Social Cognition, Vol. 43, No. 5, 2025, pp. 409–424
 https://doi.org/10.1521/soco.2025.43.5.409

Social Utility Modeling: A Tool to Gain Insight Into Social Motives

Elijah P. Galván and Alan G. Sanfey
Donders Institute, Radboud University, Nijmegen, Netherlands
Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

Computational modeling is an emerging analysis technique with the potential to offer important insight into how researchers in the cognitive sciences approach important questions—particularly questions about how people make choices. However, for those researchers who are interested in utilizing computational models in their own research, learning the how's and why's of the approach can seem prohibitively difficult. In the current work, we address these concerns by firstoutlining the basic principles of computational modeling in plain and accessible language. We then propose criteria to consider when adopting a computational model to answer a research question, illustrating that Social Utility Models have many applications in the field of social cognition and enjoy useful advantages over conventional analysis approaches. Finally, in a step-by-step tutorial, we explain how to implement a computational modeling analysis and demonstrate this approach by using an example data set.

Keywords: computational modeling, individual differences, decision-making

COMPUTATIONAL MODELS OF DECISION-MAKING

Recently, the study of decision-making has seen an increasing emphasis on the use of computational modeling to explain and understand choice behavior. When applied to the study of decision-making, computational models aim to represent the process of choice arbitration; that is, how people decide between different options (Konovalov et al., 2018; Vahed et al., 2024). Specifically, computational models of decision-making aim to represent *how* choice options are evaluated, with the effect that they make specific predictions about how people decide. Herein, computational models of decision-making differ from linear models of decision-making in the sense that they do not relate independent variables to a dependent variable. Instead, they relate choices, or dependent variables, to a single decision

Address correspondence to Elijah P. Galván, Radboud University, 29 Kapittelweg, Nijmegen, GE 6525EN, Netherlands. E-mail: epgalvan20@g.ucla.edu

variable, which is then utilized as the criteria upon which an option is chosen, as Figure 1B illustrates.

Decision-making is the process of choosing between two or more available options. In the current work, we will focus on the specific process of value-based decision-making, which involves choosing between options based on personal preferences. Here, the goal of computational models is to capture how people choose between different available options; that is, why one choice is subjectively better than others. Studying preferences in decision-making paradigms can allow researchers to investigate questions that cannot be reliably answered by self-report, especially questions pertaining to social cognition. In general, there is a broad range of questions that computational models of decision-making can be used to answer, as Table 1 illustrates.

A key advantage of utilizing computational models to study decision-making is that these models are generative. Generative models are psychological models that generate predictions by extrapolating from a set of assumptions with formalized logic. There are two classes of generative models. One is as-if models, such as Cumulative Prospect Theory, which fit parameters on the decisions themselves, which precludes inferences about the underlying psychological processes that are being estimated per se. In contrast, the other class constitutes psychological models, which estimate parameters by model-based derivatives of the behavior, taking the shape of a latent decision variable. Thus, generative models output predictions based on formalized hypotheses about the underlying data generation process but can differ in the nature of the inferences they support. Descriptive models can also output predictions based on hypotheses (i.e., we can simulate data that reflect a hypothesis that X affects Y), although these hypotheses are not formalized and, consequently, do not represent the underlying data generation process. Because generative models make direct claims about the data generation process, they bridge the gap between behavior and the underlying psychological processes at play. As generative models of human psychology, computational models of decision-making offer great potential for increasing our understanding of a myriad social cognitive processes.

Despite the considerable upside that computational modeling offers for the field of social cognition, these models have not yet been widely adopted by researchers in this field. This is likely due to perceptions that there are prohibitively large barriers of entry into computational modeling. In the current article, we aim to address these perceived barriers by illustrating how a computational model can be implemented and the advantages this approach offers researchers. Specifically, we will focus on the case of Social Utility Models. Social Utility Models offer the possibility of identifying the psychological values underlying an individual's choices, making them a powerful tool for increasing our understanding of social cognitive behavior.

In general, Utility Models are formal (i.e., algorithmic) models of choice evaluation which assume that people choose the option that maximizes the utility, or subjective satisfaction, that they get from the potential choice outcomes. Social Utility Models specifically hypothesize that this utility can derived from either adhering to the norm of self-interest or, alternatively, adhering to one or more norms that

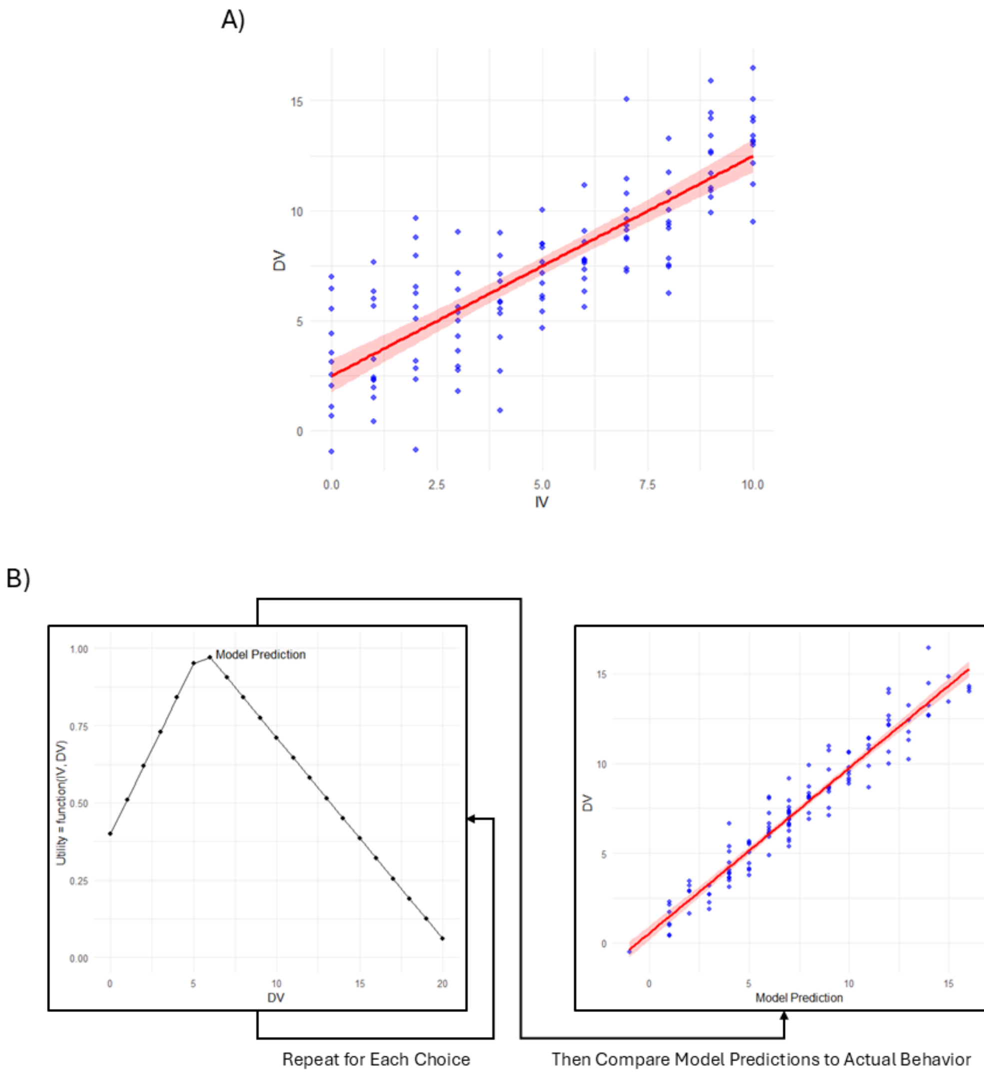

FIGURE 1. Computational modeling's differing approach in decision-making. Linear models are models relating a behavior to a measured or manipulated variable (Figure 1A), while computational models of decision-making are models of how people evaluate choices against each other (Figure 1B).

encompass more prosocial motives. The extent to which a choice (i.e., a possible value of the dependent variable) follows a norm is determined by the context in which the decision takes place (i.e., the values of the relevant independent variables and constants). Social Utility Models typically represent the hypothesis that individuals vary in terms of the utility that they derive from a given norm; these differences in preferences are represented by free parameters. Thus, one fundamental distinction between Social Utility Models and standard Linear Models is

TABLE 1. Computational Models of Decision-Making

Model Type	Use Case	Foundational Model(s)	Examples in Social Cognition Research
Reinforcement Learning	Repeated interactions; second-order risky choice	Rescorla-Wagner Model (Wagner & Rescorla, 1972)	Chang et al. (2010)
Foraging Models	Sequential choices; risky choice	Marginal Value Theorem (Charnov, 1976) Optimal Foraging Theory (Emlen, 1966; MacArthur & Pianka, 1966)	Turrin et al. (2017)
Risky Choice Models	First-order risky choice	Cumulative Prospect Theory (Tversky & Kahneman, 1992)	Nguyen et al. (2016)
Intertemporal Choice	Deterministic choice; outcomes temporally separated	Hyperbolic discounting (Breeden, 1979; Koopmans, 1960)	Levin (2014)
Motivation Models	Deterministic choice; at least one outcome requires effort	Demand Model (Hursh, 1993)	Lockwood et al. (2017)
Process Models	Deterministic choice; process data available (i.e., eye-or-mouse-tracking)	Drift-Diffusion Model (Ratcliff, 1978)	Zhang et al. (2024)
Social Utility Models	Deterministic choice; outcomes affect other people	Inequality-Aversion (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999)	van Baar et al. (2019)

Note. First-order risk refers to gambles undertaken with known outcomes and known probabilities. Second-order risk refers to gambles undertaken with known outcomes but unknown probabilities. Deterministic choice refers to decisions undertaken with known and certain outcomes.

how they handle individual differences: A priori, Social Utility Models commit to a precise definition about the exact relationships between independent variables and dependent variables, but they do not hypothesize about which people will have which relationships. By contrast, when Linear Models are typically used to study individual differences, researchers hypothesize about which people will differ in their relationship between independent and dependent variables, but do not commit to a specific relationship between independent and dependent variables.

ACTIVE TUTORIAL

In this section, we will present a high-level tutorial of Social Utility Modeling techniques in order to provide a practical example of this process. Many of the difficulties that researchers encounter when learning and utilizing this approach come from a lack of practical guidance about how to implement this programmatically (Wilson & Collins, 2019). Therefore, in order to supplement this work, we have developed an online, freely accessible *Handbook of Social Utility Modeling* (https://social-utility-modeling.readthedocs.io/en/latest/). This handbook contains a step-by-step guide on how to implement this modeling approach, as well as four tutorials that users can practice using actual data from real experiments (Crockett et al., 2014; Galván & Sanfey, 2024; Li et al., 2022; van Baar et al., 2019). These tutorials are available in Python, MatLab, and R to allow for implementation in whichever language users are most familiar with, and each tutorial is completed and can be viewed on the Tutorial page (https://epgalvan.github.io/social-utility-modeling/). The present article was written to develop conceptual understanding of what to do and why to do it; in concert, the online handbook should develop competence in independently implementing this approach in order to be applied to your own research questions.

The first step in Social Utility Modeling is to develop a research question about how people make value-based decisions that impact other people. Here, we will follow the example of van Baar et al.'s (2019) article, which sought to answer the following research question: "What motivates people to reciprocate trust, even when there are no external incentives to do so?" The research question meets this demand because it asks about a specific value-based decision (i.e., to reciprocate trust or not), which impacts both the decision-maker and other people.

To answer this question, we need to identify some plausible answers. First, it is possible that people do not actually reciprocate at all, preferring to keep the money for themselves—we can term this strategy "Greed." Second, it is possible that people do reciprocate, and do so because they dislike when they have more money than their partner, as they likely would if keeping all the money—this strategy can be called "Inequality-Aversion." Third, it is possible that people reciprocate because they would feel guilty for disappointing those who have trusted them. Trust indicates an expectation that one's vulnerability will not be exploited, so people may simply seek to avoid violating others' expectations—we can call this strategy "Guilt-Aversion."

Next, a task must be developed where these different explanations make differing predictions. To examine questions about reciprocity, it is possible to employ a well-studied experimental task known as the Trust Game (Berg et al., 1995), which is depicted in Figure 2A. In this game, one player (the Investor) is endowed with some money and told that they can invest some of it in a second player (the Trustee). Any amount they invest is multiplied (for instance by a factor of 4) and transferred to the Trustee, while anything they do not invest is kept by the Investor as a guaranteed payout. After receiving their money, the Trustee then has the opportunity to return money to the Investor, but, importantly, need not do so. For the Investor, the decision to send money is based on the trust they have that their investment will be repaid by the Trustee, thus making their risky investment pay off. However, for the Investor, they must simply decide how much of their windfall they want to share with the Investor. Do they reciprocate the Investor's transfer, or do they keep the money for themselves? The amount that they return to the Investor is how reciprocity is operationalized in the Trust Game. Here, Greed varies from Inequality-Aversion and Guilt-Aversion; however, the latter two make identical predictions because Investors expect Trustees to reciprocate equally.

To create a task where these different psychological preferences produce distinct patterns of behavior, we can amend the Trust Game such that participants are forced to make a choice about being either Inequality-Averse or Guilt-Averse. To accomplish this, van Baar and colleagues (2019) made the investment multiplier either 2, 4, or 6. However, they withheld this information from Investors, telling them that the multiplier was always 4. This Hidden Multiplier Trust Game (HMTG) is shown in Figure 2B. Because Investors have a false belief about how much the Trustee has when the multiplier is either 2 or 6, Inequality-Aversion and Guilt-Aversion make differing predictions. As Figure 3 demonstrates, for each of Greed, Inequality-Aversion, and Guilt-Aversion, there is clearly a unique relationship between the investment and the amount that people are expected to return.

The next step is to create the Social Utility Model. As mentioned previously, Social Utility Models represent hypotheses about what people value. Using the example hypothesis that people only value money, it is quite straightforward to create a utility equation that embodies the Greed hypothesis. Namely, utility can be expressed as a linear function of the amount of money received from one's payout. For reasons that we will discuss later, we improve this function by normalizing to be a number between 0 and 1, where 1 means that you have maximized your Payout (Equation 1).

$$\text{Equation 1: Utility}_{\text{Choice}} = \text{Greed}_{\text{Choice}} = \frac{My\ Payout_{\text{Choice}}}{\max\ (My\ Payout)}$$

The most relevant principle for modeling social utility is decreasing sensitivity to loss in utility. The loss in utility becomes proportionally smaller as this violation becomes larger: If Choice A violates the norm twice as bad as Choice B, the loss in utility is less than twice as bad for Choice A compared to Choice B.

There are four steps in quantifying the utility of following social norms:

1. Define what it means to follow the norm we want to represent in the model and express this mathematically, such that complete adherence to the norm results in the norm being equal to 0.
2. Normalize this term to be between 0 and 1 by dividing the expression established in Step 1 by its maximum possible absolute value.
3. Invert this ratio by subtracting it from 1: Following the norm now results in higher utility, meaning that the model predictions align with the hypothesis that people will choose options that follow our social norm.
4. Apply transformation(s) to our ratio that are consistent with psychological principles.

In what follows, we apply Steps 1–4 to explain behavior in van Baar et al.'s (2019) experiment. By following these instructions, we arrive at a different formulation of the model compared to the one which appears in the article. Importantly, however, the models fit the data equally well and are functionally equivalent.

Representing the utility of following social norms, rather than self-interest, is a slightly more complicated process. To this end, the four steps above outline how this can be done. We will apply this process to Guilt-Aversion first, which prescribes giving one's partners exactly what they expect. As Figure 2 illustrates,

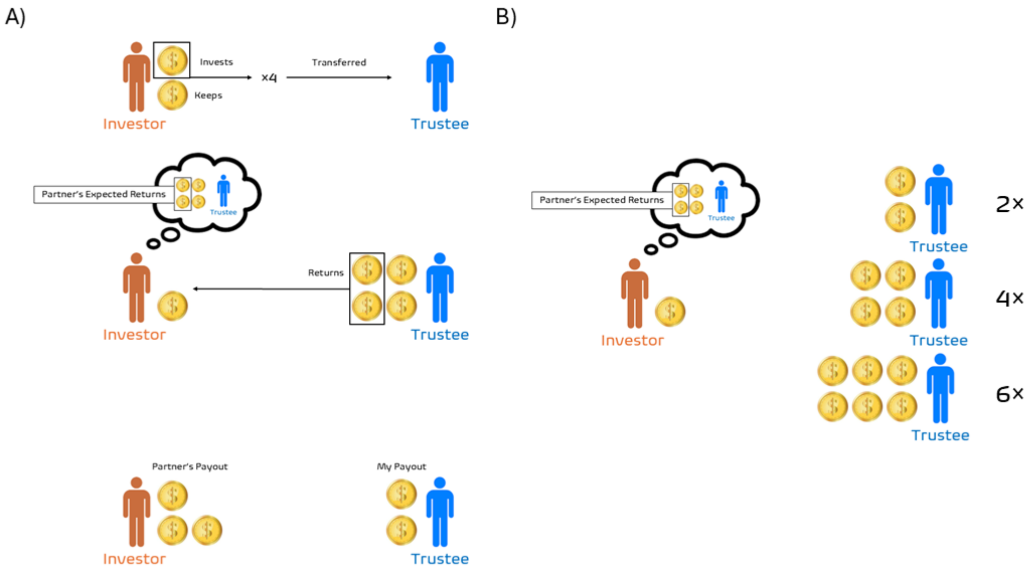

FIGURE 2. Trust Game. Figure 2A shows the canonical Trust Game first proposed by Berg, Dickhaut, & McCabe (1995). Figure 2B shows the amendments made by van Baar, Chang, & Sanfey (2019) to make their Hidden Multiplier Trust Game.

Investors expect Trustees to return half of what they believe the multiplied investment is, and maximally violating this norm involves returning nothing. To quantify the violation of this norm, we can subtract this expectation term from what the Investor actually receives. We then want to normalize this between 0 and 1 by dividing the maximum possible absolute difference between the expectation and what the Investor could receive. So this term will be 0 when investors' expectations are met and 1 when these expectations are most strongly violated. Ensuring that all terms in the model are normalized between 0 and 1 is ideal because it often eliminates the need to recursively search for ideal boundaries for free parameters (a process described by Wilson & Collins, 2019), as these free parameters should, consequently, range from 0 to 1 as well. We need these to be inverted, so we just subtract 1 from the term we have, meaning that a value of 1 means you have fully met expectations and a value of 0 represents you fully violating expectations. Finally, to incorporate diminishing sensitivity to violating a partner's expectations, the Guilt-Aversion term is squared, which is in line with previous models of Guilt-Aversion (Battigalli & Dufwenberg, 2007; Dufwenberg & Gneezy, 2000). The model of Guilt-Aversion shown in Equation 2 is the final outcome of this process.

$$\text{Equation 2: } Utility_{Choice} = Guilt_{Choice} = 1 - \left(\frac{Partner\ Payout_{Expectation} - Partner\ Payout_{Choice}}{\max\left(Partner\ Payouts_{Expectation} - 0, 0\right)}\right)^2$$

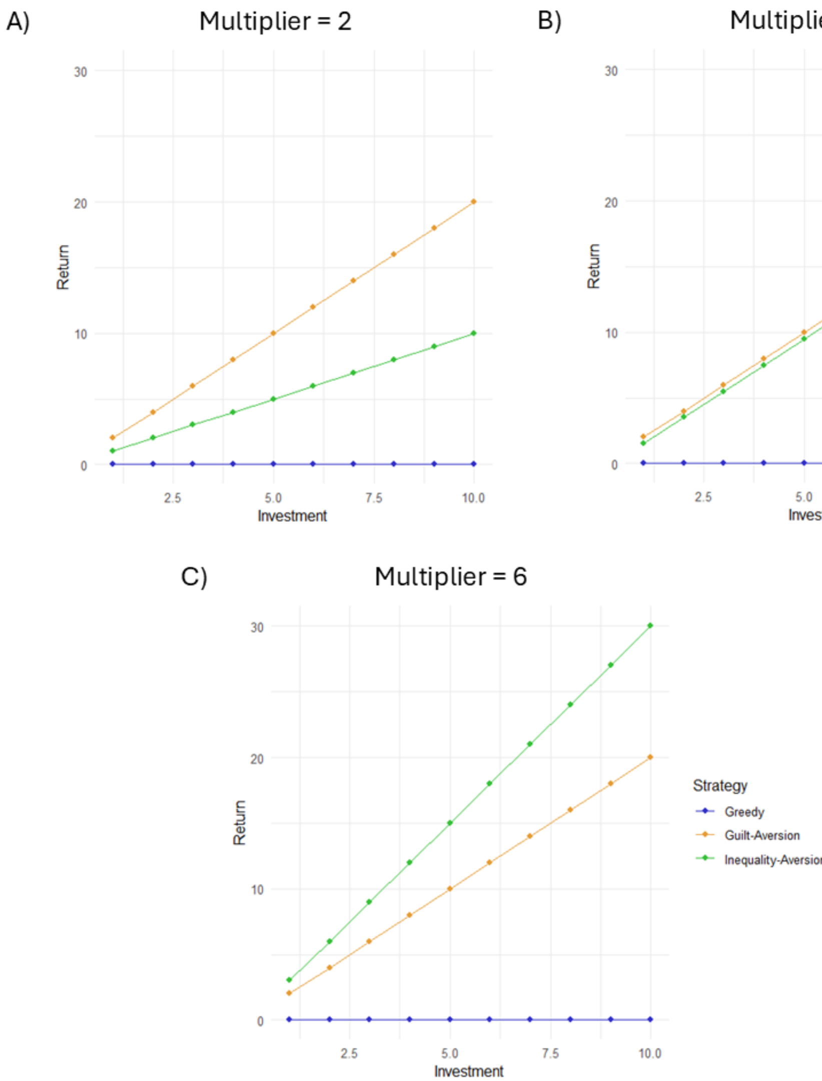

FIGURE 3. Reciprocation amounts prescribed in the HMTG. In the HMTG, Guilt-Aversion and Inequality-Aversion make differing predictions when the actual multiplier is different from the believed multiplier, which is always 4.

The four steps in Box 1 can be repeated again for Inequality-Aversion. Perfect Inequality-Aversion is achieved when the payouts of participants and their partner are equal. The maximum possible violation is when the absolute difference between the payouts of both you and your partner are the greatest. We then invert the term by subtracting it from 1. Finally, we apply the square transformation to capture the decreasing loss in sensitivity as proposed by previous formulations of Inequality-Aversion (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999), which leads to Equation 3.

$$\text{Equation 3: Utility}_{\text{Choice}} = \text{Inequality}_{\text{Choice}} = 1 - \left(\frac{My\ Payout_{Equality} - My\ Payout_{Choice}}{\max\left(|My\ Payouts - My\ Payout_{Equality}|\right)} \right)^2$$

There are now three different Utility Equations that each represent a distinct hypothesis, namely, that people will follow the same norm when making reciprocation decisions. However, these hypotheses are not necessarily wholly exclusive across multiple trials of the task: People can potentially choose to follow different norms under different circumstances when making reciprocation decisions. We can thus usefully create a single model that represents the hypothesis that people may use all three norms to guide their reciprocity decisions. This can be accomplished by following three fundamental steps, listed below, for specifying a multinorm utility equation:

1. Generate functions that quantitatively represent how well the different choices (i.e., the values of the dependent variable) adhere to social norms or self-interest, given the factors (i.e., the independent variables) in your experiment.
2. Describe dimensions, or free parameters, where these norms differ from each other; in other words, answer the question: "What psychological preferences would lead to a person choosing to follow each norm?"
3. Pair the free parameters with the relevant norm terms by multiplying them together, then add these interactions together.

The first step is already completed: The functions from Equations 1–3 can be taken to represent Greed, Guilt-Aversion, and Inequality-Aversion. The output of each function is plotted in Figure 4, demonstrating that the norms prescribe different reciprocation behavior on different trials of the HMTG.

For Step 2, we then want to identify how these functions are conceptually different: Greed is different from Guilt-Aversion and Inequality-Aversion in the sense that Greed follows from selfishness while the other two prioritize doing the "right" thing over being selfish: Greed is differentiated from the other two norms on Dimension 1. Guilt-Aversion and Inequality-Aversion are differentiated by how they define what is "right," so this can be considered as Dimension 2. Greek letters are used to represent free parameters in these models, so Dimension 1 will be free parameter Θ in the model and Dimension 2 will be Φ.

Moving to Step 3, free parameters must be assigned to each norm and mathematically represented in the utility equation. Θ is assigned to Greed, while the inverse of Θ is assigned to both Inequality-Aversion and Guilt-Aversion. Φ is assigned to Inequality-Aversion, and the inverse of Φ is assigned to Guilt-Aversion. If we multiply the free parameters with the functions they are paired to and then add all of these terms together, we arrive at our Social Utility Model as shown in Equation 4. Including random effects in a linear mixed model can enable us to capture behavior caused by Greed, Inequality-Aversion, and Guilt-Aversion. The crucial distinction here is that descriptive models represent the hypothesis that people differ in the relationship between the independent

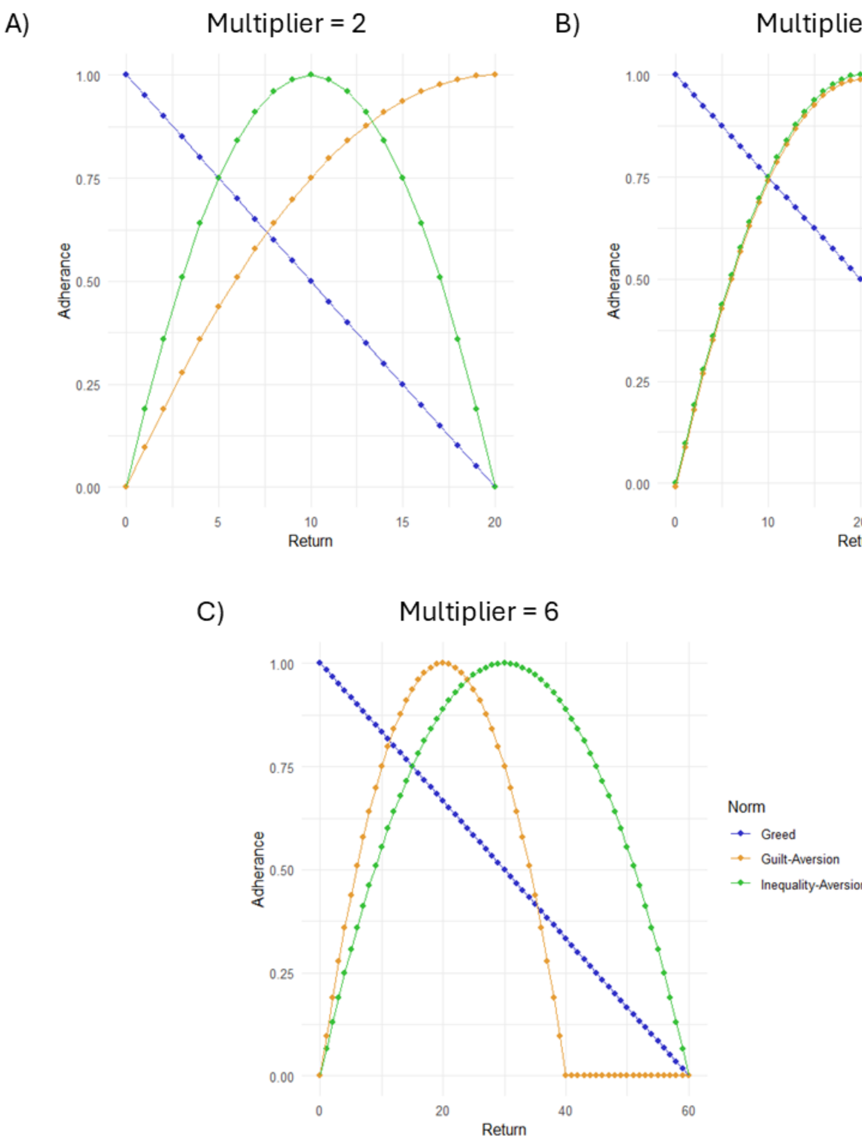

FIGURE 4. Quantification of Greed, Guilt-Aversion, and Inequality-Aversion. Adherence to greed is always highest when returning the lowest possible amount. Adherence to Guilt-Aversion is always highest when returning 20 tokens. Adherence to Inequality-Aversion is always highest when returning half of the multiplied investment.

variables and the dependent variable without committing to a constrained set of predictions (a) about what psychological processes drive these decisions and, therefore, (b) what these relationships imply about the criteria considered in the decision-making process.

Equation 4: $Utility_{Choice} = \Theta \times Greed_{Choice} + (1 - \Theta) \times (\Phi \times Inequality_{Choice} + (1 - \Phi) \times Guilt_{Choice})$

This approach notably deviates from what has informed many previous implementations of Social Utility Models. In these previous implementations, a model is specified *per condition*, allowing researchers to study how a certain set of preferences is modulated in certain contextual factors (Crockett et al., 2014; Lockwood et al., 2017; Stallen et al., 2018). While this approach is certainly the most straightforward and the handbook provides guidance on implementing this approach (https://social-utility-modeling.readthedocs.io/en/latest/3_2_0.html), we argue for a different approach. This approach sees these modulatory effects integrated in the model, which we believe is preferable for three primary reasons: The first is empirical while the second and third are theoretical.

1. Simulations demonstrate that, under certain circumstances, a per condition model may be incorrectly identified as the correct model over an integrated model (https://epgalvan.github.io/integration-simulations/).
2. The per condition approach cannot generalize to new contexts, while the integrated model can; van Baar and colleagues (2020) demonstrate this in a follow-up study.
3. The integrated model represents hypotheses about *how* and *why* preferences change between conditions, while the per condition model simply captures these changes *if* preferences change between conditions.

Following the creation of the utility equation, it is necessary to create a set of trials; here, the important information is the number of trials in the experiment and what those trials will look like to participants. With both a utility equation and a trial set, it is possible to simulate data. Simulating data is the process of generating model predictions for a hypothetical set of free parameters. A useful way of viewing free parameters is that they are the dimensions on which people can meaningfully differ from each other in their tasks. These dimensions alone represent all of the hypothesized ways that people can behave differently in their experiments. Therefore, the process of simulating data answers the question, "What would a person with this specific set of preferences do in my task?"

After model predictions are generated, one crucial step is still needed prior to model implementation. This extremely important step is to ensure that free parameters can be accurately estimated from the data that created them. Parameter estimation is the process of determining the "best" set of free parameters for the data. For Social Utility Models, the "best" set of parameters is one that minimizes the difference between the expected utility and the observed utility. Expected utility refers to the utility of choosing the model's predictions, whereas observed utility refers to the utility of the observed choice itself. Thus, the free parameter values that minimize this difference, the "best" set of parameters, is the one that would have generated the data if the decision-maker maximized utility according to the utility equation, as the model fundamentally assumes. Parameter recovery is the process of estimating free parameters from simulated data: By demonstrating that free parameters can be reliably estimated from the data that created them, researchers can ensure (a) that their model makes distinct predictions across values of free parameters (i.e., that these parameters work as intended) and (b) that the optimization algorithm can reliably find true free parameter values. If parameters cannot be reliably recovered, adjustments to the model, trial set, or free parameter values are often necessary. For a list of several common causes for poor parameter recovery and how to fix them, see the "Fixing Nonspecific Models" dropdown on the "Recovering Free Parameters" page, which can be reached by clicking on this text on the left-side panel (https://social-utility-modeling.readthedocs.io/en/latest/1_6_0.html).

Parameter estimation can be accomplished by using one of two general kinds of optimization algorithms: global or local. Optimization algorithms take an objective function and data as an input and functionally try out many different sets of free parameters to identify the best set. This best set of free parameters is the set that either minimizes or maximizes the output of the objective function, depending on the user's input: In Social Utility Models, the best set of free parameters minimizes the output of the objective function, which is the difference between expected and observed utility. Global algorithms sample sets of free parameters across the entire parameter space, while local algorithms require an initial guess and sample the parameter space starting from this point. To read more about objective functions and optimizers, see the "Recovering Free Parameters" page (https://social-utility-modeling.readthedocs.io/en/latest/1_6_0.html).

In addition, rules can be applied for assigning people to a categorical strategy group based on their free parameters. Instructions about how to do these steps and what to consider are available in the handbook on the "Grouping" page (https://social-utility-modeling.readthedocs.io/en/latest/1_7_0.html).

After it is demonstrated that parameters can be recovered reliably, data collected from participants can be analyzed. The first step in this process is to estimate free parameters for each participant. This implementation, where free parameters are fit to individual participants but not to the population, is akin to running multiple regressions per participant. Hierarchical structuring mitigates the risk of overfitting these free parameters (Singmann & Kellen, 2019), so it is necessary to demonstrate that the estimated free parameters can reliably predict data that the researchers have not yet seen—a practice known as out-of-sample prediction. More information about parameter estimation is available in the handbook on the "Estimating Free Parameters" page (https://social-utility-modeling.readthedocs.io/en/latest/2_1_0.html).

The most conventional approach to out-of-sample prediction is a process known as k-fold cross-validation. Herein, the entire data set is split into k number of folds: For each of these k folds, the data in the fold are withheld and the model is trained on the other folds. Then the free parameters estimated from the other folds are tested on the current fold, which was excluded from the training. In this tutorial, we cross-validated more than five fold per participant, which is often referred to as *fivefold validation* for shorthand. A guide on how to perform k-fold cross-validation is available in the handbook under the "Validating Parameter Recovery Process" dropdown on the "Validate the Best Model" page (https://social-utility-modeling.readthedocs.io/en/latest/2_4_0.html).

With valid and reliable free parameter estimates, it is possible to test hypotheses. The central approach to hypothesis testing is termed *model comparison* wherein we compare models in terms of how well they explain the decisions people make. Social Utility Models are formalized hypotheses drawn from theories about the underlying psychological processes involved in making decisions. Consequently, to test two models that are equivalent apart from one hypothesis is to put that hypothesized psychological process to the test in the behavioral data: If it leads to a significant improvement in the ability to explain the choices that people have made, the inference that follows is that said psychological process is involved in decision-making.

We test models using metrics called Model Fit Indices (MFIs), which quantify model quality by accounting for both performance and parsimony. *Performance* is defined as the model prediction error for the data (i.e., the sum-of-squares between the model's predictions and participants' decisions), and *parsimony* is simply the number of free parameters employed. Because smaller prediction error and fewer free parameters are more ideal, lower MFI values indicate a better model. MFIs can be utilized to statistically test hypotheses because they can be computed per participant per model. The rationale for conducting statistical tests on MFIs is that, in doing so, we are comparing each instance of the data generation process (i.e., a participant) in terms of how well a certain hypothesis (i.e., the single hypothesis that Model A and Model B do not share) can explain the observed data.

Revisiting the research question and hypotheses specified earlier, we want to know what motivates reciprocity and we hypothesize that there are potentially three motives. To test this hypothesis, we compare the MFIs for the model specified in Equation 4 to simpler derivative models: In other words, we must compare the three-norm model to all possible two-norm models and all possible one-norm models. In the Tutorial pages (https://epgalvan.github.io/social-utility-modeling/), we do this comparison and replicate van Baar et al.'s (2019) findings that the three-norm model has a significantly higher MFI compared to the derivative models. In other words, all three reciprocity motives are important for reciprocation decisions.

By utilizing Social Utility Models, van Baar and colleagues (2019) were able to directly answer their research question; namely, by testing the three-norm-model against simplified derivatives of the model, the authors demonstrate that Greed, Inequality-Aversion, and Guilt-Aversion each motivate reciprocation decisions. Only a Social Utility Modeling approach would have facilitated such a direct testing of this hypothesis. While alternative approaches could have enabled the researchers to identify and explain the heterogeneity in choices, Social Utility

Modeling allows for unique conclusions to be drawn about the psychological processes underlying the motives in social decisions.

ADVANTAGES AND LIMITATIONS

We see three main advantages of using Social Utility Models to study social cognitive processes. The first of these advantages arises as a function of how individual differences are handled. In Social Utility Models, models are fit per participant, therefore the estimated free parameters reflect what individual people value: That is, individual differences are treated as an important feature of the model. Because Social Utility Models are generative psychological models, the free parameters that are estimated can be taken to reflect the psychological distinctions that were hypothesized a priori. More to the point, these free parameters can function as individual-level variables that can be used in several different ways.

Second, Social Utility Models enable us to detect the role of multiple different competing processes in a single behavior. As Figure 3 illustrates, Social Utility Models incorporate multiple norms and thereby simultaneously predict multiple relationships between the independent variable and dependent variable of interest. Accommodating multiple relationships is a strength of this approach because different people often use the same information in different ways—a trend that is apparent in the example of van Baar et al. (2019).

The third advantage of using Social Utility Models is that it allows for direct hypothesis testing. Computational models embody hypotheses: They are specific, committed, formal interpretations about what people will do in a certain situation. Therefore, when we want to test a hypothesis via a computational modeling approach, we compare models that differ only in the specific hypothesis we want to test. Thus, the outcome of hypothesis testing is a direct psychological conclusion about what the data say. In the example presented in the active tutorial, the outcome of hypothesis testing was a direct answer to the research question—which would not have been possible with a different approach.

While there are many constraints that limit the domain of Social Utility Models' applications, which are discussed throughout the handbook, there is one key limitation. This limitation is, namely, that it is exceptionally difficult to model the decision-making process in tasks where participants make decisions that are based on unforeseen and unpredictable criteria. Of course, adding more terms to a model is always possible post hoc, but if every person uses different information, adding a unique term for each participant is not possible. In such situations, Social Utility Models have very limited usefulness: They are less interpretable and less useful, if they are even useful at all.

CONCLUDING REMARKS

When applied to the study of social cognitive processes, Social Utility Modeling offers the potential to evaluate psychological factors that can powerfully shape differences in social behavior. Even though the use of formal mathematics can be daunting, the requisite knowledge to begin learning Social Utility Modeling is actually quite minimal, and learning this technique is straightforward. Computational Modeling is a user-driven process, which opens up the opportunity to directly answer questions about what people value, under what circumstances these values change, and if people decide broadly similarly or if there are key differences across individuals. Our contention is that Social Utility Modeling is well suited to answering these questions and offers advantages over alternative possibilities.

One particular advantage is that it allows us to do model testing in an integrated, multimodal way. Specifically, although we validate the model using behavioral data, we can validate the model by showing that individual differences in our laboratory experiment explain individual differences in other kinds of data. For instance, it is common practice to test these models in the brain by using free parameters to predict different patterns of neural activation. It is also possible to demonstrate the ecological validity of your task and your experimental findings by using free parameters to predict the real-world behavior you want to learn about in the first place (Galván & Sanfey, 2024).

Computational models offer tremendous potential for research in social cognition, although this approach has not, as yet, been widely adopted. Undeniably, several key factors responsible for this trend are the perceived steep learning curve and a lack of clarity about what these models aim to do. For the benefit of the field of social cognition, addressing these issues and lowering the barrier to entry for computational modeling should remain a priority. The current work is one such effort, highlighting one specific kind of computational model that social cognition researchers may find particularly informative.

REFERENCES

Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *American Economic Review, 97*(2), 170–176. https://doi.org/10.1257/aer.97.2.170

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122–142. https://doi.org/10.1006/game.1995.1027

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *91*(1), 166–193. https://doi.org/10.1257/aer.90.1.166

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, *7*(3), 265–296. https://doi.org/10.1016/0304-405X(79)90016-3

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. https://doi.org/10.1016/j.cogpsych.2010.03.001

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320–17325. https://doi.org/10.1073/pnas.1408988111

Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior, 30*(2), 163–182. https://doi.org/10.1006/game.1999.0715

Emlen, J. M. (1966). The role of time and energy in food preference. *American*

Naturalist, *100*(916), 611–617. https://doi.org/10.1086/282455

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Galván, E. P., & Sanfey, A. G. (2024). Value-based preferences predict views on redistribution. *PsyArXiv*. https://doi.org/10.31234/osf.io/6d4km

Hursh, S. R. (1993). Behavioral economics of drug self-administration: An introduction. *Drug and Alcohol Dependence*, *33*(2), 165–172. https://doi.org/10.1016/0376-8716(93)90058-x

Konovalov, A., Hu, J., & Ruff, C. C. (2018). Neurocomputational approaches to social behavior. *Current Opinion in Psychology*, *24*, 41–47. https://doi.org/10.1016/j.copsyc.2018.04.009

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society*, *28*(2), 287–309.

Levin, S. A. (2014). Public goods in relation to competition, cooperation, and spite. *Proceedings of the National Academy of Sciences*, *111*(supplement_3), 10838–10845. https://doi.org/10.1073/pnas.1400830111

Li, Y., Hu, J., Ruff, C. C., & Zhou, X. (2022). Neurocomputational evidence that conflicting prosocial motives guide distributive justice. *Proceedings of the National Academy of Sciences*, *119*(49), e2209078119. https://doi.org/10.1073/pnas.2209078119

Lockwood, P. L., Hamonet, M., Zhang, S. H., Ratnavel, A., Salmony, F. U., Husain, M., & Apps, M. A. (2017). Prosocial apathy for helping others when effort is required. *Nature Human Behaviour*, *1*(7), 0131. https://doi.org/10.1038/s41562-017-0131

MacArthur, R. H., & Pianka, E. R. (1966). On optimal use of a patchy environment. *American Naturalist*, *100*(916), 603–609. https://doi.org/10.1086/282454

Nguyen, Q., Villeval, M. C., & Xu, H. (2016). Trust under the prospect theory and quasi-hyperbolic preferences: A field experiment in Vietnam. *Economic Development and Cultural Change*, *64*(3), 545–572. https://doi.org/10.1086/685434

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). Routledge.

Stallen, M., Rossi, F., Heijne, A., Smidts, A., De Dreu, C. K., & Sanfey, A. G. (2018). Neurobiological mechanisms of responding to injustice. *Journal of Neuroscience*, *38*(12), 2944–2954. https://doi.org/10.1523/JNEUROSCI.1242-17.2018

Turrin, C., Fagan, N. A., Dal Monte, O., & Chang, S. W. (2017). Social resource foraging is guided by the principles of the Marginal Value Theorem. *Scientific Reports*, *7*(1), 11274. https://doi.org/10.1038/s41598-017-11763-3

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. https://doi.org/10.1007/BF00122574

Vahed, S., Galván, E. P., & Sanfey, A. G. (2024). Computational modeling of social decision-making. *Current Opinion in Psychology*, *60*, 101884. https://doi.org/10.1016/j.copsyc.2024.101884

van Baar, J. M., Chang, L. J., & Sanfey, A. G. (2019). The computational and neural substrates of moral strategies in social decision-making. *Nature Communications*, *10*(1), 1483. https://doi.org/10.1038/s41467-019-09161-6

van Baar, J. M., Klaassen, F. H., Ricci, F., Chang, L. J., & Sanfey, A. G. (2020). Stable distribution of reciprocity motives in a population. *Scientific Reports*, *10*(1), 18164. https://doi.org/10.1038/s41598-020-74818-y

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–336). Academic Press.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, *8*, e49547. https://doi.org/10.7554/eLife.49547

Zhang, Y., Zhao, H., Huang, A., Li, X., Shu, X., He, Y., & He, Q. (2024). Social anxiety moderates the process of social evaluation expectations: A drift-diffusion model perspective. *Journal of Psychological Science*, *47*(5), 1044. https://doi.org/10.16719/j.cnki.1671-6981.20240503