Individuals with ventromedial frontal damage have more unstable but still fundamentally transitive preferences

Abbreviated title: Transitive preferences in VMF damage

Authors: Linda Q. Yu[1], Jason Dana[2] and Joseph W. Kable[1]
[1] Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104
[2] Yale School of Management, Yale University, New Haven, CT, 06520
Corresponding author:
Linda Q. Yu
Kable Lab – Goddard 5
433 S. University Ave
Philadelphia, PA, 19104
lindaqyu@sas.upenn.edu

Number of pages: 38
Number of figures: 5
Number of tables: 2
Number of words for Abstract: 249
Number of words for Introduction: 649
Number of words for Discussion: 1475
Conflict of interest: The authors declare no competing financial interests.
Acknowledgments: We would like to thank Avinash Vaidya and Arthur Lee for constructive discussion on aspects of data analysis. We would also like to thank Lesley Fellows for facilitating access to participants in Montreal, Christine Déry and Eileen Cardillo for coordinating participants in Montreal and in Philadelphia, and all of the participants themselves, without whom this work would not be possible. This work was supported by NIDA R01-DA029149 to JWK and an NSERC PGS-D award to LQY.

Abstract

Transitivity of preferences (i.e., if one prefers A over B, and B over C, one should prefer A over C) is a hallmark of making rational, value-based decisions. Damage to the ventromedial frontal lobes (VMF) has been shown in previous studies to increase intransitive choice cycles (i.e., choosing A over B and B over C, but C over A). However, past studies have examined transitivity by treating preferences as deterministic rather than probabilistic, which could mask an important distinction in the critical role of the VMF in value-based choices: are individuals with VMF damage prone to choosing irrationally, or are they transitive, but simply more variable in what they prefer? We present individuals with focal VMF damage, controls with other frontal damage, and healthy controls with incentive compatible stimuli (artwork, brands of chocolate, and gambles) and have them make repeated choices between all possible pairs. Using cutting edge tests of a model of stochastic transitivity, and replicating previous analyses of transitivity that treat preferences as deterministic, we find that individuals with VMF damage made decisions consistent with stochastic transitivity. We also replicate previous findings that these individuals more frequently violate deterministic notions of transitivity. Our results are consistent with the hypothesis that individuals with VMF damage are not, in fact, more irrational, but do have noisier preferences. The implication is that the VMF is critical to maintaining the stability of preferences across time and context during decision-making, rather than for the ability for choices to reflect preferences at all.

Significance statement

To best satisfy one's goals (i.e., to maximize value), it is necessary to be transitive in one's preferences. Nearly all normative and descriptive theories of decision making are transitive. Damage to the VMF has been shown to result in more inconsistent preferences, and functional neuroimaging studies have identified signals in VMF correlated with subjective value. However, the observed inconsistent choices after VMF damage have not been thoroughly characterized, and the contribution of VMF to value-based choice is not well understood. Our study shows that the VMF affects the noisiness with which value is assessed, but not the consistency with which value is sought. This finding has implications for both clinical outcomes and for decision neuroscience theory.

Introduction

A central assumption of many theories of choice is that decision-makers compare different options on a single dimension of subjective value and choose the highest valued option. Satisfying this assumption is equivalent to the observed choices being transitive (Samuelson, 1937). An example of transitivity is the following: if you choose to listen to Adele (A) over Britney Spears (B), and Britney over Celine Dion (C), then you would also choose Adele (A) over Celine (C). There is a strong argument that choices ought to be transitive, as an intransitive chooser could be exploited (e.g., as a "money pump") and would get caught in choice cycles that do not advance towards any goal. Given this, one might expect that organisms develop internal representations of subjective value to ensure transitivity. Key studies in neuroeconomics have identified neural signals in the ventromedial frontal lobe (VMF) that scale with subjective value across different goods, in the firing rate of single neurons in the orbitofrontal cortex in monkeys (Padoa-Schioppa & Assad, 2006) and in the BOLD signal of ventromedial prefrontal cortex in humans (Bartra, McGuire, & Kable, 2013; Levy & Glimcher, 2012).

Consistent with the idea that neural signals in the VMF support value maximization, inconsistency has long been recognized as a hallmark of VMF damage: Phineas Gage was "capricious and vacillating" (Harlow, 1868) and EVR would drive on a single street for hours trying to decide on a restaurant (Eslinger & Damasio, 1985). More recently, individuals with VMF damage have been shown to make more intransitive choices than healthy controls or individuals with damage elsewhere in the frontal lobe (Camille, Griffiths, Vo, Fellows, & Kable, 2011; Fellows & Farah, 2007; Henri-Bhargava, Simioni, & Fellows, 2012). In the above example with the songstresses, an individual with VMF damage would be more likely to choose C(eline) over A(dele).

Axioms of rational choice, like transitivity, are usually stated deterministically. In contrast, behavior in experiments is probabilistic, because people can make different choices given the same pair of options over time (Luce, 1959; McFadden, 1980; Regenwetter, Dana, & Davis-Stober, 2011; Tversky, 1969). There are different ways to recast transitivity in probabilistic terms (Regenwetter et al., 2011; Tversky, 1969); however, testing any probabilistic model requires observing repeated choices over many instances of the same stimulus pairs. Noting a cycle (e.g., choosing C over A when one has chosen A over B and B over C) is not sufficient to disentangle whether one has fundamentally intransitive preferences versus variable preferences.

Previous studies have only asked individuals with VMF damage about their preferences between each pair of stimuli a single time. Therefore, the greater tendency of individuals with VMF damage to make intransitive choices in these experiments is consistent with two very different possibilities from a probabilistic perspective. One possibility is that the choices of individuals with VMF damage are fundamentally intransitive. In this case, their choices would not satisfy probabilistic notions of transitivity (e.g., by consistently and reliably choosing C > A above). This could occur if individuals with VMF damage chose according to stimulus-response associations or rules that lack any higher order transitive structure, such as the lexicographic semiorder heuristic (Tversky, 1969). A second possibility is that the choices of individuals with VMF damage are fundamentally transitive, but noisier. In this case, their choices would satisfy probabilistic notions of transitivity despite violating deterministic notions more often (e.g., they might choose A over C above with greater than 50% probability, but not 100% of the time). This could occur if individuals with VMF damage chose according to underlying values, but did so less reliably.

Here we test which of these two possibilities holds. The answer is both clinically relevant, as it sheds light on the nature of "capricious and vacillating" behavior after VMF damage, and theoretically relevant, as it determines whether VMF is necessary for choices to reliably reflect underlying values or for choices to be value-based at all.

Materials and Methods

Experimental Design

Participants. Fourteen individuals with focal damage to the frontal lobes were recruited from the Focal Lesion Database (FoLD) at the University of Pennsylvania, and ten were recruited from the Cognitive Neuroscience Research Registry at McGill University (Fellows, Stark, Berg, & Chatterjee, 2008). Individuals were eligible to participate if they had a lesion primarily affecting the frontal lobes. One individual was excluded due to incomplete data collection (the individual completed one session and was not able to be scheduled for the second). Fourteen females and 9 males were included in the final sample. Participants were tested a minimum of 5 months after injury (median = 10.29 years, range: 5 months to 17.75 years).

Participants were divided into two groups *a priori* based on location of damage, assessed with MR or computed tomography images by a neurologist blind to task performance.

The ventromedial frontal lobe (VMF) group consisted of individuals who sustained damage to the VMF, while the frontal control group (FC) consisted of individuals who sustained damage to the frontal lobe sparing the VMF. Lesions were drawn on a common space [Montreal Neurological Institute (MNI) brain] by neurologists at the research sites blind to task performance. The overlap images for the groups are found in **Figure 1**. Damage in the VMF group was caused by aneurysm or subarachnoid hemorrhage in 5 cases, stroke in 2 cases, tumor resection in 3 cases, glioma in one case, and meningioma in 2 cases. Damage in the FC group was caused by hemorrhage, stroke or infarct in 7 cases, glioma in 2 cases, and meningioma in one case.

Age and education matched healthy controls (HC) were recruited from the corresponding Normal Control Databases of the University of Pennsylvania (N = 14) and McGill University (N = 6), including 15 females and 5 males (**Table 1**). They were free of neurological and psychiatric disorders. All subjects provided informed consent and were compensated for their time. The study protocol was approved by the institutional review boards of both the University of Pennsylvania and McGill University.

Apparatus. All tasks were programmed using EPrime 2.0 (Psychology Software Tools). Participants were tested at the Hospital of the University of Pennsylvania, at the MNI, or at their own home in the greater Philadelphia or Montreal area. Participants saw stimuli on a laptop monitor and responded using the 1 and 0 keys of the keyboard.

Stimuli. Stimuli consisted of images of artwork, chocolate bars, and pie charts representing gambles. There were two sets of stimuli: 10-11 stimuli for each of the categories (10 for chocolate bar brands, 11 for art and gambles) used in non-repeated choices that allow deterministic tests of transitivity (set A), and 5 stimuli for each of the categories (art, chocolate bar brands, gambles) used in repeated choices that allow probabilistic tests of transitivity (set B). Choices constructed using set A and set B stimuli were intermingled in each block. For each category, we strove to design option sets in which the options were close in preference, as intransitive choices are less likely between items that have widely different values.

The artwork stimuli were paintings that were rated highly by participants in Vaidya and Fellows (2015a). The set B stimuli consisted of 5 paintings by Monet, which were all within the top 20 most highly rated paintings by those subjects. We selected Monet as he was the artist that occurred most frequently in the top 20 rated paintings of Vaidya and Fellows (2015a). The 5 selected paintings were roughly similarly preferred (i.e., chosen with close to the same frequency in pair-wise choices across the whole sample) in a sample of 107 participants recruited from Amazon Mechanical Turk. Set A consisted of paintings of the similar style/era (Impressionist, Romantic periods) in the top 40 ranked paintings of the Vaidya and Fellows (2015a) stimuli set.

The chocolate bars were from five brands (Lindt, Godiva, Ghirardelli, Dove, and Cadbury). We selected five brands that were roughly similarly preferred across the population. These brands were being sold for similar prices, were rated similarly on a seven-point scale by a sample of 103 participants from Amazon Mechanical Turk (mean rating = 5.76), and were selected at roughly similar frequencies in pair-wise choices across another sample of 101

Mechanical Turk participants. Milk chocolate bars from each of the 5 brands were in set B, while dark chocolate and dark chocolate almond bars from each brand were in set A. The stimuli consisted of publicly available pictures of the front side of the chocolate bar packaging.

We used sixteen gambles of equal expected value ($8.80). The stimuli consisted of a pie chart showing the probability of winning, with text on top indicating both the cash amount to be won and the probability of winning. The five set B gambles were the “Cash II” set in Regenwetter et al. (2011), which used contemporary monetary equivalents of the Tversky (1969) five gamble set. The probabilities were 28%, 32%, 36%, 40%, and 44%. Set A consisted of 11 other gambles with the same expected value (probabilities of 8%, 17%, 25%, 33%, 42%, 50%, 58%, 67%, 75%, 83%, 92%).

Procedure. Participants completed a binary forced choice task. On each trial, participants first saw a central fixation point for 1s, then a screen with two choice stimuli (placed to the left and the right of the center). Participants indicated which stimulus they preferred, by pressing buttons for left or right. Participants had as much time as they needed to make their selection. Following their selection, there was an inter-trial interval of 1s where a black screen was presented.

For set A stimuli, participants faced all possible pairings of either 10 (for brands) or 11 (for art and gambles) options, constituting 45 and 55 pairs in total, respectively. Each pair was faced once. For set B stimuli, participants faced all possible pairings of 5 options, constituting 10 pairs, and each pair was repeated 15 times. Therefore, there were 195 (for brands) or 205 (for art and gambles) total choices in each category across the entire experiment.

Choice trials were presented in blocks, in which participants made choices between items within a single category (art, brands, gambles). There were five blocks of choices for each category, containing 39 (for brands) or 41 (for art and gambles) trials each. Each block contained 9 or 11 choices composed from set A and 30 choices composed from set B. Choices from set A and set B were intermingled with each other within a block, with the set A stimuli inserted into a block of B stimuli in positions randomly selected from a uniform distribution. We took a number steps to reduce any potential memory effects for choices constructed with set B stimuli. We designed the sequence of trials so that: (1) the same pairing was not repeated within a minimum of 3 trials; (2) the same stimulus rarely appeared on immediately adjacent trials (no more than 9 times throughout the entire experiment); and (3) when the same pairing was repeated the choices immediately preceding and following that pairing differed from its previous occurrence (to minimize contextual memory). Furthermore, the side on which stimuli were presented was counterbalanced across repetitions. Finally, we divided the experiment into two sessions, held on separate days for every subject except two (due to scheduling constraints). The two sessions were held on average 8.09 (sd = 11.73) days apart (excepting the two who were tested on the same day, the sessions ranged from 1 day to 57 days apart). We did not observe a significant correlation between total number of intransitive choices made across all participants (see explanation of measure below) and days between the two sessions (r = 0.24, p = 0.12).

Statistical Analysis

Deterministic tests of transitivity. All data was analyzed with MATLAB (Mathworks). We used the set A choices to perform deterministic tests of transitivity, replicating previous studies. We first determined the preference ordering within each category for each subject. The 10 or 11 options within each category were ranked according to how many times each was chosen by that subject. Then, for each trial, a choice was counted as intransitive if a lower-ranked item was chosen over a higher-ranked item. Following Henri-Bhargava et al. (2012), ties were maintained in the rankings (i.e., more than option could have the same rank) to provide a more conservative definition of intransitive choices. Because the intransitive choice counts are not normally distributed, we used non-parametric statistics to test for group differences. We used Kruskal-Wallis tests to detect effects between groups, followed by one-tailed Wilcoxon ranked sum *post hoc* pairwise tests as appropriate (as several previous studies have found increased intransitive choices after VMF damage, we had strong hypotheses about the direction of the results). To test for within-subject effects, we used repeated measures analysis of variance (ANOVA) on rank-transformed data for the omnibus test and Wilcoxon signed-rank *post hoc* tests as appropriate.

Probabilistic tests of transitivity. We used the set B choices to perform probabilistic tests of transitivity, extending on previous studies. We first obtained the proportion of choices (out of a possible total of 15 choices) for each of the 10 choice pairs afforded by all possible pairings of the 5 options in each category. We then tested the random mixture model of preference by noting whether the choices violated the linear ordering polytope (LOP) (Regenwetter et al., 2011). The random mixture model states that a person's response comes from a probability distribution over all possible orderings of the stimuli. Thus, at any one time, preferences are transitive, but the transitive state that one is in can vary. The probability of a person choosing one option (X) over another (Y) in a binary choice is the sum of all the preference states in which X is preferred to Y. In a two alternative forced choice task, this is constrained by the triangle inequalities. For every distinct X, Y, and Z in a choice set:

$$P_{xy} + P_{yz} - P_{xz} \leq 1$$

Where P_{xy} denotes the probability of choosing X over Y, etc. For up to 5 options in a 2AFC task, satisfying the triangle inequalities, which together define the LOP, is necessary and sufficient for a set of choices to be consistent with the random mixture model.

For choice probabilities that did not satisfy the triangle inequalities, we used the Q-test (Regenwetter et al., 2014) software to determine whether the data were significantly outside of the LOP. Q-test uses maximum likelihood estimation to find the goodness of fit of the data at each vertex in the polytope, using a chi-squared bar distribution with simulated weights (Regenwetter, Dana, & Davis-Stober, 2010; Regenwetter et al., 2014). Any subject with choices in a category that produced $p < 0.05$ in this test were considered as significantly violating the LOP and thus, the random mixture model of preference.

Sensitivity of probabilistic tests. We performed several simulations to determine the sensitivity of the probabilistic test of transitivity, i.e., the rate at which this test would declare different forms of random or heuristic-based choice to be transitive. First, following Regenwetter et al. (2011), we randomly picked a choice probability for every pair from a uniform distribution (from 0 to 100%). As previously shown in Regenwetter et al. (2011), only about 5% of the choice datasets simulated in this manner satisfy the triangle inequalities. That is, only 5% of the possible set of choice proportions for 10 pairs/5 stimuli satisfy the random mixture model.

Second, we simulated an intransitive chooser who has an entirely consistent preference within each pair (i.e., choosing A 100% of time when it is paired with B) that is unconstrained by any higher order transitive structure (i.e., the preference in each pair is independent from that of all other pairs). This type of intransitive chooser only satisfies the triangle inequalities about 12% of the time for choice proportions for 10 pairs/5 stimuli as in our dataset.

Third, we simulated an intransitive chooser using the lexicographic semiorder heuristic (LS; Tversky, 1969). The LS heuristic is easiest to demonstrate with the gambles stimulus set. Following Tversky (1969), we defined our LS rule as follows: if two gambles are adjacent (i.e., next to each other in the set in terms of probabilities/payouts), always choose the gamble with the higher payout (amount); for all other (non-adjacent) gamble pairs, always select the gamble with the higher probability. Such a chooser would never satisfy the triangle inequalities in our dataset. Together, the first three sets of simulations show that our probabilistic test is very sensitive to different forms of intransitive choice.

Finally, we simulated a completely random chooser (i.e., someone who flips a coin on every single trial). The choice proportions for such a random chooser are given by the binomial probabilities with $p=0.5$. Such a chooser satisfies the triangle inequalities 80% of the time in our dataset (5 stimuli, 10 choice pairs repeated 15 times). This high percentage is not unexpected, as 50% choice probabilities across all pairs is consistent with the random mixture model (i.e., 0.5 + 0.5 -0.5 < 1). We use this rate below to assess whether the behavior of VMF subjects is consistent with completely random choice.

Drift diffusion modelling and analysis of reaction times. We calculated ranks of options similar to the method we used in the set A (deterministic transitivity) above, where the option that was chosen most often overall was ranked first, and the option chosen second-most was ranked second, etc., and broke ties by looking at which options were more often chosen more than half of the time in every pair (Henri-Bhargava et al., 2012). It was necessary to break ties here for the purposes of calculating the effect of value distance on reaction times (RTs). Three subjects still had tied ranks after this process, in one category each: two are HC subjects in the gambles domain, the other is a VMF patient from the Art domain. These subjects in these categories only are dropped from the ANOVA analysis and drift diffusion modelling below.

We fit a drift diffusion model (Ratcliff, 1978) to the choices and RTs from all set B choices for every other subject and category in our experiment. We modelled the decision process as a decision variable (DV) that increased linearly with a slope $d*v^{\alpha}$, where d was the drift rate, v was the value difference of the options (expressed as the absolute rank difference between the two items for that individual), and α was an exponent accounting for potential non-linearities in the effect of rank difference. We also assume that at each time step there is Gaussian noise added to the DV, with a standard deviation of ε. We assumed 10ms time steps. We also assume there is a non-decision time (*ndt*) before accumulation begins, and an initial value (*int*) of the DV that is constant across trials. Choices are made when the DV crosses a threshold.

Thus there are five free parameters: *d,* α, ε , *int* and *ndt*. Note that the threshold was a fixed parameter across subjects, as one of the threshold, *d*, or ε must be fixed for the other two parameters to be estimable. We chose to fix threshold after a model-comparison process showed that option to provide the best model fits. Threshold was held constant at (+/-) 0.15. Values for *d* are sampled between 0 and 1, for ε are sampled between 0 and 1, for α are sampled between 0 and 3, for *int* are sampled between the threshold bounds, and for *ndt* are sampled between 0 and the minimum RT minus 10ms for that subject.

To fit these free parameters, we first calculated the cumulative probability that the DV crossed the threshold for the subject's choice (T_{correct} or $T_{\text{incorrect}}$, where "correct" was defined as choosing the option of higher rank) across all time steps. For each trial, we then calculated the joint likelihood of the subject's choice at the time which they made that choice (their trial RT, minus *ndt*), by taking the derivative of this cumulative probability at the timestep of the subject's choice (every 10ms to the maximum RT for the subject). The model was then fit using the MATLAB function *fmincon*, where the cost function was defined as the sum of the negative log likelihoods of the instantaneous probabilities of the subject's choices and RTs in all trials. The fitting procedure was repeated 10 times for each subject, with each iteration varying in randomly sampled starting values for the free parameters as specified above; the parameters with the lowest log likelihood out of the 10 was taken for that subject. The model was fit individually to each of the three reward categories (art, brands, gambles) for each subject.

To look at differences in DDM parameters between groups across categories, we performed a mixed ANOVA on each of the free parameters, with group as the cross-subject factor and reward category as the within-subject factor.

Finally, we performed a mixed ANOVA with group and value distance as factors to look for the effect of value distance on RTs across groups.

Results

Deterministic Tests of Transitivity

Individuals with frontal damage exhibit more choice cycles. A subset of the choices in our experiment, Set A, consists of a single instance of all pairwise choices from a total of nine or ten items within a category, which allows us to first replicate two previous studies of transitivity (Fellows and Farah, 2007; Henri-Bhargava et al., 2012). Combining all three categories (art, brands, gambles) in our experiment, we replicate the finding that individuals with VMF damage make more intransitive choices, though we do not replicate that this effect is selective to VMF damage in the frontal lobe. There was a moderate difference in intransitive choices in set A summed across all three categories (Kruskal-Wallis H = 5.05, p = 0.08). Because three previous studies have found increased intransitive choices after VMF damage (Fellows and Farah, 2007; Henri-Bhargava et al., 2012; Camille et al., 2011), we conducted planned comparisons between groups. Similar to previous studies, our VMF group (mean = 9.93%, sd =6.65) made more intransitive choices than the HC group (mean = 5.71%, sd = 4.05; Wilcoxon ranked sums Z = 1.64, p = 0.05). Unlike previous studies though, our FC group (mean = 9.09%, sd = 3.74) also made more intransitive choices than the HC group (Z = 2.05, p = 0.02) and the difference between VMF and FC and was not significant (Z = 0.12, p = 0.45).

Figure 2a.

Differences among reward categories. However, the analysis above obscures differences across individuals and choice categories that point to more specific effects of VMF damage. We first examined how intransitive choices in set A differ across choice categories. In the one choice category used in previous studies of transitivity, art, there was significant difference in intransitive choices across groups (Kruskal-Wallis H = 7.62, $p = 0.02$), which replicated the previously reported pattern of selective VMF deficit. The VMF group (mean = 9.93%, sd = 1.86) made significantly more intransitive choices in the art category than both the FC group (mean = 4.73%, sd = 1.36; Wilcoxon ranked sum Z = 1.91, $p = 0.03$) and the HC group (mean = 3.64%, sd = 0.97; Wilcoxon ranked sum Z = 2.62, $p = 0.004$). In contrast, in the two categories that have not been used in previous studies, brands and gambles, we did not find significant differences between the three groups (brands, H = 2.42, $p = 0.29$; gambles, H = 3.01, $p = .22$ respectively).

In **Figure 2b-d,** it appears that number of intransitive choices is relatively stable across categories in the VMF and HC groups, but variable across categories in the FC group. Indeed, the effect of reward category is significant for the FC group ($F(2,18) = 3.88$, $p = 0.04$), but not for the VMF ($p = 0.92$) or the HC group ($p = 0.27$). In the FC group, the number of intransitive choices in the gamble category was significantly greater than in the art category (Z = 2.40, $p = 0.02$), while the differences between gambles and brands ($p = 0.19$) and art and brands (p = 0.18) were not significant.

Differences among individuals. We then examined how intransitive choices in set A differ across individuals. To do this, we considered each individual with a VMF or FC lesion as a single case, and compared their total number of intransitive choices (i.e., across all three categories) against healthy controls. We made this comparison using case-control t-tests (Crawford & Howell, 1998) which are modified to compare an individual against a normative group when the sample size is small. In the VMF group, four individuals made significantly more intransitive choices than healthy controls, before corrections for multiple comparisons (Subject 350: t(19) = 2.04, *p* = 0.03; Subject 10403: t(19) = 3.28,*p* = 0.002, Subject 12402: t(19) = 3.13, *p* = 0.003; Subject 775: t(19) = 3.13, *p* = 0.003). These differences remained significant in the latter three individuals after correcting for multiple comparisons using FDR (corrected *p = 0.023* for all three individuals*).* Lesion extent of these three subjects are shown in **Figure 3.** In contrast, in the FC group, none of the individuals made significantly more intransitive choices than healthy controls (all *p* >= 0.05 before multiple comparison correction).

This result suggests that a subset of individuals with VMF damage show the most pronounced increase in intransitive choices. However, we did not find evidence to support any particular account of this heterogeneity. The total number of intransitive choices (i.e., across all three categories) was not significantly correlated with lesion size (in cc's), whether considering all subjects with lesions (Spearman's rho = -0.14, *p* = 0.51) or only those with VMF damage (rho = -0.13, *p* = 0.67). Within the VMF group, the total number of intransitive choices was also not significantly correlated with lesion volume within a vmPFC mask defined based on value effects in fMRI studies (Bartra, McGuire & Kable, 2013; rho = -0.06, p = 0.83). Finally, across all subjects, the total number of intransitive choices was not significantly correlated with any of the demographic variables (gender, point biserial r = 0.13, p = 0.39; age, rho = 0.14, p = 0.35; education, rho = 0.24, p = 0.11).

Probabilistic Tests of Transitivity

Individuals with VMF damage make choices consistent with probabilistic models of transitivity. After replicating the finding that individuals with VMF damage make an increased number of intransitive choices, we next turned to the central question motivating our study, which is whether or not the choices of these individuals violate probabilistic notions of transitivity. To do this, we examined the subset of choices in our experiment, Set B, which involve 15 repetitions each of 10 different binary choices in each of the three categories. Set B provides sufficient data for evaluating whether the choices each participant made are consistent with the random mixture model, a probabilistic model of transitive choice. None of the individuals with VMF damage violated the random mixture model in any of the three domains (a total of 39 tests, see Table 2). Similarly, none of the individuals with frontal damage outside the VMF violated the random mixture model in any of the three domains (a total of 36 tests).

Interestingly, two healthy controls significantly violated the random mixture model in the gambles domain (p = 0.002 and p = 0.01, respectively). One of these individuals followed Tversky's (1969) lexicographic semiorder heuristic exactly and the other followed this heuristic partially. Their results demonstrate the sensitivity of our test to detect individuals choosing on the basis of attribute-based heuristics that lack higher order transitive structure.

Individuals with VMF damage are not choosing randomly. One possible explanation for why individuals with VMF damage conform to probabilistic models of transitivity despite making a greater number of individual intransitive choices is that they are simply choosing randomly, as completely random choices fulfill the random mixture model 80% of the time in our experimental design (see methods). However, individuals with VMF damage are not simply choosing randomly. First, the probability that a group of random choosers the size of the VMF group (N=13) would all make choices consistent with the random mixture model in all three domains is extremely low, *p = 1.66e-04*. Second, we can evaluate directly the likelihood that an individual is choosing randomly by comparing their choice proportions (N=10 in each category) against those expected under the binomial distribution. For every single individual with VMF damage, and in all three domains, the likelihood that their choice proportions arose from completely random choice was extremely low (all p < 1e-06).

Individuals with VMF damage do not have systematically different preferences. A second possible explanation for why individuals with VMF damage conform to probabilistic models of transitivity despite making a greater number of individual intransitive choices is that they have systematically different preferences. For example, we might expect that a risk-neutral chooser would be more likely to make occasional intransitive choices in our gambles category than a strongly risk averse chooser. However, individuals in the VMF group did not make systematically different types of choices than individuals in the other groups. In a MANOVA on the choice proportions for each of the 10 binary choices the participants faced in each category, there were no significant differences between groups in the art category [Wilks' Lambda = 0.64, F(18,64) = 0.9, p = 0.58], the brand category [Wilks = 0.64, F(18,64)=0.90, p = 0.58], or the gambles category [Wilks = 0.46, F(18,64) = 1.67, p = 0.07].

Individuals with VMF damage have noisier preferences. A third possible explanation for why individuals with VMF damage conform to probabilistic models of transitivity despite making a greater number of individual intransitive choices is that they are noisier choosers. That is, their choices reflect underlying transitive preference orderings, but they vacillate among preference orderings more than other choosers. To further test this possibility, we fit each individual's choices and RTs in Set B to a drift diffusion model (DDM), which assumed that choices and RTs were a probabilistic function of the rank distance in preference ordering between the two options. These fits revealed that individuals with VMF damage were noisier choosers. The only parameter of the DDM that was significantly different across groups was the noise parameter ε [F(2,37) = 6.25, p = 0.005]. Specifically, the VMF group (mean = 0.12, sd = 0.03) had significantly higher ε than HC (mean = 0.09, sd = 0.04)[t(28) = 2.08, p = 0.047] and FC (mean = 0.07, sd = 0.02) [t(20) = 3.94, p < 0.001]. No other parameters differed between the three groups. **Figure 4.**

Individuals with VMF damage show a less pronounced effect of value on reaction times. RTs in individuals with VMF damage also showed a less pronounced effect of ranked value distance, consistent with the increased noise parameter observed in the DDM fits. We performed a mixed ANOVA on median RTs with value distance and group as factors. We found a significant main effect of value distance [F(3, 111) = 28.63, $p < 0.0001$], a significant main effect of group [F(2,37) = 4.93; $p = 0.01$], and a significant interaction between the two [F(6,111) = 3.76; $p = 0.002$].

The significant effect of value distance reflected the expected decrease in RTs as the distance in preference ordering rank gets larger. The average median RT for a rank difference of 1 (mean = 2800ms, sd = 1458) was significantly slower than a rank difference of 2 (mean = 2500ms, sd = 1211) [Z = 4.86, p <0.0001], which in turn was slower than the rank difference of 3 (mean = 2300ms, sd = 1166) [Z = 3.59, p <0.001], which in turn was slower than a rank difference of 4 (mean = 2180ms, sd = 1045) [Z = 2.78, p=0.005].

The effect of group reflected longer RTs in the FC group. RTs in the FC group (mean = 3380ms, sd = 1596 ms) were significantly slower than in VMF group (mean = 1883ms, sd = 469ms) (Z = 3.13, $p = 0.002$), and a similar slowing relative to the HC group (mean = 2439ms, sd = 1116ms) exhibited a non-significant trend (Z = 1.70, $p = 0.09$). RTs in the VMF and HC groups were not significantly different (Z = 1.16, $p = 0.24$).

The interaction between value distance and group reflected a reduced effect of value distance on RTs in the VMF group. We took the Spearman correlation between RT and the difference in preference ordering rank as an index of the value distance effect. The VMF group (mean rho = -0.16) exhibited a flatter value distance-RT relationship than the HC group (mean rho = -0.22) [t(28) = 2.20; $p = 0.04$]. The value distance-RT relationship in the FC group (mean rho = -0.19) was intermediate and not significantly different from the VMF ($p = 0.42$) or HC ($p = 0.61$) groups. As shown in **Figure 5,** these differences can be accounted for by the DDM fits described above.

Discussion

Individuals with damage to the ventromedial frontal lobes (VMF) have been shown previously to be more inconsistent in their choices (Camille et al., 2011; Fellows & Farah, 2007; Henri-Bhargava et al., 2012). These previous findings, however, are consistent with two possible patterns of behavior, with very different implications for the function of the VMF. One possibility is that individuals with VMF damage are fundamentally intransitive: that they reliably choose in an intransitive manner when given the same choice between the same options repeatedly. A second possibility is that individuals with VMF damage are more variable in their choices, yet still fundamentally transitive. Here we distinguished between these two possibilities by testing whether the choices of individuals with VMF damage satisfy probabilistic notions of transitivity, as the first possibility predicts they do not and second predicts they do. We overwhelmingly find evidence for the second possibility, as all individuals with VMF damage make choices in all domains that are consistent with probabilistic models of transitivity.

The first possibility, that individuals with VMF damage are fundamentally intransitive choosers, implies that the VMF is necessary for choices to be value-based, as transitivity is the key hallmark of a value-based choice (Samuelson, 1937; Von Neumann & Morgenstern, 1945). According to this view, individuals with VMF damage would only be able to choose in a non-value-based manner, for example, according to rules or heuristics. Our data, however, provide strong evidence against this possibility. This result is difficult to reconcile with the view that VMF is *the* critical substrate for value-based choice.

In contrast, we found strong evidence for the second possibility, that individuals with VMF damage are fundamentally transitive, that is their choices satisfy probabilistic models of transitivity, even though they make more intransitive choices according to deterministic notions of transitivity. Furthermore, we showed that this pattern was not due to individuals with VMF damage choosing in an entirely randomly manner, nor was it due to these individuals having preferences that were systematically different from those of the other groups. Rather, this pattern was due to individuals with VMF damage being noisier or more variable choosers. This is consistent with the suggestion of Henri-Bhargava et al. (2012), that “values are unstable, fluctuating from trial to trial in those with VMF damage.” We illustrated this by fitting a drift diffusion model (Ratcliff, 1978) to each individual’s choices. In this model, the VMF group had a significantly higher noise term, i.e., more variance around the decision variable, than healthy individuals or those with frontal damage outside the VMF. Importantly, the VMF group did not differ on the value of any other parameters. Reaction times in the VMF group were also similar to healthy controls, arguing against accounts of their behavior based on impulsivity (faster RTs) or indecision (slower RTs). Overall our modeling further strengthens the conclusion that the

VMF serves to make preferences more stable, so that individuals would be less likely to select an option that is typically less preferred.

These results are easier to reconcile with a framework in which valuation and value-based choice are distributed processes, to which multiple regions of the brain contribute in some respect (Hunt & Hayden, 2017). This framework would predict that others regions can compensate for damage to the VMF, so that such damage does not fundamentally abolish the transitivity of preferences. The modest effect size in deterministic tests, which is typically an increase of around 5% in the number of intransitive choices in the VMF group relative to control groups in our study and previous ones (Fellows & Farah, 2007; Henri-Bhargava et al., 2012), is also more consistent with this view. As making transitive choices that maximize value is incredibly important to the survival of an organism, it would make sense that value is a highly conserved process that is not abolished by damage to one part of the cortex. Future studies could more directly test hypotheses about compensation by examining activity in inconsistent individuals with fMRI, as it is also possible that regions that compensate are in the still intact parts of VMF rather than in other regions entirely.

Our results do not speak to how exactly the VMF supports choice stability. One possibility is that VMF contributes some part of the composition of subjective value. If subjective value is computed through the interaction of several brain regions, the loss of VMF may make this computation noisier and less reliable, akin to the greater noise we see in our DDM results. Alternatively, as a flattening of the value distance-RT relationship is consistent with greater indifference between options, the VMF could amplify or enhance the differences in value between different options (Henri-Bhargava et al., 2012). It is also possible that the VMF contributes a unique, specific component to valuation. For example, it has been suggested that the VMF contributes emotional content when making aesthetic judgments (Vaidya, Sefranek, & Fellows, 2017), and in other contexts that it contributes motivational salience that can distinguish close options from one another more clearly (Manohar & Husain, 2016; Pujara, Philippi, Motzkin, Baskaya, & Koenigs, 2016; Vaidya & Fellows, 2015b).

Another broad set of possibilities can be generated by considering the nature of the random mixture model that individuals with VMF damage satisfy. In this model, choosers are allowed to have different preference orderings in different contexts or at different points in time. It is possible, therefore, that VMF somehow contributes to the same preference ordering being repeated reliably. For example, individuals might use episodic memories of their previous choices (e.g., “I remember choosing A over B before”) to guide their decisions. Although we tried to reduce the influence of such memories, it is difficult to eliminate their influence entirely (Birnbaum, 2011) and VMF has been implicated in episodic memory processes (Bertossi, Tesini, Cappelli, & Ciaramelli, 2016). Alternatively, VMF could support a representation of the context of the experiment that in turn activates a specific set of preferences, such as in a schematic network. Consistent with this idea, previous work has shown VMF involvement in schema formation (Schlichting & Preston, 2016; Spalding et al., 2018).

Finally, we extended previous studies that considered only deterministic notions of transitivity by identifying heterogeneity in these effects both across individuals and across domains. There was considerable heterogeneity within the VMF group, where some participants made as few intransitive choices as healthy controls, while other participants made significantly more intransitive choices. We did not find any systematic differences in lesion location or size that accounted for this heterogeneity. The lesions of the three individuals in the VMF group who made significantly more intransitive choices overall did not overlap much in their location, and the overlap areas were in the same location where other individuals had sustained lesions. The lesions of the three most inconsistent individuals in the VMF group did tend to extend more posteriorly towards the basal forebrain and ventral striatum, though given the sample size in our study this potential explanation will need to be rigorously evaluated in future work with a larger number of subjects. Future studies could also test alternative explanations that we were unable to assess by using more advanced imaging to test whether damage to specific white matter tracts or disruptions in specific connectivity networks are linked to making more intransitive choices.

There was also considerable heterogeneity across domains, with the pattern of intransitive choices being most consistent with previous studies (i.e., showing a deficit selective to VMF damage) in the one the domain, art, that had been used in those studies. The greatest heterogeneity across domains, though, was in the frontal control group. This group looked similar to healthy controls in the art domain but made the most number of intransitive choices in the gamble domain. The frontal control group includes individuals with damage to the dorsomedial or dorsolateral prefrontal cortex, and both of these regions have been previously shown to be involved in decisions about risk (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009; Hsu, Krajbich, Zhao, & Camerer, 2009). Previous studies have started to consider how the brain regions necessary for preference consistency may vary across domains (Fellows & Farah, 2007; Henri-Bhargava et al., 2012), and our results further highlight the need to examine a variety of domains in future work.

In conclusion, we found that individuals with VMF damage make choices that are noisier, but still fundamentally transitive. This result both characterizes how erratic choices manifest after damage to the VMF (Eslinger & Damasio, 1985; Harlow, 1868), as well as potentially explains why studies using similar decision-making paradigms in individuals with VMF-damage can yield different results (Fellows, 2011). In addition, our findings further clarify and define the necessary role the VMF plays in value-based decision-making. Specifically, though each choice still reflects some subjective preference ordering after VMF damage, an intact VMF is necessary for preference orderings to remain stable and reliable across time and contexts.

References

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage, 76*, 412-427.

Bertossi, E., Tesini, C., Cappelli, A., & Ciaramelli, E. (2016). Ventromedial prefrontal damage causes a pervasive impairment of episodic memory and future thinking. *Neuropsychologia*. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2016.01.034

Birnbaum, M. H. (2011). Testing mixture models of transitive preference: Comment on Regenwetter, Dana, and Davis-Stober (2011).

Camille, N., Griffiths, C. A., Vo, K., Fellows, L. K., & Kable, J. W. (2011). Ventromedial frontal lobe damage disrupts value maximization in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 31*(20), 7527-7532. doi:10.1523/JNEUROSCI.6527-10.2011

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *Journal of Neuroscience, 29*(40), 12574-12583.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*(4), 482-486.

Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation. *Patient EVR, 35*(12), 1731-1731. doi:10.1212/wnl.35.12.1731

Fellows, L. K. (2011). 16 The Neurology of Value. *Neurobiology of sensation and reward*, 351.

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral cortex (New York, N.Y. : 1991), 17*(11), 2669-2674. doi:10.1093/cercor/bhl176

Fellows, L. K., Stark, M., Berg, A., & Chatterjee, A. (2008). Patient registries in cognitive neuroscience research: Advantages, challenges, and practical advice. *Journal of cognitive neuroscience, 20*(6), 1107-1113.

Harlow, J. M. (1868). Recovery from the passage of an iron bar through the head. *Publications of the Massachusetts Medical Society, 2*(3), 274-281.

Henri-Bhargava, A., Simioni, A., & Fellows, L. K. (2012). Ventromedial frontal lobe damage disrupts the accuracy, but not the speed, of value-based preference judgments. *Neuropsychologia, 50*(7), 1536-1542. doi:10.1016/j.neuropsychologia.2012.03.006

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *Journal of Neuroscience, 29*(7), 2231-2237.

Hunt, L. T., & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews Neuroscience, 18*(3), 172.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology, 22*(6), 1027-1038. doi:https://doi.org/10.1016/j.conb.2012.06.001

Luce, R. (1959). Individual Choice Behaviour. A Theorical Analysis: New York. Willey.

Manohar, S. G., & Husain, M. (2016). Human ventromedial prefrontal lesions alter incentivisation by reward. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 76*, 104-120. doi:http://dx.doi.org/10.1016/j.cortex.2016.01.005

McFadden, D. (1980). Econometric Models for Probabilistic Choice Among Products. *The Journal of Business, 53*(3), S13-S29.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226.

Pujara, M. S., Philippi, C. L., Motzkin, J. C., Baskaya, M. K., & Koenigs, M. (2016). Ventromedial prefrontal cortex damage is associated with decreased ventral striatum volume and response to reward. *Journal of Neuroscience, 36*(18), 5047-5054.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2010). Testing transitivity of preferences on two-alternative forced choice data. *Frontiers in psychology, 1*, 148.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review, 118*(1), 42-56. doi:10.1037/a0021150

Regenwetter, M., Davis-Stober, C. P., Lim, S. H., Guo, Y., Popova, A., Zwilling, C., . . . Messner, W. (2014). QTest: Quantitative testing of theories of binary choice. *Decision, 1*(1), 2.

Samuelson, P. A. (1937). A note on measurement of utility. *The review of economic studies, 4*(2), 155-161.

Schlichting, M. L., & Preston, A. R. (2016). Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiology of Learning and Memory, 134*, 91-106.

Spalding, K. N., Schlichting, M. L., Zeithamova, D., Preston, A. R., Tranel, D., Duff, M. C., & Warren, D. E. (2018). Ventromedial prefrontal cortex is necessary for normal associative inference and memory integration. *Journal of Neuroscience, 38*(15), 3767-3775.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*.

Vaidya, A. R., & Fellows, L. K. (2015a). Testing necessary regional frontal contributions to value assessment and fixation-based updating. *Nature communications, 6*, 10120. doi:10.1038/ncomms10120

Vaidya, A. R., & Fellows, L. K. (2015b). Ventromedial frontal cortex is critical for guiding attention to reward-predictive visual features in humans. *The Journal of Neuroscience, 35*(37), 12813-12823. doi:http://dx.doi.org/10.1523/JNEUROSCI.1607-15.2015

Vaidya, A. R., Sefranek, M., & Fellows, L. K. (2017). Ventromedial Frontal Lobe Damage Alters how Specific Attributes are Weighed in Subjective Valuation. *Cerebral Cortex*, 1-11.

Von Neumann, J., & Morgenstern, O. (1945). *Theory of games and economic behavior*: Princeton University Press Princeton, NJ.

Legends

Figure 1. Overlap images of the VMF and frontal control lesion groups. Numbers below slices indicate the MNI z-coordinates. Colors indicate extent of overlap. L = left; R = right.

Figure 2. Group average and individually plotted intransitive choices in (deterministic) set A across a) all domains, and b-c) in each reward domain. Filled-in dots encircled in gray denote the VMF subjects whose errors were significantly higher compared to the HC group, and whose lesion extents are depicted in Figure 3. Error bars are standard errors of the mean.

Figure 3. Lesion tracings of the three individuals with VMF lesions who had significantly more intransitive choices compared to healthy control subjects, as determined by case-control t-tests. Red denotes areas where at least one of these subjects had a lesion; yellow denotes the areas where at least one of these subjects had lesions *outside* of all other lesion subjects. There

was very little overlap in lesions within the three subjects (only maximally two out of three on only in a small number of voxels). Numbers below axial slices indicate the MNI z-coordinates.

Figure 4. DDM parameter fits: noise, drift rate, initial starting point, non-decision time, and alpha (exponent on rank distance). Error bars are standard errors of the mean.

Figure 5. Value distance effect on RT, by group. Dotted line are simulated RTs from DDM parameter fits. Error bars are standard errors of the mean.

Tables

Group (n)	Gender	Mean Age (sd)	Education in yrs
VMF (13)	7F:6M	59 (15)	14
FC (10)	7F:3M	66 (8)	14
HC (20)	15F:5M	62 (8)	15

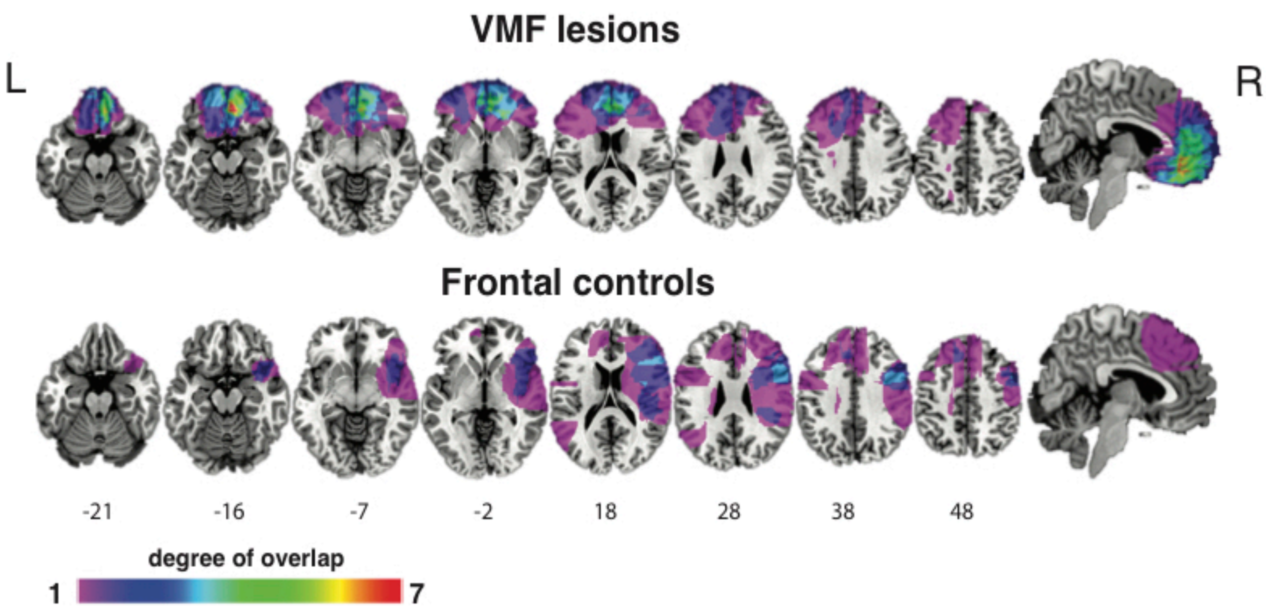
Table 1: demographics of participants.

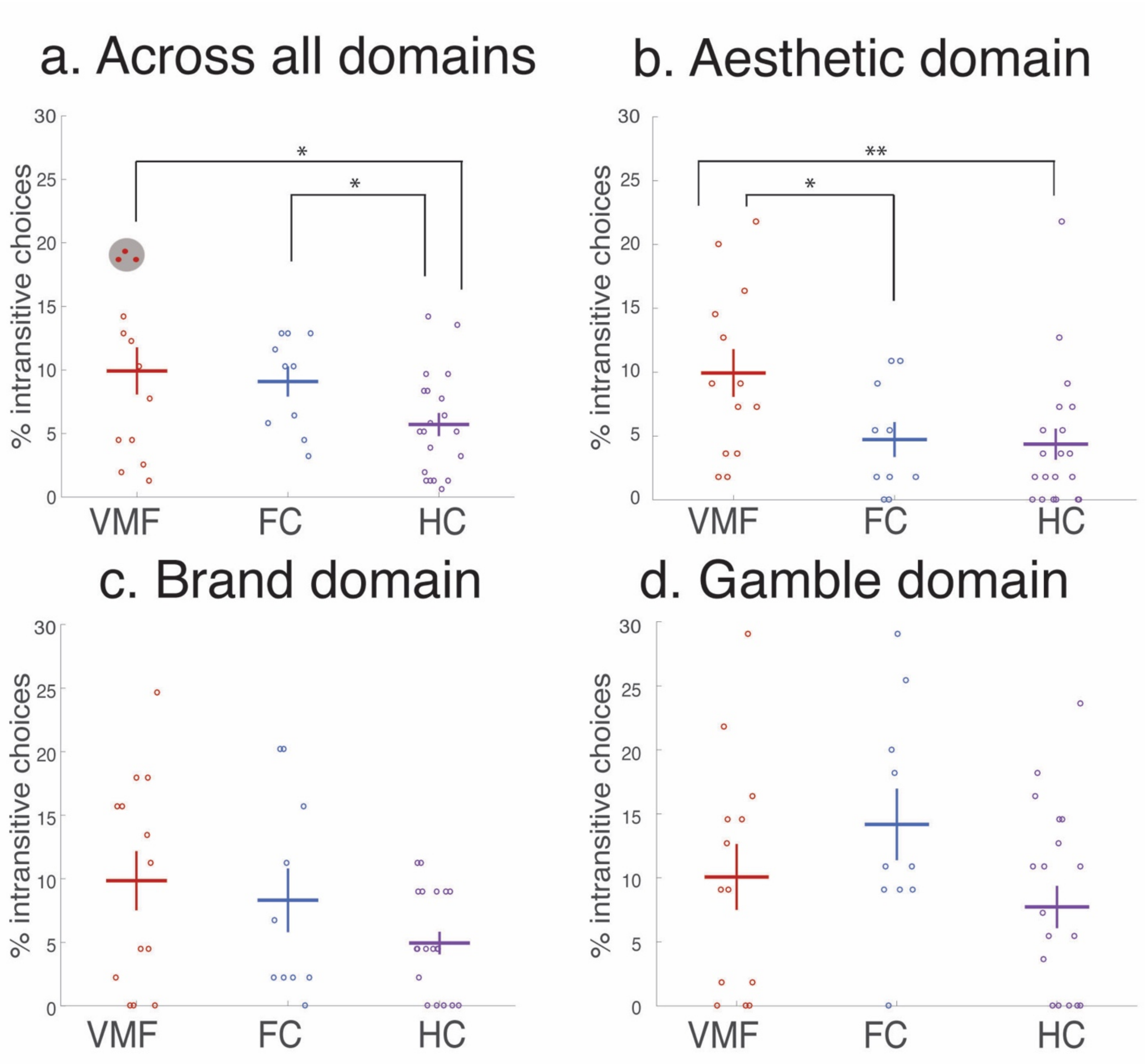
Group	Art	Food	Gamble
VMF	0/13	0/13	0/13
FC	0/12	0/12	0/12
HC	0/20	0/20	2/20

Table 2. Number of participants whose choices violated the random mixture model ($p < 0.05$)

Figures

Figure 1.





a. Across all domains
b. Aesthetic domain
c. Brand domain
d. Gamble domain
% intransitive choices
0
5
10
15
20
25
30
*
**
VMF
FC
HC

Figure 3.

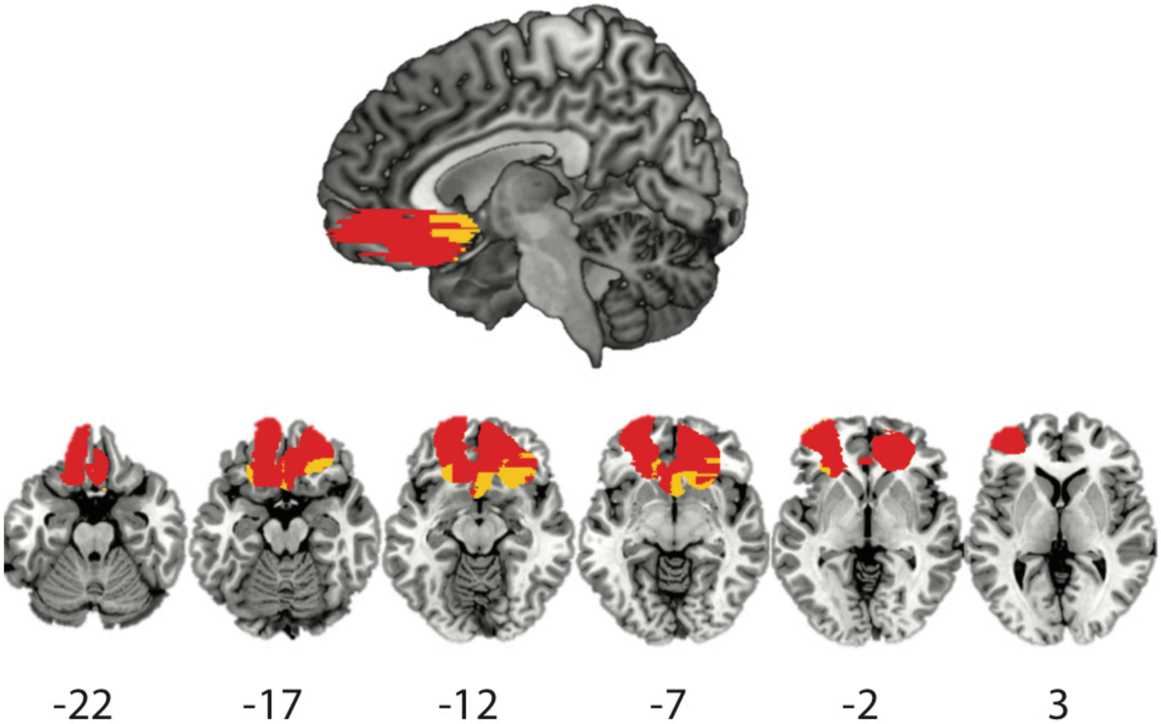


Figure 4.

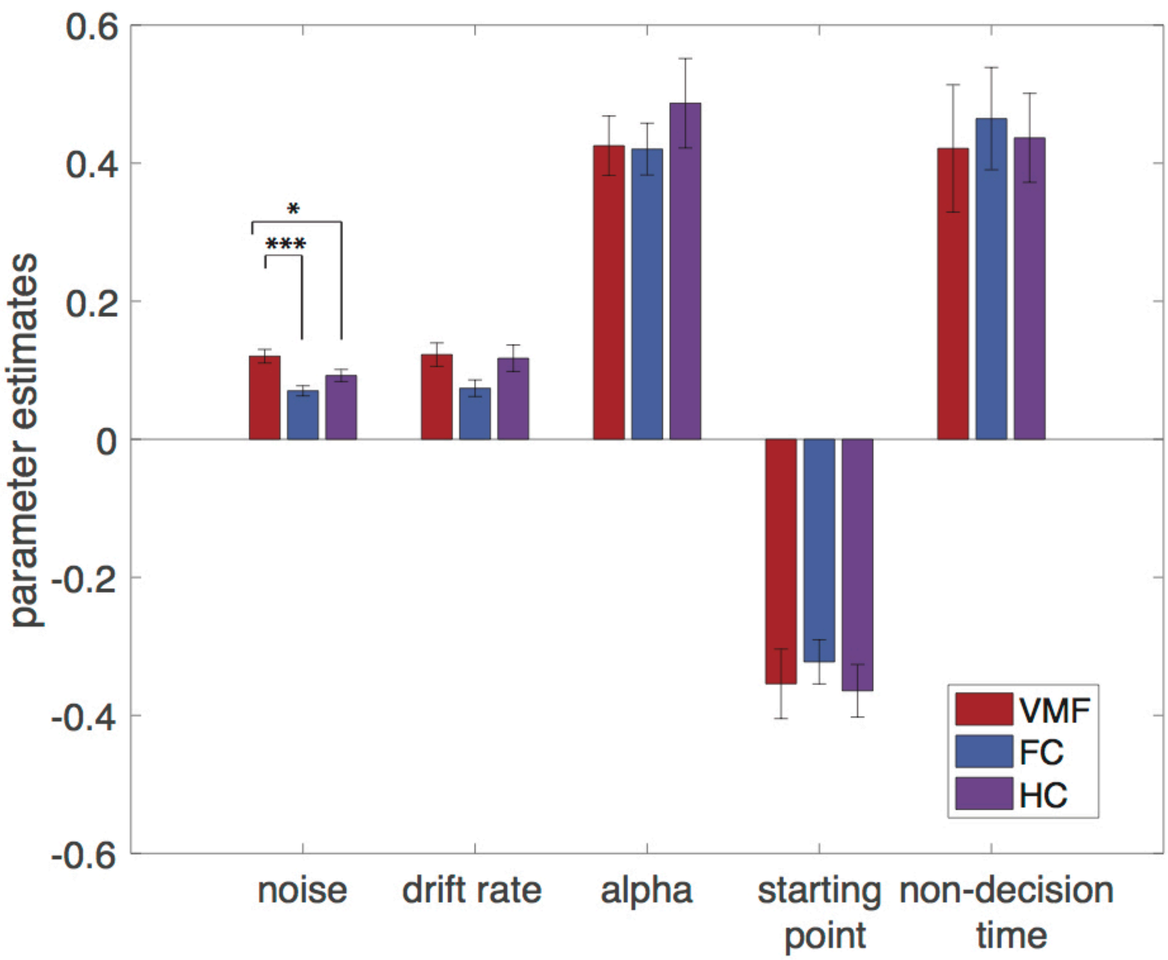
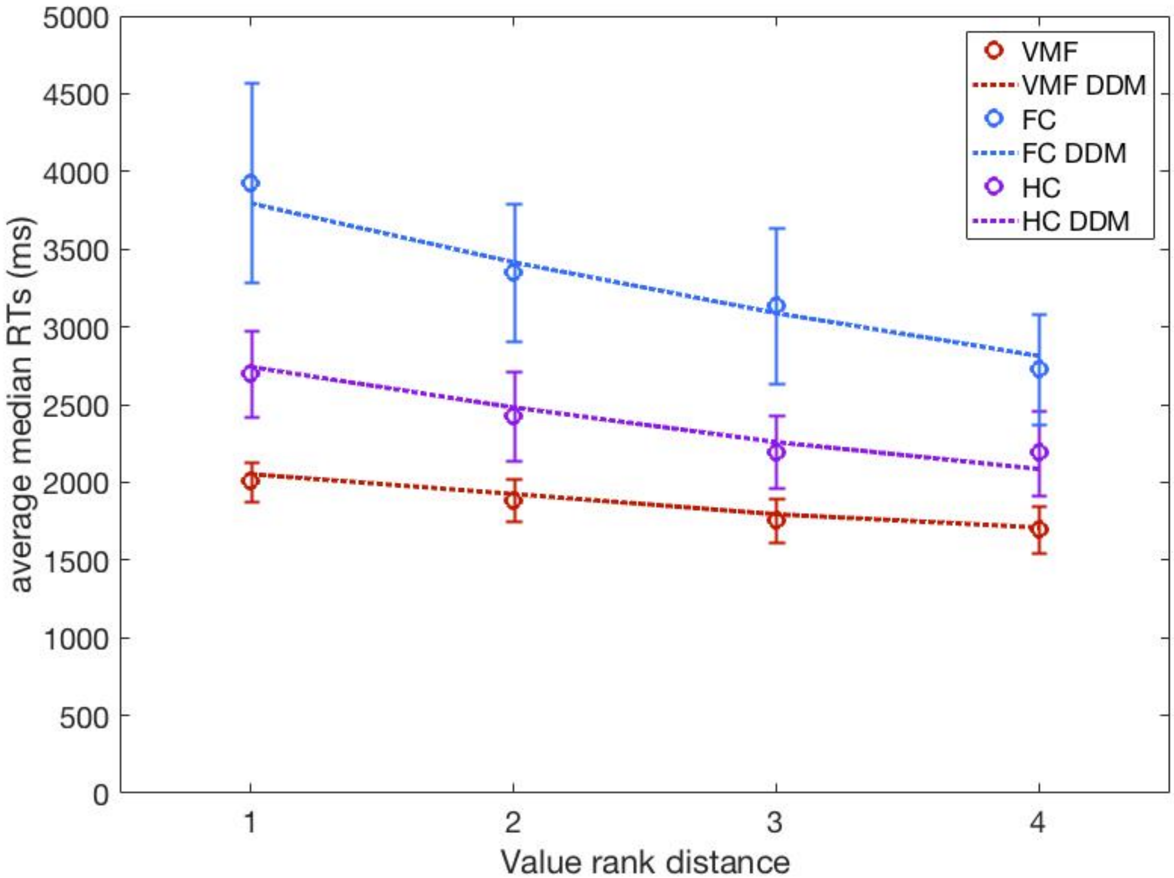

0.6
0.4
0.2
0
-0.2
-0.4
-0.6
parameter estimates
*
***
VMF
FC
HC
noise
drift rate
alpha
starting point
non-decision time

Figure 5.



5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
average median RTs (ms)
1
2
3
4
Value rank distance
VMF
VMF DDM
FC
FC DDM
HC
HC DDM