To each their own theory: Exploring the limits of individual differences in decisions under risk

Joshua C. Peterson[1] **(joshuacp@princeton.edu)**
Marina Mancoridis[1] **(marina.mancoridis@gmail.com)**
Thomas L. Griffiths[1,2] **(tomg@princeton.edu)**
[1]Department of Computer Science
[2]Department of Psychology
Princeton University

Abstract

Theories in cognitive science are primarily aimed at explaining human behavior in general, appealing to universal constructs such as perception or attention. When it is considered, modeling of individual differences is typically performed by adapting model parameters. The implicit assumption of this standard approach is that people are relatively similar, employing the same basic cognitive processes in a given problem domain. In this work, we consider a broader evaluation of the way in which people may differ. We evaluate 23 models of risky choice on around 300 individuals, and find that most models—spanning various constructs from heuristic rules and attention to regret and subjective perception—explain the behavior of different subpopulations of individuals. These results may account for part of the difficulty in obtaining a single elegant explanation of behavior in some long-studied domains, and suggest a more serious consideration of individual variability in theory comparisons going forward.

Keywords: decisions under risk; theory development; cognitive modeling; individual differences

Introduction

A primary aspiration of psychology has been to identify general principles of behavior that apply universally across the population. One of the oldest searches for such principles have focused on *decisions under risk*—fundamental kinds of decisions where the outcomes of our choices are uncertain: *e.g.,* would you rather take $1 or have a chance at winning $2 with a 50/50 chance? One early account that is still influential among psychologists and behavioral economists today is Expected Utility Theory (Von Neumann & Morgenstern, 1944), which posits that people maximize the expected (average) utility of their decisions. However, at least 65 competing theories of this behavior have emerged (He, Zhao, & Bhatia, 2022) and reaching consensus as to a winner has remained challenging (Brandstätter, Gigerenzer, & Hertwig, 2008; Erev, Ert, Plonsky, Cohen, & Cohen, 2017).

One possible explanation for this proliferation of models is that the underlying behavior is more complex than we might suspect. Peterson, Bourgin, Agrawal, Reichman, and Griffiths (2021) conducted an analysis of approximately 10,000 choice problems and concluded that the complexity class of the true decision function is more complex than the majority of current theories that have been proposed. They further demonstrated that a more complex model that applies traditionally competing explanations in different problem contexts outpredicts all other models in the largest theory evaluation to date. This need for multiple explanations implies that different theories may to some extent be better understood as representing different possible cognitive strategies that a person might employ when making decisions, an idea that has been suggested by other theorists in the past as well (Brandstätter et al., 2008; Erev et al., 2017). Moreover, the prevalence of more than one decision strategy further opens up the possibility that *individual* decision makers may also employ very different strategies when making their decisions.

Evaluating theories at the level of individuals is certainly not a new idea. In fact, it is common in modeling decisions under risk. However, the typical approach is to modulate free parameters within a single model (representing a single theory) to help explain individual variability in response data. In the case of Expected Utility Theory, a single parameter is modulated to express differences in “risk preference,” where some individuals are thought to be “risk-seeking” (*i.e.,* willing to take a chance to win more) and others “risk-averse” (*i.e.,* may accept less money on average to avoid uncertainty). However, this approach doesn’t allow for the modulation of entirely different constructs used in competing theories (*e.g.,* attention, regret, or disappointment). That is, it is possible that some people make decisions that are strongly affected by say, a process of attention (Birnbaum, 2008), whereas others may be more driven by their inclination to minimize feelings of regret (Bell, 1982).

In the current work, we ask whether entirely different theories might provide explanations of entirely different people. We address this question by conducting the largest analysis of individual differences in decisions under risk to date, evaluating more than 20 theories on each of 300 individuals while taking care to avoid overfitting. We find that almost every theory we evaluate provides the best fit to some subpopulation of individuals. While this implies that many different theories may be necessary to fully explain human decisions, we find that the Mixture of Theories model (Peterson et al., 2021) was the most explanatory. Other influential models such as Prospect Theory (Kahneman & Tversky, 1979) and Regret Theory (Bell, 1982) were also highly explanatory. We discuss how these results support the idea that even fundamental kinds of behavior may vary significantly across individuals and that increasing the diversity of pools of competing theories may end up being a more fruitful strategy in the practice of theory development in some domains.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

Theory Evaluation

All theories we consider are aimed at explaining human choices between two gambles: gamble A and gamble B, where a single gamble is a collection of possible outcomes x_i and their respective probabilities p_i. We further use the notation y_j and q_j to distinguish the outcomes and probabilities respectively of gamble B from gamble A. Most theories specify how people are thought to determine the value of a gamble, such as gamble A, or $V(A)$. The probability that a decision maker chooses gamble A is proportional to $V(A)$, defined as:

$$P(A) = \frac{e^{\eta V(A)}}{e^{\eta V(A)} + e^{\eta V(B)}}, \tag{1}$$

where η is a fitted parameter that controls the degree of deterministic responding.

As discussed above, the number of competing theories of risky choice is large. Our aim in this work is not to exhaustively test all theories but to establish the general applicability of a diversity of theories to the explanation of different individuals. Thus, we select a subset of 22 theories that span many different traditions and literatures. For ease to the reader, these models are described below for the simple case of two-outcome gambles, although we apply each theory to datasets with gambles involving more than two outcomes.

A standard normative model of human decisions is **Expected Value Maximization** (EV), which encodes the assumption that people maximize the average payoff from the outcomes of their choices (*i.e.,* make the most money possible):

$$V(A) = \sum_{i=1}^{2} p_i x_i. \tag{2}$$

We next consider models across the history and literature on "Subjective Expected Utility." These include **Expected Utility Theory** (EU), discussed above (Von Neumann & Morgenstern, 1944), where $u(x_i)$ is typically a single-parameter power function (Wakker, 2010):

$$V(A) = \sum_{i=1}^{2} p_i u(x_i), \tag{3}$$

Prospect Theory (PT), famously proposed by Kahneman and Tversky (1979) to explain deviations from EU, where $\pi(x_i)$ is also a single-parameter power function (Wakker, 2010):

$$V(A) = \sum_{i=1}^{2} \pi(p_i) u(x_i), \tag{4}$$

Cumulative Prospect Theory (CPT), an upgraded version of Prospect Theory (Tversky & Kahneman, 1992):

$$V(A) = \begin{cases} \pi^{+}(p_1)\, u(x_1) + (1-\pi^{+}(p_1))\, u(x_2), & \text{if } 0 \le x_2 \le x_1 \\ \pi^{+}(p_1)\, u(x_1) + \pi^{-}(p_2)\, u(x_2), & \text{if } x_2 < 0 < x_1 \\ (1-\pi^{-}(p_2))\, u(x_1) + \pi^{-}(p_2)\, u(x_2), & \text{if } x_2 \le x_1 \le 0 \end{cases} \tag{5}$$

and the **Mixture of Theories** (MOT) model (Peterson et al., 2021), a context-dependent variation of Prospect Theory where one of two utility $u_j(\cdot)$ and weighting functions $\pi_k(\cdot)$ are soft-selected depending on the choice problem using sets of inferred convex weights ω_j and ω_k:

$$V(A) = \sum_{i \in A} \left[\sum_j \omega_j u_j(x_i)\right] \left[\sum_k \omega_k \pi_k(p_i)\right] \tag{6}$$

Also included from this literature was an attention-based model, the **Transfer of Attention Exchange** (TAX) model (Birnbaum, 2008):

$$V(A) = \frac{\left(p_1{}^{\tau} - \frac{\kappa}{3} p_2^{\tau}\right) u(x_1) + \left(p_2{}^{\tau} + \frac{\kappa}{3} p_1{}^{\tau}\right) u(x_2)}{p_1{}^{\tau} + p_2{}^{\tau}}. \tag{7}$$

In this and all following models, τ, κ, δ, ν, and α are free parameters.

We also consider theories from the literature on "Risk-as-value," including **Portfolio Theory with variance** (Markowitz, 1952):

$$V(A) = \sum_{i=1}^{2} p_i x_i - \kappa p_1 p_2 (x_1 - x_2)^2, \tag{8}$$

Portfolio Theory with standard deviation (Fishburn, 1977):

$$V(A) = \sum_{i=1}^{2} p_i x_i - \kappa \sqrt{p_1 p_2}\, (x_1 - x_2), \tag{9}$$

the **Below-Target model** (Fishburn, 1977), where $\mathrm{I}(\cdot)$ is an indicator function:

$$V(A) = \sum_{i=1}^{2} p_i x_i - \kappa \sum_{i=1}^{2} \mathrm{I}(x_i < 100\delta)\, p_i (100\delta - x_i), \tag{10}$$

the **Below-mean semivariance model** (Fishburn, 1977):

$$V(A) = \sum_{i=1}^{2} p_i x_i - \kappa p_2 \left(\sum_{i=1}^{2} p_i x_i - x_2\right)^2, \tag{11}$$

and the **Below-target semivariance model** (Fishburn, 1977):

$$V(A) = \sum_{i=1}^{2} p_i x_i - \kappa \sum_{i=1}^{2} \mathrm{I}(x_i < 100\delta)\, p_i (100\delta - x_i)^2. \tag{12}$$

We further consider theories from the literature on "Counterfactual" models, including **Regret Theory with expected value evaluation** (Bell, 1982), where $\mathrm{R}(\cdot)$ is a "regret" function:

$$V(A) = \sum_{i=1}^{2} p_i x_i + \sum_{i=1}^{2} \sum_{j=1}^{2} p_i q_j \mathrm{R}(x_i - y_j), \tag{13}$$

Regret Theory with expected utility evaluation (Loomes & Sugden, 1982):

$$V(A) = \sum_{i=1}^{2} p_i u(x_i) + \sum_{i=1}^{2} \sum_{j=1}^{2} p_i q_j \mathrm{R}(u(x_i) - u(y_j)), \tag{14}$$

Disappointment Theory without rescaling (Bell, 1985):

$$V(A)=\sum_{i=1}^{2} p_i x_i + \nu p_1 p_2 (x_1 - x_2), \tag{15}$$

Disappointment Theory with expected value evaluation (Loomes & Sugden, 1986):

$$V(A)=\sum_{i=1}^{2} p_i x_i + \kappa \sum_{j=1}^{2} p_i \operatorname{sign}\left(x_j - \sum_{i=1}^{2} p_i x_i\right)\left|x_j - \sum_{i=1}^{2} p_i x_i\right|^{\alpha}, \tag{16}$$

and **Disappointment Theory with expected utility evaluation** (Loomes & Sugden, 1986):

$$V(A)=\sum_{i=1}^{2} p_i u(x_i) + \kappa \sum_{j=1}^{2} p_j \operatorname{sign}\left(u(x_j) - \sum_{i=1}^{2} p_i u(x_i)\right)\left|u(x_j) - \sum_{i=1}^{2} p_i u(x_i)\right|^{\alpha}. \tag{17}$$

Lastly, we consider models from the literature on heuristics. In this case, rather than defining $V(A)$, we define $S(A)$, a function that outputs 1 if gamble A is meant to be selected and 0 otherwise. These models include the **Better-than-average heuristic** (Thorngate, 1980):

$$S(A)=\sum_{i=1}^{2} \mathrm{I}\left(x_i > \frac{1}{4}(x_1 + x_2 + y_1 + y_2)\right), \tag{18}$$

the **Equiprobable heuristic** (Thorngate, 1980):

$$S(A)=\frac{1}{2}\sum_{i=1}^{2} x_i, \tag{19}$$

the **Low-payoff elimination heuristic** (Thorngate, 1980):

$$S(A)=2\operatorname{sign}(x_1 - y_1) + \operatorname{sign}(x_2 - y_2), \tag{20}$$

the **Low expected payoff** elimination heuristic (Thorngate, 1980):

$$S(A)=2\operatorname{sign}(p_1 x_1 - q_1 y_1) + \operatorname{sign}(p_2 x_2 - q_2 y_2), \tag{21}$$

the **Minimax heuristic** (Thorngate, 1980):

$$S(A)=\operatorname{sign}(x_2 - y_2), \tag{22}$$

the **Maximax heuristic** (Thorngate, 1980):

$$S(A)=\operatorname{sign}(x_1 - y_1), \tag{23}$$

and the most influential of the series, the **Priority Heuristic** (Brandstätter, Gigerenzer, & Hertwig, 2006):

$$S(A)=\operatorname{sign}(x_2 - y_2) I\left(|x_2 - y_2| > \frac{\max(x_1, y_1)}{10}\right) \times 4 + \operatorname{sign}(p_1 - q_1) I(|p_1 - q_1| > 0.1) \times 2 + \operatorname{sign}(x_1 - y_1). \tag{24}$$

Dataset

We make use of one of the replication datasets originally collected by Peterson et al. (2021). It consists of choice data for 300 participants recruited through the Prolific crowdsourcing platform. Each participant completed five trials of a set of sixty choice problems (*i.e.,* pairs of gambles), for a total of 300 decisions. Following Peterson et al. (2021), each decision model outputs the probability that a particular decision maker chooses gamble A (over gamble B). Each set of sixty gambles were chosen at random from a dataset of 1,000 problems. Each choice problem consisted of a choice between two gambles, as described above, where the number of possible outcomes of any one gamble varied between 2 and 9. After each choice was made, feedback was given about which randomly sampled outcome was obtained. After the experiment, participants received a bonus proportional to a randomly sampled outcome of their decisions, which provided an incentive for participants to make the best possible decisions.

Model Fitting

Analysis 1. For our first analysis, we fit each model to each individual participant by using gradient descent to find parameter values that minimize mean-squared error (MSE) between their decisions and the output of the model. In fitting each model to each participant, we searched for the best learning rate in the set [.0001, .0005, .001, .005, .01, .05, .1, .5, 1, 10] using 10-fold cross-validation and reported performance on only the out-of-sample data. Holding out data for evaluation in this way guards against the possibility that individual models have been overfit to individuals and thus would not generalize to unseen data from that participant. This also avoids the question of whether or not to penalize more complex models, since complexity is warranted if it allows for better generalization.

Analysis 2. To minimize the possibility that parameter optimization fails to converge at the level of individual model fitting, we conduct a second analysis where models are first fit to a different but similar dataset of 10,000 choice problems, also from Peterson et al. (2021), before being fine-tuned on individual-level data. We first collapse response proportions in this dataset across participants, such that there are exactly 10,000 probabilities to predict for each of 10,000 possible choice problems. The idea is that a model that has first been fit to aggregate (shared) behavior will require fewer data points at the level of individuals to converge to a reasonable model. We use this dataset for model initialization as opposed to aggregating responses in the main dataset to avoid overlap in participants across the two. We then use the original dataset for fine-tuning to individuals because it has more trials per participant and thus more examples of individual-level behavior. We search learning rates in the set [.0001, .001, .01, .1, 1, 10] for both aggregate and fine-tuned models, resulting in a total of 36 possible settings for each individual.

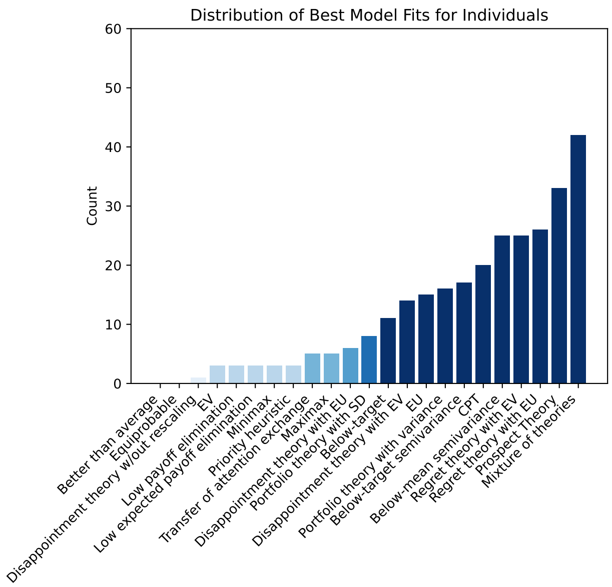

Figure 1: Distribution of the number of times each model provided the best fit for an individual.

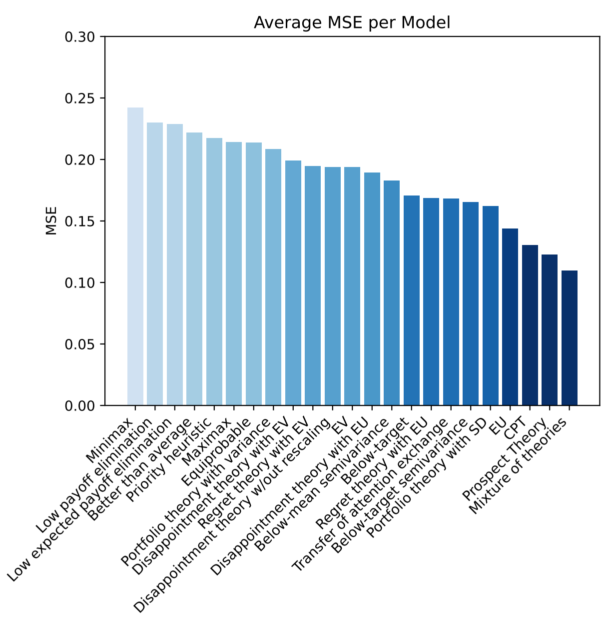

Figure 2: Average MSE across all individuals for each model.

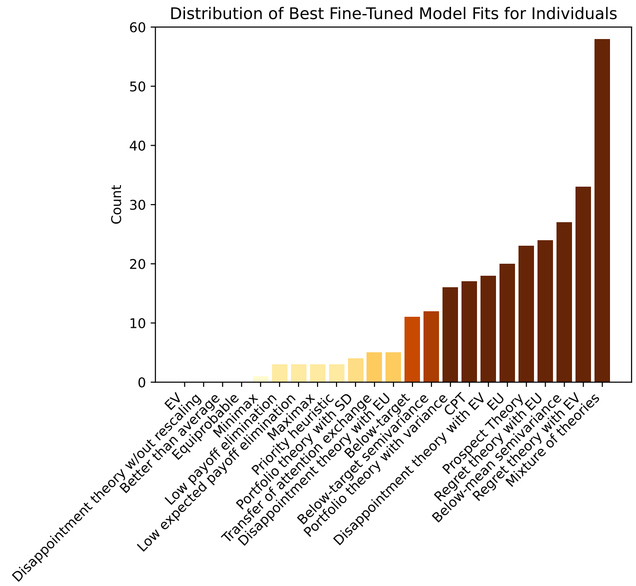

Figure 3: Distribution of the number of times each fine-tuned model provided the best fit for an individual.

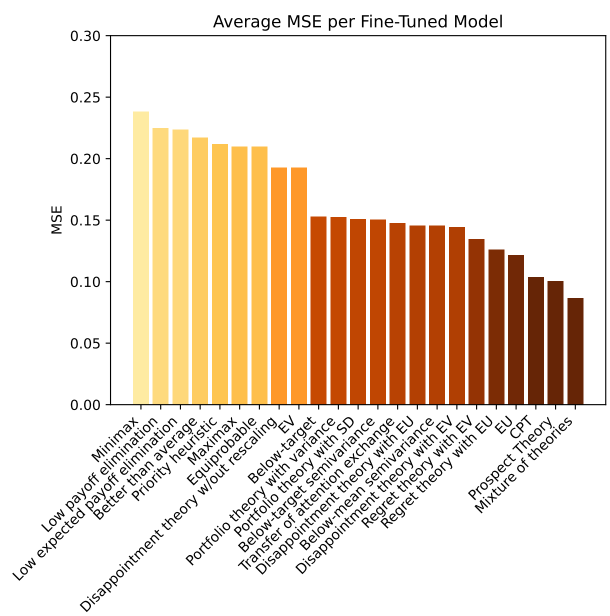

Figure 4: Average MSE across all individuals for each fine-tuned model.

Results

Analysis 1. Figure 1 displays a histogram of best-fit counts for each model across the 300 individuals. Importantly, no single model fits every individual best, meaning that most models provide best explanations of at least some subpopulation of individuals. However, some models provided best fits to a larger swath of individuals than others, in particular MOT, Prospect Theory, and Regret Theory. Heuristic models were the least explanatory. A diversity of other theories involving subjective functions, disappointment, and constructs from the Risk-as-value class all explained some subpopulation best. Figure 2, which by contrast plots average MSE across all participants for each model, paints a slightly different picture of model performance. That is, some models, such as Below-mean semivariance, have a lower performance rank under this metric. This may indicate that models can provide good fits to many individuals despite relatively bad non-best fits.

Analysis 2. Figures 3 and 4 show analogous results when models were first initialized via training on the aggregate version of the data. A quick look at Figure 4 confirms what we might expect, that overall MSE is either lower or approximately equivalent. This analysis thus likely gives a more accurate picture of the best performance and true ranking of each model across individuals. Interestingly, Figure 3 indicates that the distribution of best-fits is now more dominated by particular models, with much of the frequency mass now re-allocated rightward. This may suggest that models such as Prospect Theory require less data to characterize individuals, whereas models such as Regret Theory may require more. Providing enough data is thus essential to increase the precision of model comparison.

One important confound in the above analysis is that each participant completed a different set of choice problems. This leaves open the possibility that particular models better fit particular individuals not because of their individual characteristics but because they made decisions for similar choice problems. After all, past work indicates that different strategies may apply in the context of different choice problems (Peterson et al., 2021). To rule this out, we looked at the top 1,000 pairs of individuals with the greatest overlap in choice problems. No pair shared more than 25 choice problems. The proportion of pairs that were best fit by the same model was 11.5%. If there is no dependence between problems and models, then the probability that two decision makers are fit by the same model is simply the sum of the squared proportion of individuals that are best fit by each theory, which comes out to 11.1%. Thus, the assignment of individuals to the same best fitting model appears to be fully explained by chance.

Discussion

When specifying theories of basic cognitive processes, theorists have an understandable tendency to provide highly general explanations. We say, for example, that our theory of decisions is a theory of mental sampling, or a theory of attention. When considering individual differences, we thus expect that people differ in the manner they draw mental samples, or in the limitations or proclivities of their attention systems, but is this a reasonable assumption?

We set out to test the idea that different theories, even those employing very different constructs, could explain the behavior of different subpopulations of participants. While we found that the distribution of best-fits across individuals is not uniform, there is also no single winner. Most theories we evaluated provided best fits to at least some subpopulation of individuals, ranging from less than 10 to almost 60. This suggests that people may vary more in their decision making strategies and behavior than is typically assumed in the vast literature on decisions under risk.

The diversity in the theories that provided the largest number of best fits is interesting. Regret Theory posits that people act to minimize regret. Risk-as-value theories assume the people penalize gambles with high variability. EU, PT, and CPT all mainly appeal to subjective functions that transform objective quantities such as monetary outcomes and probabilities into subjective quantities. MOT assumes that people employ different decision strategies in different contexts. In retrospect, it seems more than merely plausible these ideas might apply to some people more than others.

Perhaps the most surprising finding aside from the relevance of a diversity of models in general was the high applicability of regret-based theories. These theories assume that a gamble has a lower value if it is likely to induce regret as a result of considering counterfactual scenarios (*i.e.,* where more money could have been earned otherwise). There is far less work evaluating the predictive performance of such models compared to *e.g.,* Prospect Theory, the Transfer of Attention Exchange model, or the Priority Heuristic. While the performance of Prospect Theory may be expected given that it is a special case of the MOT model, Regret Theory is not known to be explicitly related to MOT. This may suggest that multi-strategy models such as MOT may benefit from including a regret component in future iterations. In any case, tendency toward regret may be one of the more defining features that sets some individuals apart from others.

Even in the case that theorists indeed take our results as evidence for stronger individual differences, a number of arguments could still be made for a single winner based on various criteria. MOT, which assumes that the effects of prospect theory are context-dependent, sometimes turning on or off, provided best fits to the largest number of individuals while also maintaining lowest overall MSE. However, Regret Theory and Prospect Theory are relatively simple theories, providing good explanations with relatively minimal machinery (and requiring less data than MOT to be fit). EV and EU, the favorites of early behavioral economists, are even simpler. It could be argued that such simple theories explain the majority of behavior while subsequent ideas explain only minor biases and idiosyncrasies. Nonetheless, based on our results, the range of these idiosyncrasies is likely larger than has been considered in past work.

As we highlighted near the outset of this work, people may employ different cognitive strategies depending on the context or decision problem at hand. Our analysis suggests that decisions under risk may be even more complicated than this, moderated by both context and significant differences in individual characteristics (*i.e.,* different strategies of employing different strategies). This presents nontrivial considerations for future modeling. While it is relatively easy to vary single parameters within any particular model, there is not always a library of historical theories at hand. Thus, theorists should consider leaving room when possible for greater flexibility when considering how a pool of individuals may differ.

Our analysis is not without limitations. First, while we aimed to examine as broad a sampling of decision theories as possible, we likely excluded some important ones. For example, one recent promising theory is BEAST (Erev et al., 2017), which was not considered here simply because it takes much more time to fit—and thus, to evaluate fairly—than other theories we consider. Second, the need for pre-fitting models on aggregate data also suggests that analyses would benefit more from larger data at the level of individual resolution beyond just large numbers of participants. Lastly, and perhaps most importantly, it is possible that some portion of the ranking based on number of best fits may be due to noise. However, if this were the case, we would expect better performing models to be highly similar, such that only chance factors determine the final "best fit" to a particular individual. Our results suggest a different story. For example, Regret Theory, Prospect Theory, and MOT are relatively different models. Thus, perhaps a more likely explanation of the results is that some participants do indeed worry more about potentially regretting their decisions later than others. Nonetheless, future work should focus on reducing the chance of noise through repeated fits to different subsets of participants' trials or larger datasets at the level of trials per participant.

Conclusion

Theories of behavior have long sought to highlight both the common principles of cognition that are shared across the population as well as the characteristics that make individuals so unique and interesting. However, variability at the level of constructs is usually reserved for debates and model comparisons concerning human behavior in general, while variability at the level of individual differences is generally considered to be much smaller in scope. Our results suggest that this may not always be a good assumption, and that theorists should take care to consider that even fundamental cognitive strategies may differ greatly across individuals.

Acknowledgments

This research project and related results were made possible with the support of the NOMIS Foundation and the National Science Foundation (grant no. 1718550).

References

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations research*, *30*(5), 961–981.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations research*, *33*(1), 1–27.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological review*, *115*(2), 463.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. *Psychological review*, *113*(2), 409.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky choice with heuristics: reply to birnbaum (2008), johnson, schulte-mecklenbeck, and willemsen (2008), and rieger and wang (2008).

Erev, I., Ert, E., Plonsky, O., Cohen, D., & Cohen, O. (2017). From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological review*, *124*(4), 369.

Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *The American Economic Review*, *67*(2), 116–126.

He, L., Zhao, W. J., & Bhatia, S. (2022). An ontology of decision models. *Psychological Review*, *129*(1), 49.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*, 278.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, *92*(368), 805–824.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. Retrieved 2023-01-29, from `http://www.jstor.org/stable/2975974`

Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, *372*(6547), 1209–1214.

Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science*, *25*(3), 219–225.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), 297–323.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.