



HHS Public Access

Author manuscript

Nat Hum Behav. Author manuscript; available in PMC 2021 June 15.

Published in final edited form as:

Nat Hum Behav. 2020 November ; 4(11): 1118–1120. doi:10.1038/s41562-020-00941-5.

Divisive normalization *does* influence decisions with multiple alternatives

Ryan Webb¹, Paul W. Glimcher^{2,3}, Kenway Louie^{2,3,*}

¹Rotman School of Management and Department of Economics, University of Toronto, Toronto, Canada

²Center for Neural Science, New York University, New York, USA

³Neuroscience Institute, New York University School of Medicine, New York, USA

Divisive normalization (DN), an algorithm prominent in sensory processing and viewed as a canonical neural computation¹, has been recently proposed to play a similarly important role in decision making, capturing aspects of both neural value coding and choice behaviour^{2–13}. In a previous study, DN was proposed to explain context-dependent Independence from Irrelevant Alternatives (IIA) violations, where the relative preference between two given alternatives depends on the value of a third (distracter) alternative¹⁴; however, Gluth et al.¹⁵ conclude in a replication study that there is no effect of distracter value on choice behaviour in their dataset, instead finding an effect on reaction times. Here, we reanalyze the replication dataset, graciously made available by the authors, using econometric techniques specifically designed to test for the presence of divisive normalization in multi-alternative choice datasets; while we observe some differences between their data and the dataset originally reported by Louie et al.¹⁴, overall we find compelling evidence for a DN effect in choice in their dataset. This discrepancy arises primarily from unnecessarily restrictive assumptions used in prior analyses^{14,15}, emphasizing the importance of deriving the analysis directly from the choice theory itself.

Testing for normalization in choice data

In sensory processing, DN is a prevalent computation observed in multiple brain regions, sensory modalities, cognitive processes, and species from invertebrates to humans¹. Recent studies suggest that DN also plays a crucial role in the neural representation of subjective value, and consequently, choice behaviour. At the neural level, DN captures context-dependent value coding^{8,12}, characterizes reward-related neural dynamics⁹, and falls within the class of provably optimal methods for encoding value under neural noise¹⁰. At the behavioural level, DN predicts both within-choice and across-choice context effects observed in nematodes⁴, monkeys^{13,14}, and humans^{2,5–7,14}. An early example of DN-

*Corresponding author: Kenway Louie, Center for Neural Science, New York University, 4 Washington Place, Room 621, New York, NY 10003, klouie@cns.nyu.edu.

Author contributions

R.W. and K.L. analyzed the data. R.W., P.W.G., and K.L. wrote the paper.

Competing interests

The authors declare no competing interests.

predicted choice behaviour is a negative distracter effect in trinary choice, in which distracter value diminishes relative choice accuracy¹⁴. However, variable and even opposite distracter effects have also been observed^{2,3}, emphasizing the importance of replication studies and robust tests for DN in behaviour.

The Gluth et al.¹⁵ replication employed a three-alternative choice experiment that was analyzed with a binary Logit regression on the average valuations of each alternative V_1 , V_2 , and V_3 in the choice set. This model was estimated both at the subject level (with random effects) and via a hierarchical analysis that allows each subject's behaviour to arise from a distribution estimated at the level of the pooled sample. While a similar analysis approach was used in Louie et al.¹⁴, we have more recently noted that this method of analysis faces a number of challenges¹¹.

1. A logistic regression is derived from a model of binary choice. The reported experiment offers a choice between three alternatives. While including the third alternative is particularly important when choices of the lowest-valued alternative are frequent, it still improves model estimation even if the third alternative is never chosen. In the analyses reported by Louie et al.¹⁴ and Gluth et al.¹⁵, trials on which subjects chose the lowest valued option were simply dropped from the analysis and the binary Logit model was imposed on the remaining datapoints, a strategy we no longer employ.
2. DN, by design, specifies a non-linear relationship between inputs and outputs. A Logit regression imposes a linear relationship between the log-odds and the regressors (V_1 , V_2 , and V_3). Given that normalization necessarily predicts a non-linear effect on the log-odds, the Logit regression is thus a severe misspecification of the theory.
3. In order to test the normalization model using stochastic choice data, an assumption on the distribution of errors must be made. Louie et al.¹⁴ described the predictions of DN under an independent normally-distributed error term. Recently, we have employed a more sophisticated error model that has broadened our understanding of DN and the IIA violations it induces. For example, Webb et al.¹¹ examined the predictions of DN when the errors follow the Gumbel distribution (a standard noise assumption in discrete choice analysis) as well as the more flexible multivariate normal distribution. Ensuring the analysis is robust to assumptions on the error term is critical because any analysis of normalization in choice data is a *joint* test of normalization *and* the assumed error distribution¹¹.

A reexamination of the Gluth et al data

To relax Gluth et al.'s assumptions and to take a broader look at normalization effects in choice, we applied a simplified version of the econometric model described in Webb et al.¹¹ to the data from the replication experiment. This choice model uses the same random utility form that underlies standard Logit and Probit regression, but takes as arguments the normalized rather than unnormalized values.

$$U_i = \frac{V_i}{\sigma + \omega \sum_n V_n} + \varepsilon_i, \quad (1)$$

Importantly, this specification allows a fit parameter, $\omega \in [-\infty, \infty]$, to govern the degree of normalization in utility. Choice is determined by which of these utilities is largest, with the probability of choosing alternative i given by

$$P_i(V) = \int \mathbb{1} \left[\frac{V_i - V_j}{\sigma + \omega \sum_n V_n} > \varepsilon_j - \varepsilon_i, \forall j \neq i \right] f(\varepsilon) d\varepsilon \quad (2)$$

where $f(\varepsilon)$ is the joint probability density of the errors and must be assumed.

If fits of the model indicate that $\omega = 0$ and $f(\varepsilon)$ is Gumbel, then this indicates that there is no normalization and the model reduces to the standard multinomial Logit in which IIA must hold. Therefore the model allows for a nested hypothesis test (via log-likelihoods) for the presence of normalization, and by extension, whether IIA is violated. We emphasize that - if normalization is *not* present in a dataset - then ω would be zero and all of the variance in the choice data should be captured by the parameter σ . On the other hand, a non-zero ω would support a contextual choice effect captured by normalization: a positive ω replicates the divisive normalization result of Louie et al.¹⁴ whereby larger distracter values (V_3) decrease relative choice performance (P_1/P_2); a negative ω , while rarely employed in sensory normalization models and harder to interpret, predicts - in at least some scenarios - a contrasting facilitatory effect of distracter value that has also been previously reported^{2,3}.

We estimated two versions of this model. The first uses the dataset pooled over all subjects and considers three different distributions of the error term: the Gumbel, the independent standard normal (N), and the multivariate normal (MVN). Across *all* of them, we find a highly significant effect of normalization (Table 1). The best-performing model uses the Gumbel distribution, where we find a departure from IIA with ω significantly larger than zero ($p = .0041$, $\chi^2(1)$); note that in this analysis, ω was not constrained to be positive. Normalization accounts for 14% of the variance in choice behaviour (calculated at the average of V_1 , V_2 , and V_3). Figure 1 shows the estimated probability ratio P_1/P_2 as V_3 is varied for the best-fitting model from the two datasets in question: DN with MVN errors for the Louie et al. dataset (i.e. as previously described¹¹), and DN with Gumbel errors for the Gluth et al. dataset (i.e. the estimates reported in Table 1). Two findings are evident: (1) clearly, including normalization allows the model to capture an initial decrease in P_1/P_2 as V_3 increases in both datasets, and (2) the pattern of IIA violations differs between the datasets, with a notable u-shaped non-monotonicity in the Louie et al. dataset. Note that the latter non-monotonicity is consistent with both computational simulations of a normalization model and observed choice patterns in the original dataset¹⁴.

Our second analysis estimates a hierarchical version of the model that allows both σ and ω to follow a distribution over subjects (i.e. a random effect for both parameters). In this analysis, we assumed Gumbel errors and allowed the ω parameter to follow a Normal distribution over subjects; the σ parameter followed a Gamma distribution. The mean of the

distribution for $\omega = 0.007$ was significantly positive ($p < .0001$, $\chi^2(2)$), and allowing the density for ω to depart from zero captures 36% of the variance in choice behaviour (calculated at the average of V_1 , V_2 , and V_3 and the mean of the densities for ω and ϕ).

Conclusion

In our re-examination of the Gluth et al. data, all of our analyses display significant evidence of normalization, regardless of assumptions. We believe this discrepancy arises from two factors. First, we employed a more rigorous regression approach that unambiguously indicates that DN is observed in these data. Second, we examined the complete pattern of choice behaviour implied by the model fits and identified a differences between the original and replication datasets. This difference in observed choice behaviour may reflect a general difference between the datasets, consistent with the reported difference in reaction time effects. Furthermore, this difference aligns with the reported variability in the nature of trinary IIA violations³ and may reflect underlying complexity in the generating mechanisms. The shape of IIA violations is a critical and deeply interesting question that we applaud Gluth and colleagues for raising. Like these authors, we are devoting significant energy to understanding the experimental conditions under which IIA violations occur and the proper methodology for analyzing them. However, we believe this new dataset does not materially challenge the evidence for DN: In all of the specifications we estimated, we observed a significant role for normalization in accounting for the variance in choice behaviour. This leads us to the conclusion that in this new dataset, as in others, divisive normalization influences decisions with multiple alternatives.

References

1. Carandini M. & Heeger DJ Normalization as a canonical neural computation. *Nat Rev Neurosci* 13, 51–62, (2012).
2. Chang LW, Gershman SJ & Cikara M. Comparing value coding models of context-dependence in social choice. *Journal of Experimental Social Psychology* 85, 103847, (2019).
3. Chau BK, Kolling N., Hunt LT, Walton ME & Rushworth MF A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nat Neurosci* 17, 463–470, (2014). [PubMed: 24509428]
4. Cohen D. et al. Bounded rationality in *C. elegans* is explained by circuit-specific normalization in chemosensory pathways. *Nat Commun* 10, 3692, (2019). [PubMed: 31409788]
5. Furl N. Facial-Attractiveness Choices Are Predicted by Divisive Normalization. *Psychol Sci* 27, 1379–1387, (2016). [PubMed: 27565534]
6. Hunt LT, Dolan RJ & Behrens TE Hierarchical competitions subserving multi-attribute choice. *Nat Neurosci* 17, 1613–1622, (2014). [PubMed: 25306549]
7. Khaw MW, Glimcher PW & Louie K. Normalized value coding explains dynamic adaptation in the human valuation process. *Proc Natl Acad Sci U S A* 114, 12696–12701, (2017).
8. Louie K., Grattan LE & Glimcher PW Reward value-based gain control: divisive normalization in parietal cortex. *J Neurosci* 31, 10627–10639, (2011).
9. Louie K., LoFaro T., Webb R. & Glimcher PW Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. *J Neurosci* 34, 16046–16057, (2014).
10. Steverson K., Brandenburger A. & Glimcher P. Choice-theoretic foundations of the divisive normalization model. *J Econ Behav Organ* 164, 148–165, (2019). [PubMed: 32076358]
11. Webb R., Glimcher PW & Louie K. The normalization of consumer valuations: context-dependent preferences from neurobiological constraints. *Manage Sci*, (2020).

12. Yamada H., Louie K., Tymula A. & Glimcher PW Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nat Commun* 9, 162, (2018). [PubMed: 29323110]
13. Zimmermann J., Glimcher PW & Louie K. Multiple timescales of normalized value coding underlie adaptive choice behaviour. *Nat Commun* 9, 3206, (2018). [PubMed: 30097577]
14. Louie K., Khaw MW & Glimcher PW Normalization is a general neural mechanism for context-dependent decision making. *Proc Natl Acad Sci U S A* 110, 6139–6144, (2013). [PubMed: 23530203]
15. Gluth S., Kern N., Kortmann M. & Vitali CL Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nat Hum Behav* 4, 634–645, (2020). [PubMed: 32015490]

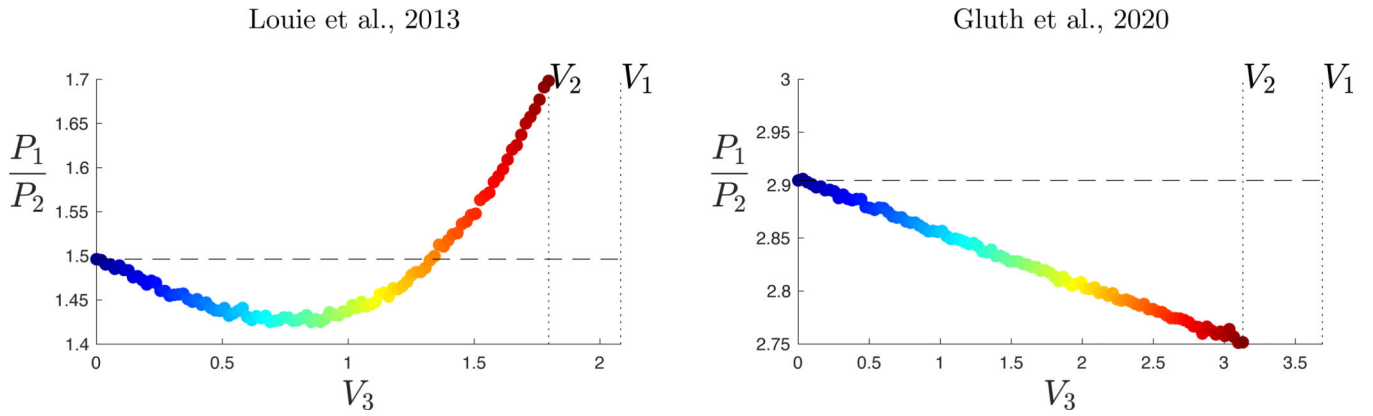


Figure 1. Best-performing model fits of the P_1/P_2 ratio as V_3 is varied. Left, the original Louie et al. dataset; right, the Gluth et al. replication. IIA is represented by the dashed line. These fits are generated for the average V_1 and V_2 observed in each sample. In the Louie et al. dataset, Divisive Normalization with the MVN error captures a u-shaped non-monotonic pattern in the P_1/P_2 ratio. In the Gluth et al. dataset, Divisive Normalization with the Gumbel error captures a decreasing pattern. Overall, both datasets show evidence for normalization but differ in the specific pattern of context-dependence, suggesting subtle differences in the choice data.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

Maximum likelihood estimates for the Gluth et al. replication dataset, pooled over subjects. Standard errors are in parentheses

$f(e)$	No Normalization			Divisive Normalization		
	Gumbel	N	MVN	Gumbel	N	MVN
σ	1.072 (.012)	1.064 (.011)	.994 (.016)	.927 (.051)	.794 (.046)	.783 (.044)
ω	0	0	0	.018 (.006)	.033 (.006)	.027 (.006)
C_1			.547 (.038)			.556 (.039)
C_2			.957 (.019)			.942 (.018)
LL	-16727	-16917	-16898	-16723	-16900	-16886
p-value on $\chi^2(1)$ test for $\omega > 0$.0041	.0000	.0000
AIC	33456	33836	33802	33450	33804	33780