectors obtained by the PCA was 2.7 s divided by the analysis window size for the probability and magnitude; 27, 54, and 135 eigenvectors in 0.1, 0.05, or 0.02 s time window, respectively. (**B**) Examples of eigenvectors at time of *i* th analysis window for the probability and magnitude, whose direction indicates a signal characteristic at that time represented on the population ensemble activity. EV: expected values (45º, 225º), M: magnitude (90º, 270º), P: probability (0º,180º), R-R: risk-return (135º, 315º). (**C**) Characteristics of the eigenvectors evaluated quantitatively. Angle: vector angle from horizontal axis from 0º to 360º. Size: length of the eigenvector. Deviance: difference between vectors.



869 Fig. 4. Neural populations provide stable expected-value signals in the VS and cOFC.
870 (**A**) Cumulative variance explained by PCA in the four neural populations. Dashed line
871 indicates percentages of variances explained by PC1 and PC2 in each neural population.
872 (**B**) Overlay plots of the series of eigenvectors for PC1 and PC2 in the four neural
873 populations. a.u. indicates arbitrary unit. (**C**) Box plots of the deviance from the mean
874 vector estimated in each neural population for the PC1 (left) and PC2 (right). (**D**) Box plots
875 of the vector size estimated in each neural population for the PC1 (left) and PC2 (right).
876 (**E-H**) Same as **A-D**, but for the PCA under the shuffled condition 1. See Methods for
877 details. (**I-L**) Same as **A-D**, but for the PCA under the shuffled condition 2. In **C-D**, **G-H**,
878 and **K-L**, asterisks indicate statistical significance between two populations using
879 Wilcoxon rank-sum test with Bonferroni correction for multiple comparisons (**,*, and §
880 indicate statistical significance at $P < 0.01$, $P < 0.05$, and $0.05 < P < 0.06$ (close to
881 significance), respectively). The results are shown by using 0.1 s analysis window.
882

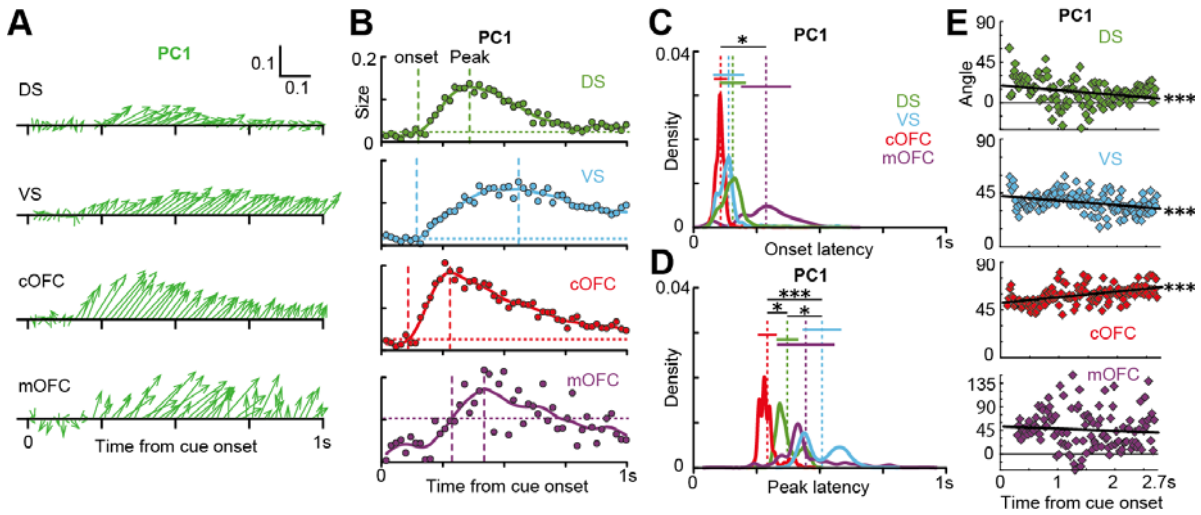


Fig. 5. Gradual and sharp evolutions of neural population signals in the VS and cOFC. (**A**) Plots of the time series of eigenvectors for PC1 in 0.02 s analysis windows shown in a sequential order during 1 s after the cue onset. Horizontal and vertical scale bars indicate the eigenvectors for the probability and magnitude in arbitrary units, respectively. (**B**) Plots of the time series of vector size during 1 s after the cue onset. Horizontal dashed lines indicate three standard deviations of the mean vector size during the baseline period, a 0.3 s time period before the cue onset. Solid colored lines indicate interpolated lines using a cubic spline function to provide a resolution of 0.005 s. Vertical dashed lines indicate the onset (left) and peak (right) latencies for the changes in vector sizes. (**C**) Probability densities of the onset latencies of the four neural population signals. The probability densities were estimated using bootstrap re-samplings. Vertical dashed lines indicate the means. Horizontal solid lines indicate the bootstrap standard errors. (**D**) same as **C**, but for the peak latencies of the four neural population signals. (**E**) Plots of the time series of vector angle from the detected onset to the onset of outcome feedback. Solid black lines indicate regression slopes. In **C** and **D**, asterisks indicate statistical significance estimated using bootstrap re-samplings (*** and * indicate statistical significance at $P < 0.001$ and P

< 0.05, respectively). In **E**, triple asterisks indicate a statistical significance of the regression slope at $P < 0.001$. Data for the PC2 is not shown.

References

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci, 7*(4), 404-410.

Burnham, K., & Anderson, D. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Method Res., 33*, 261–304.

Chen, X., & Stuphorn, V. (2015). Sequential selection of economic good and action in medial frontal cortex of macaques during value-based decisions. *Elife, 4*. doi:10.7554/eLife.09418

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature, 487*(7405), 51-56. doi:10.1038/nature11129

Efron., B., & Tibshirani., R. (1993). *An Introduction to the Bootstrap*: Chapman & Hall/CRC.

Elsayed, G. F., & Cunningham, J. P. (2017). Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat Neurosci, 20*(9), 1310-1318. doi:10.1038/nn.4617

Eshel, N., Tian, J., Bukwich, M., & Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nat Neurosci, 19*(3), 479-486. doi:10.1038/nn.4239

Fouragnan, E. F., Chau, B. K. H., Folloni, D., Kolling, N., Verhagen, L., Klein-Flugge, M., . . . Rushworth, M. F. S. (2019). The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. *Nat Neurosci, 22*(5), 797-808. doi:10.1038/s41593-019-0375-6

Gardner, M. P. H., Conroy, J. C., Sanchez, D. C., Zhou, J., & Schoenbaum, G. (2019). Real-Time Value Integration during Economic Choice Is Regulated by

Orbitofrontal Cortex. *Curr Biol, 29*(24), 4315-4322 e4314. doi:10.1016/j.cub.2019.10.058

Glimcher, P. W., Camerer, C. F., Fehr, E., & Poldrack, R. A. (2008). *Neuroeconomics: Decision Making and the Brain*. New York: Elsevier.

Goense, J. B., & Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr Biol, 18*(9), 631-640. doi:10.1016/j.cub.2008.03.054

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4-26. doi:10.1038/npp.2009.129

Hirokawa, J., Vaughan, A., Masset, P., Ott, T., & Kepecs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. *Nature, 576*(7787), 446-451. doi:10.1038/s41586-019-1816-9

Houthakker, H. S. (1950). Revealed Preference and the Utility Function. *Economica, 17*(66), 159-174.

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc Natl Acad Sci U S A, 112*(16), 5195-5200. doi:1503550112 [pii]
10.1073/pnas.1503550112

Howard, J. D., & Kahnt, T. (2017). Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. *J Neurosci, 37*(10), 2627-2638. doi:JNEUROSCI.3473-16.2017 [pii]
10.1523/JNEUROSCI.3473-16.2017

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *J Neurosci, 29*(7), 2231-2237. doi:10.1523/JNEUROSCI.5296-08.2009

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47*, 313–327.

Levy, D. J., & Glimcher, P. W. (2012). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J Neurosci, 31*(41), 14693-14707. doi:10.1523/JNEUROSCI.2218-11.2011

Lopatina, N., McDannald, M. A., Styer, C. V., Peterson, J. F., Sadacca, B. F., Cheer, J. F., & Schoenbaum, G. (2016). Medial Orbitofrontal Neurons Preferentially Signal Cues Predicting Changes in Reward during Unblocking. *J Neurosci, 36*(32), 8416-8424. doi:10.1523/JNEUROSCI.1101-16.2016

Ma, W. J., & Jazayeri, M. (2014). Neural coding of uncertainty and probability. *Annu Rev Neurosci, 37*, 205-220. doi:10.1146/annurev-neuro-071013-014017

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature, 503*(7474), 78-84. doi:10.1038/nature12742

Milham, M. P., Ai, L., Koo, B., Xu, T., Amiez, C., Balezeau, F., . . . Schroeder, C. E. (2018). An Open Resource for Non-human Primate Imaging. *Neuron, 100*(1), 61-74 e62. doi:10.1016/j.neuron.2018.08.039

Murray, J. D., Bernacchia, A., Roy, N. A., Constantinidis, C., Romo, R., & Wang, X. J. (2017). Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci U S A, 114*(2), 394-399. doi:10.1073/pnas.1619449114

Noonan, M. P., Chau, B. K. H., Rushworth, M. F. S., & Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal Cortex Lesions on Credit Assignment and Decision-Making in Humans. *J Neurosci, 37*(29), 7023-7035. doi:10.1523/JNEUROSCI.0692-17.2017

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454. doi:10.1126/science.1094285

Ongur, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex, 10*(3), 206-219.

Papageorgiou, G. K., Sallet, J., Wittmann, M. K., Chau, B. K. H., Schuffelgen, U., Buckley, M. J., & Rushworth, M. F. S. (2017). Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. *Nat Commun, 8*(1), 1886. doi:10.1038/s41467-017-01833-5

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*(6741), 233-238.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat Neurosci, 16*(9), 1170-1178. doi:10.1038/nn.3495

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J Neurosci, 29*(42), 13365-13376. doi:29/42/13365 [pii]

10.1523/JNEUROSCI.2572-09.2009

Rudebeck, P. H., & Murray, E. A. (2014). The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron, 84*(6), 1143-1156. doi:10.1016/j.neuron.2014.10.049

Samuelson, P. A. (1950). The Problem of Integrability in Utility Theory. *Economica, 17*(68), 355–385.

Savage, L. J. (1954). *The Foundations of Statistics*. New. York: John Wiley and Sons.

Stephens, D., & Krebs, J. (1986). *Foraging Theory*. New Jersey: Princeton Univ. Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning*. Cambridge: The MIT press.

Takei, T., Confais, J., Tomatsu, S., Oya, T., & Seki, K. (2017). Neural basis for hand muscle synergies in the primate spinal cord. *Proc Natl Acad Sci U S A, 114*(32), 8643-8648. doi:10.1073/pnas.1704328114

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307*(5715), 1642-1645.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315*(5811), 515-518. doi:315/5811/515 [pii]

10.1126/science.1134239

Von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. New Jersey: Princeton Univ. Press.

Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat Neurosci, 15*(1), 13-19. doi:10.1038/nn.2956

Xie, J., & Padoa-Schioppa, C. (2016). Neuronal remapping and circuit persistence in economic decisions. *Nat Neurosci, 19*(6), 855-861. doi:10.1038/nn.4300

Yamada, H., Inokawa, H., Hori, Y., Pan, X., Matsuzaki, R., Nakamura, K., . . . Minamimoto, T. (2016). Characteristics of fast-spiking neurons in the striatum of behaving monkeys. *Neurosci Res, 105*, 2-18. doi:10.1016/j.neures.2015.10.003

Yamada, H., Inokawa, H., Matsumoto, N., Ueda, Y., Enomoto, K., & Kimura, M. (2013). Coding of the long-term value of multiple future rewards in the primate striatum. *J Neurophysiol, 109*(4), 1140-1151. doi:10.1152/jn.00289.2012

Yamada, H., Louie, K., Tymula, A., & Glimcher, P. W. (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nat Commun, 9*(1), 162. doi:10.1038/s41467-017-02614-w

Yamada, H., Tymula, A., Louie, K., & Glimcher, P. W. (2013). Thirst-dependent risk preferences in monkeys identify a primitive form of wealth. *Proc Natl Acad Sci U S A, 110*(39), 15788-15793. doi:1308718110 [pii]
10.1073/pnas.1308718110

Yan, C., Su, L., Wang, Y., Xu, T., Yin, D. Z., Fan, M. X., . . . Chan, R. C. (2016). Multivariate Neural Representations of Value during Reward Anticipation and Consummation in the Human Orbitofrontal Cortex. *Sci Rep, 6*, 29079. doi:10.1038/srep29079

Yoo, S. B. M., & Hayden, B. Y. (2020). The Transition from Evaluation to Selection Involves Neural Subspace Reorganization in Core Reward Regions. *Neuron, 105*(4), 712-724 e714. doi:10.1016/j.neuron.2019.11.013

Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2019). Rat Orbitofrontal Ensemble Activity Contains Multiplexed but Dissociable Representations of Value and Task Structure in an Odor Sequence Task. *Curr Biol, 29*(6), 897-907 e893. doi:10.1016/j.cub.2019.01.048

Supplemental Information

Neural population dynamics underlying expected value computation

Hiroshi Yamada*, Yuri Imaizumi, Masayuki Matsumoto

Supplemental Figures

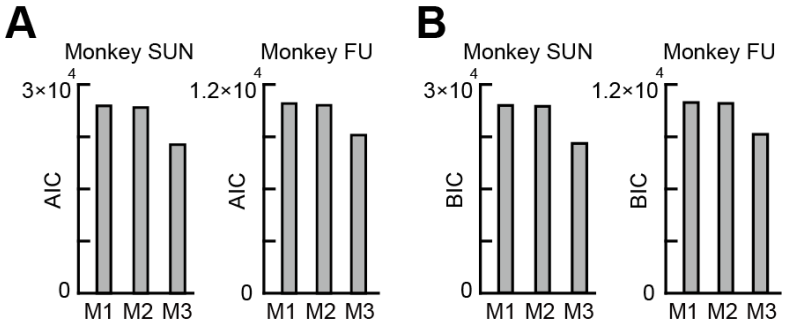


Fig. S1. Model fits to the choice behavior of two monkeys.

(**A**) Plots of the AIC values in monkey SUN and FU estimated in the three behavioral models, M1: number of pie. M2: probability and magnitude. M3: the expected values. (**B**) Same as **A**, but for the BIC values.

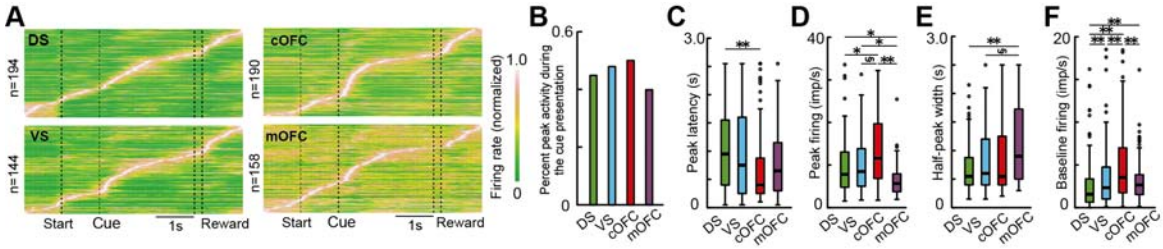


Fig. S2. Basic firing properties of neurons in the DS, VS, cOFC, and mOFC.

(**A**) Color map histograms of neuronal activities recorded from the four brain regions. Each horizontal line indicates the neural activity aligned to the cue onset, that was averaged for all lottery conditions. Firing rates of a neuron were normalized to the peak activity. (**B**) The percentages of neurons that showed an activity peak during the cue presentation. (**C**) Latency plots of the peak activity after the cue presentation. (**D**) Firing rates of the peak activity observed during the cue presentation. (**E**) Plots of the half-peak width, which indicates the phasic nature of activity changes. (**F**) Plots of the baseline firing rates during 1 s time period before the onset of the central fixation target. In **C-F**, asterisks indicate statistical significance among two neural populations using Wilcoxon rank-sum test with Bonferroni correction for multiple comparisons (**,*, and § indicate statistical significance at $P < 0.01$, $P < 0.05$, and $0.05 < P < 0.06$ (close to significance), respectively).

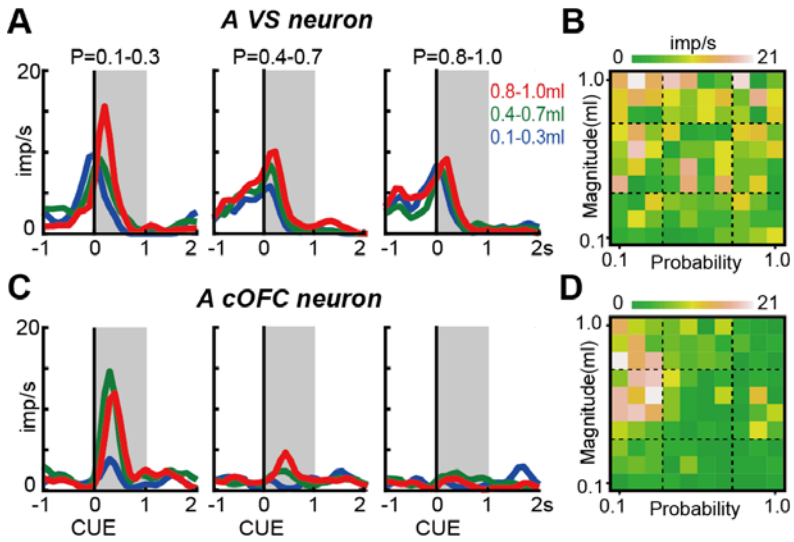


Fig. S3. Risk-return signals detected by the conventional analyses.
(**A**) An example activity histogram of a VS neuron modulated by both the probability and magnitude of reward with opposite signs (i.e., negative b_p and positive b_m). The activity aligned to cue onset is represented for three different levels of the probability (0.1-0.3, 0.4-0.7, 0.8-1.0) and magnitude (0.1-0.3 mL, 0.4-0.7 mL, 0.8-1.0 mL) of rewards. Gray hatched areas indicate a 1 s time window to estimate the neural firing rates shown in **B**. The neural modulation pattern was defined as the Risk-Return type based on the linear regression and AIC-based model selection, and as Magnitude type based on the BIC-based model selection. Regression coefficients were -2.44 (P = 0.039) and 4.86 (P < 0.001) for the probability and magnitude, respectively. (**B**) Activity plots of the VS neuron during the 1 s time window shown in **A** against the probability and magnitude of rewards. (**C-D**) same as **A-B**, but for a cOFC neuron. The neural modulation type was defined as the Risk-Return type based on all the three analyses. Regression coefficients were -6.65 (P < 0.001) and 3.82 (P < 0.001) for the probability and magnitude, respectively.

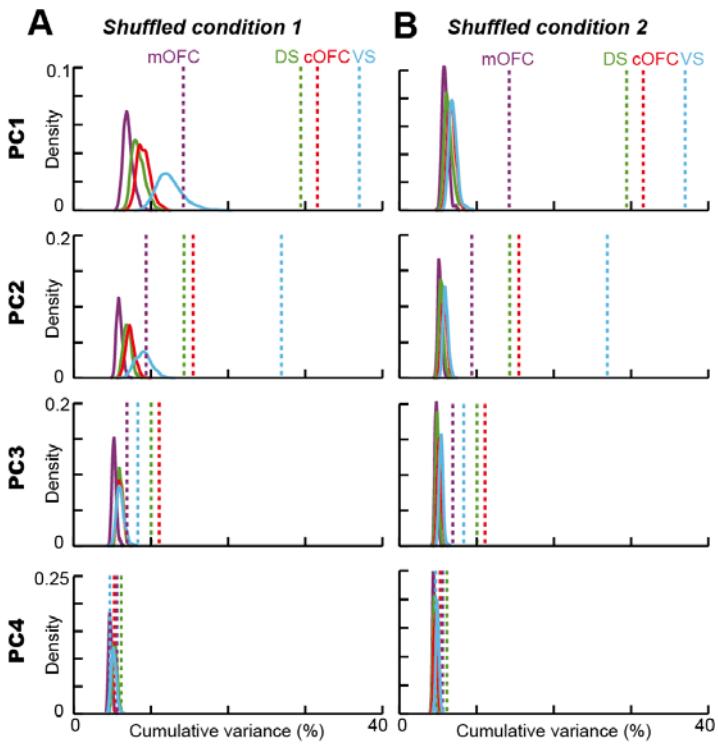


Fig. S4. Probability density of the explained variances by PCA in the shuffled controls.

(**A**) Probability density of variances explained by PCA for PC1 to PC4 under the shuffled condition 1 (see Methods for details). The probability density was estimated with 1,000 repeats of the shuffles in each neural population. (**B**) Probability density of variance explained by the PCA for PC1 to PC4 under the shuffled condition 2 (see Methods for details). The probability density was estimated with 1,000 repeats of the shuffles in each neural population. In **A** and **B**, dashed lines indicate the variances explained by PCA in each of the four neural populations without the shuffle.

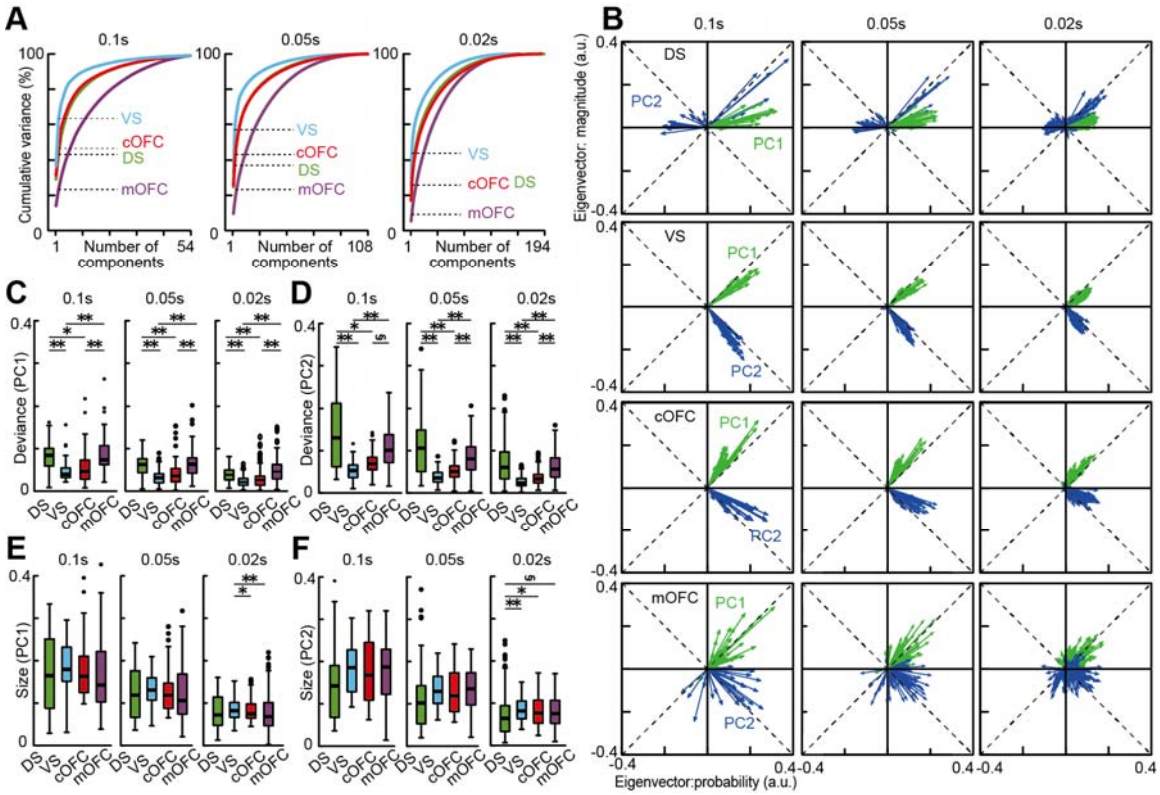


Fig. S5. Effects of the analysis window size on the PCA.

(**A**) Cumulative variances explained by PCA in the four neural populations. Dashed lines indicate the percentages of variance explained by PC1 and PC2 in each neural population. The size of the analysis windows was 0.1, 0.05, and 0.02 s, respectively. (**B**) Overlay plots of the series of eigenvectors in the four neural populations. The eigenvectors for PC1 and PC2 are shown. The size of the analysis windows was 0.1, 0.05, and 0.02 s, respectively. a.u. indicates arbitrary units. (**C**) Box plots of the deviance from the mean vector estimated in each neural population are shown for the PC1. (**D**) same as (**C**), but for the PC2. (**E**) Box plots of the vector size estimated in each neural population are shown for the PC1. (**F**) same as (**E**), but for the PC2. In **C-F**, asterisks indicate statistical significance between two neural populations using Wilcoxon rank-sum test with Bonferroni correction for multiple comparisons (**, *, and § indicate statistical significance at $P < 0.01$, $P < 0.05$, and $0.05 < P < 0.06$ (close to significance), respectively).

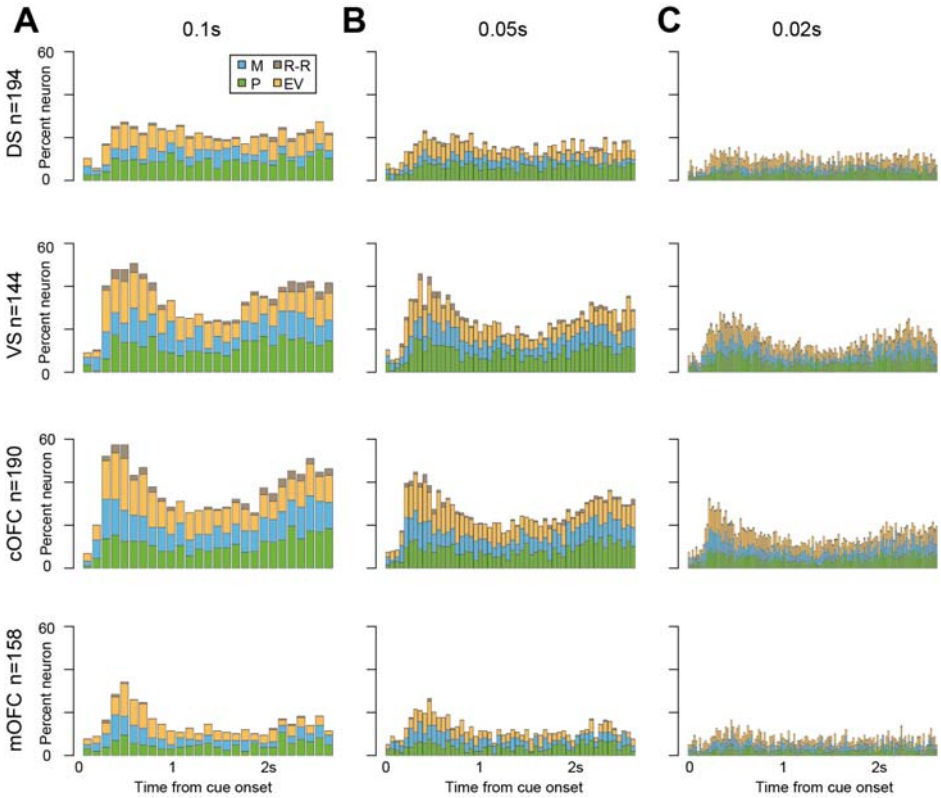


Fig. S6. Effects of the analysis window size on the conventional analyses.
(**A**) Percentage plots of the neural modulation type defined by using the model selection based on BIC values. The data for 0.1 s analysis windows is shown for the four neural populations. M: Magnitude type, P: Probability type, EV: Expected value type, R-R: Risk-Return type. (**B**) same as **A**, but for the data in 0.05 s analysis window. (**C**) same as **A**, but for the data in 0.02 s analysis window.

Supplemental table

EV	Probability									
	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.0
Magnitude (ml) .10	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
.20	.02	.04	.06	.08	.10	.12	.14	.16	.18	.20
.30	.03	.06	.09	.12	.15	.18	.21	.24	.27	.30
.40	.04	.08	.12	.16	.20	.24	.28	.32	.36	.40
.50	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
.60	.06	.12	.18	.24	.30	.36	.42	.48	.54	.60
.70	.07	.14	.21	.28	.35	.42	.49	.56	.63	.70
.80	.08	.16	.24	.32	.40	.48	.56	.64	.72	.80
.90	.09	.18	.27	.36	.45	.54	.63	.72	.81	.90
1.0	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.0

Table S1. Payoff matrix regarding probability and magnitude of rewards.
Payoff matrix of the probability and magnitude of rewards indicated by the blue and green pie chart. The expected values [EV] defined as probability times magnitude are shown for the 100 possible combinations. All the 100 possible combinations of the reward probability and magnitude were used during the single cue and choice task.