Do people use social information to improve predictions about everyday events?

Stephen Payne (paynesc@rose-hulman.edu)
Department of Computer Science and Software Engineering
Rose-Hulman Institute of Technology

Alan Jern (jern@rose-hulman.edu)
Department of Humanities, Social Sciences, and the Arts
Rose-Hulman Institute of Technology

Abstract

Following Griffiths and Tenenbaum (2006), we explore whether people use relevant social information to improve their already nearly optimal predictions about quantities in everyday events. We tested this question in two experiments involving quantities in three domains: cake baking times, movie runtimes, and podcast lengths. In Experiment 1, we found that participants were sensitive to the difference between relevant and irrelevant social information. In Experiment 2, we found that people consistently used relevant social information to adjust their predictions in the expected directions. We introduce an optimal social prediction model but find that it does not consistently perform better at accounting for our participants' social predictions than an optimal non-social prediction model. We conclude by discussing whether people use social information for prediction in an optimal way.

Keywords: prediction; social inference; Bayesian inference

Sometimes, the quickest way to learn something is to ask someone who knows the answer. But we don't always need to directly ask others to benefit from their knowledge. Simply observing how other people act can provide us with information about what they know, due to our ability to draw inferences using our understanding of the relationship between other people's knowledge and their behavior (e.g. Baker & Tenenbaum, 2014; Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; D. T. Gilbert, 1998; Malle & Knobe, 1997; Malle, 2006).

We can also use this ability for prediction. For example, suppose that a cake has been baking in the oven for 20 minutes. What would you predict is the total baking time of the cake? This example comes from a study by Griffiths and Tenenbaum (2006) in which they found that people's predictions across multiple domains were remarkably close to optimal Bayesian predictions, given the actual distributions of quantities like baking times from cake recipes.

Now suppose that the person who made the cake is in the kitchen leaning against the oven at the 20-minute mark. What would you now predict is the total baking time? Intuitively, it seems odd for someone to be doing this unless they expect the cake to be ready soon. As a result, in this case, you might lower your predicted baking time. If so, the additional information about the person caused you to change your prediction due to your understanding of the relationship between people's knowledge and and their behavior. More importantly, the information potentially allowed you to make a more accurate prediction. When the person is not in the kitchen, you have nothing but your general knowledge of cake baking times to go on. But when the person is in the kitchen, you have a valuable piece of indirect information to incorporate into your prediction.

Broadly speaking, this is an example of what social psychologists call informational social influence: being influenced by others because you believe they know something you do not (Detusch & Gerard, 1955; Cialdini & Goldstein, 2004). More specifically, there is evidence that people take social information into account when making inferences about things like whether an action was causally responsible for an outcome (E. A. Gilbert, Tenney, Holland, & Spellman, 2015; Goodman, Baker, & Tenenbaum, 2009; Lagnado & Channon, 2008; Kirfel & Lagnado, 2021).

In this paper, we explore whether people also take social information into account when making everyday predictions about quantities like cake baking times. We tested this question in two experiments. In the next section, we briefly review the modeling approach introduced by Griffiths and Tenenbaum (2006) for generating optimal predictions. We then describe our experiments designed to test whether people are sensitive to relevant social information and if the predictions they make by incorporating that information deviate from the model's predictions.

Predicting quantities

Griffiths and Tenenbaum (2006) introduced a Bayesian model of optimal (non-social) prediction, which we briefly review here, using the cake-baking example to illustrate. The model computes a posterior probability distribution over the total baking time t_{total} as follows:

$$p(t_{total}|t) = \frac{p(t|t_{total}) \cdot p(t_{total})}{p(t)}, \tag{1}$$

where t is the initial observed time—20 minutes in our example. The likelihood term, $p(t|t_{total})$, specifies the probability that you notice the cake at time t of the cake's total baking time. For simplicity, the model assumes that it is equally likely that you observe an event at any point in its duration. Thus, in this example, the model assumes it is equally likely that you would see the cake in the oven right after it was put in as just before it was done. In other words, the likelihood term is uniform: $p(t|t_{total}) = \frac{1}{t_{total}}$.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

$p(t_{total})$ specifies a prior probability distribution over the quantity being estimated. In our example, $p(t_{total})$ is a distribution of cake baking times. This distribution will vary depending on the scenario—cake baking times will have a different distribution than movie runtimes—so we estimated this distribution using empirical data separately for each scenario in our experiments, described later.

Following Griffiths and Tenenbaum (2006), we define the final prediction of the model t^* to be the median of the posterior probability distribution. Intuitively, t^* represents a "best guess" estimate of the duration of the event. Formally, t^* is defined such that[1]:

$$\sum_{k=t}^{t^*} p(t_{total}|k) = 0.5$$

Incorporating social information

Now suppose that o represents an observation—like the fact that someone is leaning up against an oven with a cake in it. We can modify Equation 1 to incorporate o as follows:

$$p(t_{total}|t,o) = \frac{p(t_{total}) \cdot p(t,o|t_{total})}{p(t,o)}. \quad (2)$$

If we again assume that t is uniformly distributed between 0 and t_{total}, and additionally assume that o is a function of both t and and t_{total}, the likelihood term can be expressed as:

$$\begin{aligned} p(t,o|t_{total}) &= p(t|t_{total}) \cdot p(o|t,t_{total}) \\ &= \frac{1}{t_{total}} \cdot p(o|t,t_{total}) \end{aligned} \quad (3)$$

Thus, the key difference between a non-social and social prediction model is in the term $p(o|t,t_{total})$. Later, we introduce one way of defining this term.

Our primary research question was whether people use relevant social information to optimally adjust their predictions of total quantities. But first we asked a more basic question: whether people are sensitive to the difference between *relevant* and *irrelevant* social information for the purposes of prediction. For example, seeing the person who made the cake leaning up against the oven seems relevant for predicting how much more time will pass before the cake is done. But seeing someone who had nothing to do with making the cake do the same provides much less useful information.

The term $p(o|t,t_{total})$ captures the probability of an observation like seeing someone leaning against the oven. In Experiment 1, we asked participants to make judgements about probabilities like these. Specifically, they rated the likelihood of seeing several events involving people at different times. We hypothesized that they would rate irrelevant events—like a random person leaning against the oven—as about equally likely to occur at any time. By contrast, we hypothesized that participants would rate relevant social events—like the baker of a cake leaning against the oven—as more likely to occur at later times.

[1]For details about computing $p(t_{total}|t)$, see Griffiths and Tenenbaum (2006).

Experiment 1

Method

All methods and statistical analyses were preregistered: `https://osf.io/zsmda`. All data, code, and materials from Experiments 1 and 2 are available at `https://github.com/jernlab/social-prediction`.

Participants Participants completed the experiment online through Prolific and were paid for participation. We collected data in batches and replaced participants who were excluded (one for failing an attention check described below) until achieving a pre-planned sample size of at least 60 (final N = 61).

Design and Procedure On each page of the experiment, participants saw a scenario such as "A cake has been baking in an oven for 10 minutes". Subjects were then asked to rate the likelihood (on a 1–7 scale from "not at all likely" to "very likely") of both a *relevant* social observation and an *irrelevant* social observation. Participants repeated this for the same scenario for four additional values of t (e.g., the amount of time the cake was baking in the oven). They then repeated this entire procedure for two more scenarios.

Participants saw three scenarios in a random order: a cake baking in an oven, a movie playing in a theater, and someone listening to a podcast. The first two scenarios were inspired by those used in Griffiths and Tenenbaum (2006). The full text of each scenario and the relevant and irrelevant observations used are below. In each case, X was replaced with the t value.

- **Cake**: A cake has been baking in an oven for X minutes. On the following scale, how likely do you think it would be to now see ...
  - *Relevant*: the person who made the cake is leaning against the oven.
  - *Irrelevant*: someone who didn't make the cake is leaning against the oven.
- **Movie**: A movie has been playing for X minutes. On the following scale, how likely do you think it would be to now see 10 people ...
  - *Relevant*: exit the movie's showroom.
  - *Irrelevant*: exit the showroom of a movie next door.
- **Podcast**: Someone has been listening to a podcast for X minutes. On the following scale, how likely do you think it would be to now hear ...
  - *Relevant*: the podcast host say "Welp, that's all we planned to discuss this week!"
  - *Irrelevant*: someone nearby says "Hey, I like that podcast too. Cool."

Each participant rated the likelihoods of these events for multiple values of t, which we refer to as levels. The levels

	Level				
Scenario	1	2	3	4	5
Cake	10	20	35	50	70
Movie	30	45	60	85	100
Podcast	15	30	55	75	105

Table 1: Values of t (in minutes) for each level of each scenario in Experiments 1 and 2.

of each scenario are shown in Table 1. Participants provided ratings for all levels of one scenario in increasing order before moving to another scenario.

Experiment 1 therefore employed a 3 (scenario) $\times$ 5 (level) $\times$ 2 (observation: relevant vs. irrelevant) within-subjects design.

After answering all questions, participants completed an attention check: a multiple choice question asking what topic (i.e., scenario) was not mentioned in the survey.

Results

Figure 1 shows participants' ratings for each scenario. To test our hypotheses, we fit a linear mixed effects model (Brown, 2021) to the full data set using observation type and level (and their interaction) as fixed effects and fitting by-scenario random intercepts, and by-subject random slopes and intercepts.

As predicted, there was a significant effect of level, such that each increase in level resulted in an estimated increase in likelihood rating of about 0.5 ($\hat{\beta} = 0.52$, $SE = 0.06$, $t = 9.33$, $p = 1.71 \times 10^{-15}$). However, this effect was qualified by a predicted interaction with observation type (relevant or irrelevant), also significant. Specifically, the increase in likelihood ratings as level increased was smaller for irrelevant observations than for relevant observations, resulting in a difference in estimated slope of about 0.4 ($\hat{\beta} = -0.40$, $SE = 0.06$, $t = -7.31$, $p = 4.11 \times 10^{-13}$). This interaction is evident in Figure 1 by comparing the ratings for the relevant and irrelevant conditions.

Discussion

Experiment 1 suggests that people are sensitive to the difference between relevant and irrelevant social information. Our participants did not see a relationship between t and the irrelevant observations, but did see a relationship between t and the relevant observations. This suggests that people will incorporate relevant social information into their predictions about quantities like a cake's total baking time. We tested this question in our second experiment.

Experiment 2

In Experiment 2, participants predicted durations of events, with or without relevant social information. We predicted that if participants incorporated the social information we presented them with into their predictions, they would give consistently lower predictions than participants who did not have this information.

Method

All methods and statistical analyses were pre-registered: `https://osf.io/qczes`.

Participants

300 participants (150 in each condition) completed the experiment online through Prolific and were paid for participation.

Design and Procedure Experiment 2 largely resembled Experiment 1 in design and procedure. Participants saw the same three scenarios in random order, each with five different t values in Table 1 in increasing order. There were two important differences. First, participants were asked to predict total times (like the total baking time of the cake). Second, rather than answering questions about relevant and irrelevant social observations, participants were randomly assigned to one of two between-subject conditions: a *non-social condition* or a *social condition*.

Participants in the non-social condition were only provided with the basic scenario. Participants in the social condition were also given the relevant social information from Experiment 1. Subjects were then asked to predict the total time for each scenario. For example, in the cake scenario, they were asked, "What would you predict is the *total baking time of the cake? (in minutes)*". The italicized part of the question varied by scenario. Complete descriptions of the scenarios are below. The text in brackets appeared only in the social condition. X was replaced with the t values from Table 1.

- **Cake**: A cake has been baking in an oven for X minutes. [*Social condition:* At that time, the person who made the cake is leaning against the oven.]
- **Movie**: A movie has been playing for X minutes. [*Social condition:* At that time, 10 people exit the movie's showroom.]
- **Podcast**: Someone has been listening to a podcast for X minutes. [*Social condition:* At that time, the podcast's host says, "Welp, that's all we planned to discuss this week!"]

Experiment 2 therefore employed a 3 (scenario) $\times$ 5 (level) $\times$ 2 (context: non-social vs. social) mixed design.

After answering all questions, participants completed the same attention check as in Experiment 1.

Non-social model predictions Recall that the distribution of $p(t_{total})$ in Equation 1 is likely to vary by quantity. We therefore generated empirical estimates of $p(t_{total})$ separately for each scenario:

- **Cake**: We collected data from the BBC Food website. We searched all recipes of traditional dessert cakes (excluding things like crab cakes) and recorded the recommended total baking times for each.
- **Movie**: We used The Movie Database (TMDb) 5000 dataset published in 2018. The dataset includes the durations, along with titles, popularity ratings, and other vari-

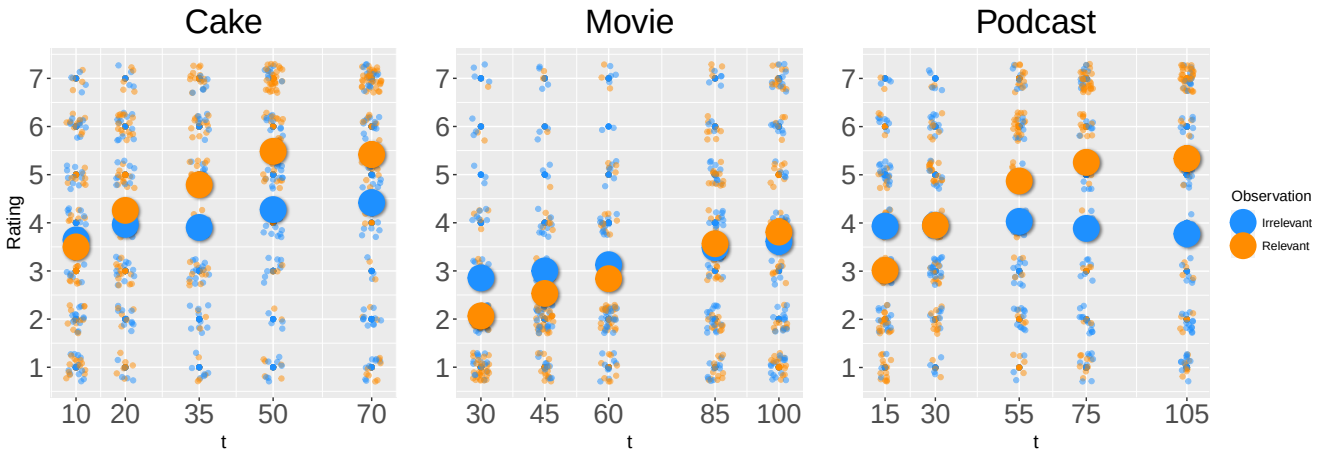

Figure 1: Experiment 1 data. The x-axes show the t values and the y-axes show the ratings on a 1–7 scale. The large points represent mean ratings with 95% confidence intervals (in most cases barely visible). Each smaller point represents a single subject's rating.

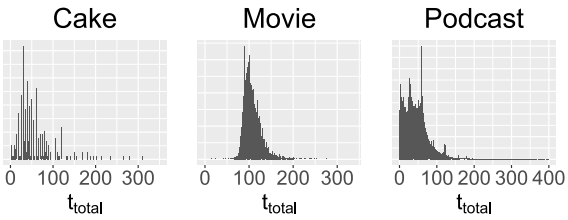

Figure 2: $p(t_{total})$ distribution for each scenario.

ables for nearly 5000 movies catalogued on the TMDb website.

- **Podcast**: We used a public iTunes podcast dataset published in 2017. The data was obtained from iTunes and includes the durations, names, descriptions, and other variables for over 10,000 podcasts.

For each scenario, we calculated relative probabilities of the raw data from each dataset as the distribution $p(t_{total})$. The resulting distributions are shown in Figure 2.

Results

Individual participant predictions were excluded as outliers if they were more than three standard deviations from the mean for a given context (social or non-social) and level (t value). Out of 4500 total data points, 83 (1.8%) were excluded.

Figure 3 shows the experiment results. In all scenarios, mean predictions were consistently lower in the social condition compared to the non-social condition.

We fit a linear mixed effects model to the full data set. Because the model in our pregistration plan resulted in a singular fit, we simplified the model by removing the random slopes from the model and re-fit it. The resulting model included context (social vs. non-social) and level (and their interaction) as fixed effects and by-scenario and by-subject random intercepts.

As predicted, there was a significant effect of context, such that social predictions were estimated to be about 17 minutes lower than non-social predictions ($\hat{\beta} = -16.5$, $SE = 1.48$, $t = -11.14$, $p < 2 \times 10^{-16}$). Also as predicted, there was a significant effect of level, such that each increase in level resulted in an estimated increase in total predicted time of about 14 minutes ($\hat{\beta} = 13.69$, $SE = 0.32$, $t = 43.31$, $p < 2 \times 10^{-16}$).

However, there was also a significant interaction between context and level. Specifically, the increase in predictions as level increased was slightly greater for the social condition than for the non-social condition ($\hat{\beta} = 2.04$, $SE = 0.45$, $t = 4.56$, $p = 4.91 \times 10^{-6}$). We had no specific predictions regarding an interaction, but we considered the possibility that we might find an interaction in the results, given that we were using three very different scenarios with different empirical distributions.

The results in Figure 3 suggested that this interaction might be limited to only certain scenarios, so we ran a follow-up exploratory analysis that we did not preregister. Specifically, we fit separate linear mixed effects models to subsets of the data for each of the three scenarios. We used the same statistical model as before, but with the actual t values as a fixed effect instead of level. This decision allowed us to get more interpretable estimates of the fixed effects.

The effects of context and t value were significant for all scenarios (all ps $< 1 \times 10^{-9}$). For the movie scenario, participants' predictions were estimated to be about 21 minutes lower than the non-social predictions ($\hat{\beta} = -21.30$, $SE = 3.68$, $t = -5.79$, $p = 8.63 \times 10^{-9}$) and there was a significant interaction between context and t level ($\hat{\beta} = 0.21$, $SE = 0.05$, $t = 3.94$, $p = 8.58 \times 10^{-5}$). For the podcast scenario, participants' predictions were estimated to be about 23 minutes lower than the non-social predictions ($\hat{\beta} = -22.78$, $SE = 1.25$, $t = 18.20$, $p < 2 \times 10^{-16}$) and there was a significant interaction between context and t level ($\hat{\beta} = 0.11$, $SE = 0.02$, $t = 5.44$, $p = 6.46 \times 10^{-8}$). For the cake scenario,

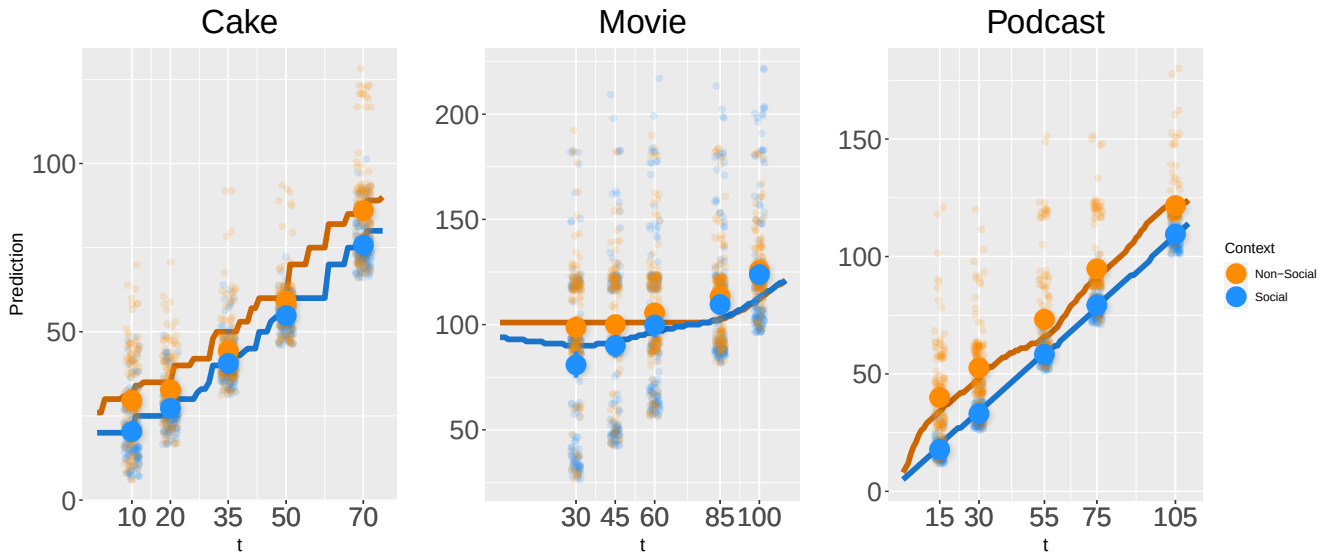

Figure 3: Experiment 2 results. The x-axes show the t values and the y-axes show the predicted t_{total} values. The large points represent mean ratings with 95% confidence intervals (in most cases, barely visible). Each smaller point represents a single subject's prediction. The orange and blue lines show the non-social and social model predictions respectively.

participants' predictions were estimated to be about 6 minutes lower in the social condition compared to the non-social condition ($\hat{\beta} = -5.98$, $SE = 0.95$, $t = -6.30$, $p = 4.03 \times 10^{-10}$), and the interaction between context and t level was not significant ($\hat{\beta} = -0.02$, $SE = 0.02$, $t = -0.84$, $p = 0.40$). In sum, it appears that the interaction between context and t value was small and driven by the movie and podcast conditions.

Discussion

When given different relevant social information, our participants consistently made lower predictions than when they did not have the social information. Additionally, we generated model predictions for the non-social conditions using the optimal Bayesian model developed by Griffiths and Tenenbaum (2006). The model's predictions matched our participants' predictions quite well (and the social predictions less so), even for distributions with different empirical distributions. Experiment 2 therefore largely replicated their results. But the results from our social condition suggest that their model needs to be modified to incorporate social information if it is to account for how people use social information for prediction. We describe one way this can be done in the next section.

One result we did not predict was the interaction between context and t, reflected in Figure 1 by the convergence of participants' mean predictions between the social and non-social conditions at higher values of t, perhaps most evident in the movie scenario. One possible explanation for this result is that the relative information value provided by the social observations in our scenarios diminished as t increased because participants already had considerable existing knowledge of the empirical distributions. For example, if you see people leaving a theater 60 minutes into a movie, that is somewhat surprising and you'd probably assume they know something you don't. But if you see people leaving 100 minutes in, it's less surprising as you likely expected the movie to be ending soon anyway.

An optimal social prediction model

While our results strongly support the conclusion that people take relevant social information into account when making predictions about quantities, an important question remains: are they doing so in an optimal way? This was the original question posed by Griffiths and Tenenbaum (2006). We now return to our specification of an optimal social prediction model in Equations 2 and 3.

Recall that fully specifying a social prediction model requires defining the term $p(o|t, t_{total})$. Here, we adopt a utility-based approach that assumes agents choose actions that increase rewards and minimize costs (Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Jara-Ettinger, Schulz, & Tenenbaum, 2020). We make a simple assumption that, for example, a baker derives greater rewards from being near the oven as a cake nears closer to being done baking. Specifically, we define reward r as:

$$r = \frac{1}{t_{total} - t}. \tag{4}$$

In many situations, the probability of someone taking action will not increase linearly with r. For example, you are not likely to see the baker of a cake waiting near the oven until the cake is nearly done. Therefore, we defined the likelihood

	Model	
Scenario	Social	Non-social
Cake	5206.0	**3179.0**
Movie	**2650.9**	5589.1
Podcast	5016.6	**4802.5**

Table 2: BIC results for each model by scenario for Experiment 2 social context data. The best-scoring model for each scenario is marked in bold.

term, $p(o|t,t_{total})$, to be a logistic function of r[2]:

$$p(o|t,t_{total}) = \frac{1}{1+\exp^{-\beta_0+\beta_1 r}}\,. \quad (5)$$

Because the probability of observing someone take an action might depend on the situation (leaving a movie theater before a movie has ended may be less likely than standing in front of the oven several minutes before a cake has finished baking), we fit parameters β_0 and β_1 to the data for each scenario of Experiment 2.

Model predictions

We fit the model to data in each scenario by performing a grid search of parameters (b_0 from -20 to 20 in increments of 0.5, and b_1 from 0 to 400 in increments of 5) minimizing mean-squared error (MSE) with participants' predictions in Experiment 2. Figure 3 shows the resulting predictions.

In an exploratory analysis that was not preregistered, we tested whether the additional assumptions of the social model provided a better fit to the data from our social condition than the original Griffiths and Tenenbaum (2006) non-social model by calculating the Bayesian Information Criterion (BIC)—which penalizes models with extra parameters—for each model in each scenario. We used the posterior distribution ($p(t_{total}|t,o)$ or $p(t_{total}|t)$) as the likelihood of each model[3]. The results are shown in Table 2.

Despite the social model's close fit to mean predictions in Figure 3, it did not compare favorably with the non-social model, resulting in a higher BIC in two of the three scenarios. Our analysis indicates that, except for in the movie scenario, the additional assumptions of our social model do not provide much additional explanatory power.

Do people use social information to make *optimal* predictions?

Does the relatively poor performance of our social model mean that people's social predictions were not optimal? Our results make it difficult to judge for several reasons.

First, our social model, which incorporated some simple assumptions about how people's behavior in multifaceted situations, is a function only of how much time is remaining before an event. This is certainly an over-simplification; our participants may have had more complex models of human behavior in mind.

An alternative to directly specifying the likelihood term $p(o|t,t_{total})$ would be to use people's own judgments as empirical likelihood estimates. For example, in our Experiment 1, we asked people how likely it was to see someone in the kitchen leaning against the oven with a cake that had already been baking in it for 20 minutes. Ratings like these could be used as inputs to a model that generates predictions of the most likely total baking times for cakes, given that someone is leaning against the oven.

We were unable to use our own data for this purpose because we did not provide participants with one key piece of information: t_{total}. As a result, they likely inferred this information and incorporated it into their ratings. To get true empirical likelihood judgments from people, they would need to be provided with the observation (e.g., the person leaning against the stove), the time t (e.g., the elapsed baking time), and the total duration of the event (e.g., the total baking time). Future work may wish to collect these judgments as a means of generating social model predictions to test if people's predictions are at least consistent with their own likelihood judgments (and with the empirical distributions of the events themselves).

Another difficulty with judging whether people use social information information to make optimal predictions using our data is that people may have been suboptimal in *both* the non-social and social conditions. For example, if our participants provided over-estimates in both conditions, it might explain why the non-social model provides a good fit to the social data.

Conclusion

When we have relevant knowledge, we generally use it. And when we don't know things, it's natural to look to others around us who are more knowledgeable. Our work suggests that people sometimes do both at the same time: combine their existing knowledge of a domain with relevant social cues from others who likely have specific additional knowledge. Our work provides additional support for the idea that, without the social cues, people are capable of making close to optimal predictions about real-world quantities. But additional research will be needed to determine if the way people incorporate social information into those predictions is also close to optimal.

Acknowledgements

Rose-Hulman's IP/ROP provided funding for this project. We thank multiple anonymous reviewers for valuable feedback.

[2]We thank an anonymous reviewer for this suggestion.

[3]Recall that we used empirical prior distributions for each scenario. Because the datasets that these distributions were based on did not include some values of t, the resulting posterior probabilities for the models were technically 0 for those values. We chose to omit data points for which the posteriors would be 0 in our BIC calculations (i.e., where people predicted values of t that were not in our data sets). The total number of omitted data points, at most, was 15.5% in one scenario.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1–10.

Baker, C. L., & Tenenbaum, J. B. (2014). Modeling human plan recognition using Bayesian theory of mind. In G. Sukthankar, R. P. Goldman, C. Geib, D. Pynadath, & H. Bui (Eds.), *Plan, activity and intent recognition: Theory and practice.* Morgan Kaufmann.

Brown, V. A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–19.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621.

Detusch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629–636.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1). New York: Oxford University Press.

Gilbert, E. A., Tenney, E. R., Holland, C. R., & Spellman, B. A. (2015). Counterfactuals, control, and causation: Why knowledgeable people get blamed more. *Personality and Social Psychology Bulletin*, *41*(5), 643–658.

Goodman, N. D., Baker, C. L., & Tenenbaum, J. B. (2009). Cause and intent: Social reasoning in causal learning. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society.*

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday life. *Psychological Science*, *17*(9), 767–773.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*(8), 589–604.

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The naive utility calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, *123*, 101334.

Kirfel, L., & Lagnado, D. (2021). Causal judgments about atypical actions are influenced by agents' epistemic states. *Cognition*, *212*, 104721.

Lagnado, D. A., & Channon, S. (2008). Judgments of casue and blame: The effects of intentionality and foreseeability. *Cognition*, *108*, 754–770.

Malle, B. F. (2006). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. MIT Press.

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.