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

Elijah P. Galván and Alan G. Sanfey

Abstract

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, firstly, outlining 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, demonstrating 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 dataset.

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. Computational models **formalize** – that is, *mathematically represent* – mental processes under consideration. By committing to a rigorous interpretation of what these mental processes are, computational modeling enables testing of hypotheses at a theoretical level. Despite the considerable upside that this approach offers for decision-making research, and social cognition in general, use of modelling has not yet been widely adopted by social cognition researchers. This is likely due to perceptions that there are prohibitively large barriers of entry into computational modeling. In the current paper, we aim to address these perceived barriers by illustrating, in hopefully a straightforward manner, how one specific kind of computational model can be implemented, and what advantages this approach offers the researcher.

Computational models are used in a myriad of different fields, each which uses its own specific and slightly different definition. In the context of the cognitive sciences, we will here define **computational models** as a *mathematical equation which represents a cognitive process*. These equations take inputs from the environment, apply a mathematical function to these inputs, and produce outputs. These outputs are called **model predictions** as they *represent the predicted outcome of the cognitive process according to the computational model*.

Decision-making is a specific cognitive process: it *is the process of choosing between multiple options*. However, the focus of this section is narrower yet: we will focus on the process of value-based decision-making. **Value-based decision-making** involves *choosing between options based on personal preferences* (i.e. opting for one marshmallow now or having two later). In the context of value-based decisions, 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.

Utilizing value-based decision-making paradigms can yield insights into individual preferences that go beyond what is possible with self-report, making them useful for studying social cognition. When it comes to studying decision-making, participants' **observed preferences**, or *what people demonstrate that they prefer via their behaviors*, are often considered to be more informative than their **stated preferences**, or *what people claim to prefer*. In general, there is a broad range of questions that computational models of decision-making can be used to answer, as the flow chart in Figure 1 illustrates. While there are potential applications for each type of computational model when studying social cognition, here we will illustrate one specific type of computational model of decision-making: the Social Utility Model. Social Utility Models offer perhaps the greatest number of potential use cases for illuminating social cognitive behavior, as they endeavor to evaluate when and why people trade-off their own benefit for the well-being of others.

Figure 1

Computational Models of Decision-Making and When to Use Them

[[Insert Figure 1 Here]]

Note. Reinforcement Learning Models include several specific models, the most widely known of which is the Rescorla-Wagner model (Wagner & Rescorla, 1972). Foraging Models include patch choice models such as Marginal Value Theorem (Charnov, 1976) and prey models such as Optimal Foraging Theory (MacArthur & Pianka, 1966; Emlen, 1966). The most notable Risk Model is Cumulative Prospect Theory (Kahneman & Tversky, 1992). Social Utility Models are a variant of Utility Models, advanced to explain prosocial behavior in microeconomic games (Fehr & Schmidt, 1999; Bolton & Ockenfels, 2000). Intertemporal Choice Models are also a type of Utility Model which aim to capture the tradeoff between reward magnitude and preference for short time horizons (Koopmans, 1960; Breeden, 1979). Motivation Models capture demand for goods as a function of the effort required to obtain them, predicting when people will or will not exert effort to pursue a goal (Hursh, 1993). Process Models focus on evidence integration and include several specific types of models, the most widely used of which are Drift-Diffusion Models (Ratcliff, 1978).

As Figure 1 indicates, Social Utility Models are well-suited to examine decision-making when there is no uncertainty and one’s choices have impact on another person’s outcomes. Just like all utility models, Social Utility Models assume that people make decisions that maximize their **subjective utility**, or *the experienced feeling of subjective satisfaction*. This can be derived from sources such as money, food, experimental tokens, or other tangible goods, but can also be derived from non-economic sources. In the context of Social Utility Models, these non-economic sources are often social in nature, such as satisfaction gained by cooperative acts, reciprocating trust, or indeed, punishing the bad behavior of others.. As we discussed previously, computational models of decision-making aim to represent the process by which people choose between different options. Thus, Social Utility Models posit that people do this by considering how much total utility they could derive from each possible choice in the current situation (i.e. derived from both economic and prosocial sources), then choosing the option which has the highest utility to them personally.

Clearly, in order to determine which choice has the highest subjective utility, we have to be able to quantify the utility of all choices. This brings us to perhaps the most formidable barrier to entry into computational modeling: developing a formal model. At their core, Social Utility Models are mathematical functions which quantify the utility of a choice given (a) the situation in which the choice takes place and (b) the personal preferences of the chooser. The **situation** in which the choice takes place is defined by the *relevant features of the environment which influence how people judge an option*. We can think of the situation as defining what it means to follow the norms that we are interested in studying: what it means for a choice to maximize payout or be “fair” or “right” according to a specifically defined rule (i.e. a norm). The **personal preferences** refer to *the values that the decider holds – usually about what is right and fair*. Equation 1 illustrates the conceptual form of a utility model.

Equation 1:

Now that we know what Social Utility Models are, we turn to looking at what advantages and limitations they have relative to alternative approaches.

Advantages and Limitations

The use cases of Social Utility Models are limited in two ways by the nature of the inputs to the model. We can, first and foremost, only use these models when we can identify and quantify all of the relevant aspects of the situation: namely the value of each independent variable(s), relevant constant(s), and all possible values of the dependent variable. For this reason, these models are typically applied to experimental paradigms in which participants are tasked with allocating money. Thus, the same set of limitations that apply to experimental monetary allocation tasks also apply to Social Utility Models, most notable of which is the potential lack of ecological validity when compared to real-life social choices.

The second limitation is that, in general, several observations of each decision scenario are required for model fitting. A key component of Social Utility Models are the presence of Free Parameters; however, people obviously do not come into the laboratory readily prepared to announce their Free Parameter values. Instead, these must be assessed from their choices, a process termed parameter recovery. **Recovering Free Parameters** is thus used to *infer which preferences have guided the decision-making process*. Thus, we need multiple observations of the decisions people make in order to reliably infer what these preferences are. Since these models are fit at the individual level, we often need many observations from each participant: depending the signal-to-noise ratio of the task, the number of available choices, and the number of Free Parameters, this can range from as few as ten to more than one hundred trials. This requirement can therefore obviously limit the type of experimental tasks that are suited to employing Social Utility models, and the types of populations these models can be used to study.

In addition to these limitations, there are also certain constraints making *valid* use of this approach. First, and perhaps obviously, it is imperative to apply sound experimental design principles: Social Utility Models can only be tested in an experiment wherein the different psychological preferences represented in the model would produce different patterns of behavior. Second, while Social Utility Modeling is a user-defined analysis approach (i.e. you write your own functions), it is still important to adhere to basic statistical principles, as assumptions such as homoscedasticity, normality of error, linearity, and error independence still apply in hypothesis testing. Finally, it is crucial to exhaustively test models. These models **embody** hypotheses, meaning that they *generate trial-by-trial predictions stemming from the hypotheses informing the equation*. Because computational models represent the decision-making process, the hypotheses of the best model are accepted. Thus, we must compare all reasonable, possible, models that could characterize the data – no stone must be left unturned.

These limitations should be taken seriously, however we believe that Social Utility Modeling does offer important advantages over alternative approaches, for example a standard Linear Mixed Effects Modeling approach. When thinking about the key differences between Social Utility Models and Linear Mixed Effects Models, it is useful to first understand what these models aim to do. As we have already discussed, computational models of decision-making attempt to capture how people decide between different possible choices. We can therefore classify Social Utility Models as **generative models**, meaning that they *generate predictions about how people behave without "knowing" the data first*. By comparison, Linear Mixed Effects Models are **descriptive** models, meaning that they *describe how one or more Independent Variables relate to one or more Dependent Variables*. The relative advantages that Social Utility Models have with respect to Linear Mixed Effects Models are a consequence of this distinction.

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 recovered Free Parameters reflect what individual people value: that is, individual differences are treated as an important feature of the model. By contrast, individual differences are handled as random effects in Linear Mixed Effects Models which serve to regress out individual differences in order to give better signal to the "real" effect, termed fixed effects in the model.

Secondly, Social Utility Models enable us to detect the role of multiple different competing processes in a single behavior - Linear Mixed Effects Models tend to average across these processes. Linear Mixed Effects Models attempt to capture the relationship between one or more Independent Variable and the Dependent Variable of interest. By contrast, Social Utility Models differ in the sense that they incorporate multiple constructs and therefore predict multiple relationships between the Independent Variable and Dependent Variable of interest. Thus, where Social Utility Models are fit based on how well an individual's choices follow the theoretical constructs in the model, Linear Mixed Effects Models are fit to the data based on the average relationship between each Independent Variable and the Dependent Variable over the entire sample.

The third advantage of using Social Utility Models instead of Linear Mixed Effects 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 which only differ in the specific hypothesis we want to test. Thus, the outcome of hypothesis testing is a direct psychological conclusion about what the data say (i.e. people apply *this* principle but not *that* principle to guide their decision-making). In contrast, Linear Mixed Effects Models indicate what variables have an effect on behavior, from which we must extrapolate conclusions about the psychological processes which produced the observed behavi0r (i.e. *this* variable but not *that* interaction has an effect on decisions, which we believe means that a certain principle may be influential in this decision).

Now we will turn our attention towards how to practically implement Social Utility Models in social cognition research, specifically examining social decision-making.

Active Tutorial

In this section, we will present a high-level tutorial of Social Utility Modelling 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 which you can practice using actual data from real experiments (Crockett et al., 2014; van Baar, Chang, & Sanfey, 2019; Li et al.; 2022; Galván & Sanfey, 2024). These tutorials are available in Python, MatLab, and R to allow for implementation in whichever language you are most familiar with. The present paper is 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, Chang, and Sanfey's (2019) paper which sought to answer the following research question: "What motivates people to reciprocate trust, even when there are no external incentives to do so?". As you can see, the research question asks about a specific value-based decision (i.e. to reciprocate or not, and if so, how much), which impacts both the decision-maker at well as other people. To examine questions about reciprocity, we can employ a well-studied experimental task known as the Trust Game (Berg, Dickhaut, & McCabe, 1995). 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?

To answer this question, we need to identify some plausible answers. First, it is possible that people don't actually reciprocate at all, preferring to keep the money for themselves – we can term this strategy as '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 their game partner. Presumably the Investor sends money in the hope of getting a return, therefore if Trustee not reciprocate they will violate their partner's expectations will let down this player – we can call this strategy 'Guilt-Aversion'.

To create a task where we 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, Greedy, Inequality-Averse or Guilt-Averse. To accomplish this, van Baar, Chang, and Sanfey (2019) made the investment multiplier either 2, 4, or

6 but they withheld this information from Investors, telling them that the multiplier was always 4. Thus, Investors expect half of what they believe to be the multiplied investment (Expectation = ½ × Investment × Believed Multiplier) but an equal distribution would differ from this expectation when the multiplier is 2 or 6 (Equality = ½ × (Investment × Multiplier + Endowment – Investment); Investment × Multiplier is what the Trustee possesses prior to the decision and Endowment – Investment is what the Investor kept, so the sum is the total amount of money in the game).

The next step is to create the Social Utility Model. As mentioned before, Social Utility Models represent hypotheses about what people value, in this case assuming you are the participant in the role of Trustee. Let's look at the example of Greed and create a utility equation which embodies the hypothesis that you only value money. Since this model intends to capture the notion that utility only comes from money, we could very simply express utility as being 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 2).

Equation 2:

As we've mentioned before, utility can also come from sources other than money. Let's take a look at how we can embody the hypotheses that you follow social norms. To represent the utility of following social norms we need to follow four steps:

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

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.[1]

[1] In what follows, we apply steps 1-4 to explain behavior in van Baar, Chang, & Sanfey's (2019) experiment. By following these instructions, we arrive at a different formulation

Let's apply this process to Guilt-Aversion wherein the highest possible utility you could have is by giving their partners exactly what (you believe) they expect: we can subtract this from what your partner actually receives. We then want to normalize this between 0 and 1 by dividing the maximum possible absolute difference between your the partner expects and you could give them. So, this term will be 0 when you meet your partner's expectations and 1 when you maximally violate these expectations. 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, previous models of Guilt-Aversion square the Guilt-Aversion term (Dufwenberg & Gneezy, 2000; Battigalli & Dufwenberg, 2007), so let's just do the same – we thus arrive at the model of Guilt-Aversion shown in Equation 3.

Equation 3:

Now we can apply the same format for Inequality-Aversion. Perfect Inequality-Aversion is achieved when the payouts of both you and your 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 (Fehr & Schmidt, 1999; Bolton & Ockenfels, 2000), which gets us to Equation 4.

Equation 4:

So now we have 3 different Utility Equations which each represent a distinct hypothesis: namely, that you will follow the same norm when making reciprocation decisions. However, these hypotheses are not necessarily wholly exclusive across multiple trials of the task: you can potentially choose to follow different norms under different circumstances when making your reciprocation decisions. We can thus usefully create a single model, which requires three fundamental steps.

1. Mathematically 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*

---

of the model compared to the one which appears in the paper. Importantly however, the models fit the data equally well and are functionally equivalent.

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

We have already done the first step: we'll just take the functions from Equations 2-4 to represent Greed, Guilt-Aversion, and Inequality-Aversion.

For step two, we then want to identify how these 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. We use Greek letters to represent Free Parameters in our model, so let's call Dimension 1 Θ in our model and Dimension 2 Φ in our model.

Moving to step three, we want to assign Free Parameters to each norm and mathematically represent these in our model. 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, then add all of these terms together, we arrive at our Social Utility Model as shown in Equation 5.

Equation 5:

Now that we have a model, we need to create a trial set; here, we determine the number of trials in the experiment and what those trials will look like to participants. Once you have done this, you can generate model predictions. In order to do this, you only need to supply Free Parameters to the model. A useful way of viewing Free Parameters is that they are the dimensions on which people can meaningfully differ from each other in your task: these dimensions alone represent all of the possible ways that people can behave differently in your experiment. Therefore, the process of generating model predictions answers the question *"What would a person with this specific set of preferences do in my task?"*.

After generating model predictions, there are still two crucial steps prior to model implementation. The first and most important step is to ensure that Free Parameters can be accurately identified from the data that created them – we do this by recovering Free Parameters from the model predictions. In a moment, we will walk through the process of Free Parameter Recovery which is the same irrespective of whether data is fake or real. The other step is to determine the possible range of values for the Free Parameters. The benefit of normalizing the terms in a model to be between 0 and 1 is that your Free Parameters should almost certainly range from 0 to 1, but this should be checked nonetheless. Additionally, 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.

This all gets us to the point of actually analyzing experimental data. At this stage, the first priority is to recover Free Parameters: the model cannot be assessed without knowing what it predicts for each participant, and we can only know what it predicts if we know what their Free Parameters are. To reiterate, Social Utility Models predict that people maximize their subjective utility: thus, we want to find a set of Free Parameters where the difference between expected-versus-observed utility is smallest. We can think of **expected utility** as *the utility predictions for a specific set of Free Parameters* and **observed utility** as *the utility which someone with a specific set of Free Parameters would have experienced if they made the choices we see*. This difference between expected and observed utility is the criteria that we use to assess how “good” the Free Parameters are. Thus, to recover parameters we first create an **objective function** which takes the participant’s Choices, IVs, and Constants alongside a proposed set of Free Parameters as inputs, and then outputs the utility error measure (usually the sum-of-squared differences between expected and observed utility).

Recovering Free Parameters is basically a trial-and-error process, but this can be speeded up by using an optimizer. **Optimizers** *find the ideal solution to an objective function*, which means that we can simply give the optimizer our objective function as well as a participant’s data and it will produce Recovered Free Parameters. This can save a great deal of time when compared to exhaustively sampling the entire parameter space (for this model, it is about 100 times faster).

Once we have recovered the Free Parameters for all participants, we can use these Free Parameters to determine model predictions per participant. Therefore, we use the data to infer the Free Parameters which we then use to assess the data with goal that the model is a good predictor of the data. This is somewhat circular of course, though less so than regression models such a Linear Mixed Effects Models which fit parameters to the data directly (rather than to a psychological variable like utility), but is still potentially problematic. To ensure that we are not relying on circular reasoning, we need to assess how well *the model predicts data that it hasn’t yet seen*, known as **out-of-sample prediction**.

The most conventional approach to out-of-sample prediction is **fivefold validation**. Here, *the data is divided into five equal parts, and for each fifth we use the other four-fifths of the data to recover Free Parameters which are then used to predict the remaining one-fifth of the data.* The goal here is to show that the fivefold validation model can explain data that it hasn't seen at a level comparable with the normal model. We also want to demonstrate that the Free Parameters that we recover are reliable estimates of our actual Free Parameters.

We now move to hypothesis testing. The central approach to hypothesis testing is termed **model comparison** wherein we *compare models in terms of how well they explain the data*. We do this 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 or the deviance) and **parsimony** is simply *the number of Free Parameters* employed. Since smaller prediction error and fewer Free Parameters mean better performance, lower MFI values indicate a better model. We can also conduct statistical tests using MFIs by computing these per participant per model, then utilizing paired t-test to assess the respective models.

To illustrate this, let's revisit our example. If we want to test the hypothesis that all three norms (Greed, Inequality-Aversion, Guilt-Aversion) are used in guiding reciprocation decisions we need to demonstrate that including all three norms significantly improves model performance over simplifications of the model. Thus, we would need to compare the MFIs for the model presented in Equation 5 with models that only have a single norm (i.e. Equations 2-4), as well as models that contain only two of the norms. Using the same logic, we can also explicitly test for individual differences by assessing whether or not recovering Free Parameters per participant significantly improves model performance compared to recovering Free Parameters for the whole sample. If we have multiple conditions (i.e. some kind of treatment or manipulation versus a control) and want to see if these affect preferences, we could test whether or not recovering Free Parameters per condition per participant significantly improves model performance compared to recovering Free Parameters for all conditions per participant.

If you want to see the results of this analysis, you can follow the first tutorial in the handbook, or just read the paper!

Concluding Remarks

When applied to the study of social cognitive processes, Social Utility Modeling offers the potential to evaluate psychological factors which can powerfully shape differences in social behavior. Despite the fact that 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 offer advantages over alternative possibilities.

One particular advantage is that it allows us to do model testing in an integrated, multimodal, way. Specifically, though 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 (i.e. Galván & Sanfey, 2024). Figure 2 illustrates the opportunities offered by Social Utility Modeling in comparison to other methods.

Figure 2

Data Types and Methodologies for Addressing Research Questions

[[Insert Figure 2 Here]]

Note. Panel A illustrates conventional modeling approaches. Panel B illustrates how Social Utility Modeling uses the same kinds of data to provide more informative insights.

Computational models offer tremendous potential for research in social cognition, though 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 which social cognition researchers may find particularly informative.

References

Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. American Economic Review, 97(2), 170-176.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, *10*(1), 122-142.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, *91*(1), 166-193.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of financial Economics*, *7*(3), 265-296.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, *9*(2), 129-136.

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.

Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. Games and economic Behavior, 30(2), 163-182.

Emlen, J. M. (1966). The role of time and energy in food preference. *The American Naturalist*, *100*(916), 611-617.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817-868.

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.

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

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.

MacArthur, R. H., & Pianka, E. R. (1966). On optimal use of a patchy environment. *The American Naturalist*, *100*(916), 603-609.

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

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

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.

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. *Inhibition and learning*, 301-336.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife*, *8*, e49547.