Integrating emotional expressions with utterances in pragmatic inference

Yang Wu (yangwu1@stanford.edu) [1,a], Michael Henry Tessler (tessler@mit.edu)[1,b], Mika Asaba (masaba@stanford.edu)[a], Peter Zhu (pzhu12@jhu.edu)[c], Hyowon Gweon (hyo@stanford.edu)[a], & Michael C. Frank (mcf@stanford.edu)[a]

[a]Department of Psychology, Stanford University, Stanford, CA 94305 USA

[b]Department of Brain and Cognitive Science, MIT, Cambridge, MA 02139 USA

[c]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA

Abstract

Human communication involves far more than words; speakers' utterances are often accompanied by various kinds of emotional expressions. How do listeners represent and integrate these distinct sources of information to make communicative inferences? We first show that people, as listeners, integrate both verbal and emotional information when inferring true states of the world and others' communicative goals, and then present computational models that formalize these inferences by considering different ways in which these signals might be generated. Results suggest that while listeners understand that utterances and emotional expressions are generated by a balance of speakers' informational and social goals, they additionally consider the possibility that emotional expressions are noncommunicative signals that directly reflect the speaker's internal states. These results are consistent with the predictions of a probabilistic model that integrates goal inferences with linguistic and emotional signals, moving us towards a more complete formal theory of human communicative reasoning.

Keywords: politeness; communicative goals; emotional expressions; affective cognition; computational modeling

Introduction

Human communication is inherently multimodal. We draw communicative inferences based not only on what others say but also what nonverbal expressions they make, such as their emotional expressions. While utterances and emotional expressions are generally aligned (e.g., someone smiles as she praises your cooking), we often experience cases where they provide conflicting cues. For instance, suppose that your friend Anne tastes a bit of a cookie you made. While she says, "It tastes good," you also see signs of disgust on her face. What is Anne trying to communicate, and what does her behavior tell you about how your cookie tastes? Despite Anne's explicitly positive comment, her expression of disgust may have powerful impact on your inferences: Perhaps your cookie isn't as good, and she was just trying to be nice. Decades of research have focused on people's pragmatic inferences from linguistic information (e.g., Grice, 1975; Goodman & Frank, 2016), but we still understand little about how listeners might integrate emotional expressions to interpret speakers' behaviors. In this study, we examine how people integrate speakers' emotional expressions with their utterances in communicative reasoning.

Language has at least two roles in human communication. First, it provides a medium for transmitting information about the world (Shannon, 1948; Grice, 1975). Second, it also serves various social functions (Ide, 1989; Brown et al., 1987; Goffman, 1967). For instance, in socially sensitive contexts, speakers may choose utterances that maintain their own or others' positive self-image or reputation rather than utterances that convey accurate information (Brown et al., 1987). Recent computational models, known as Rational Speech Act (RSA) models (Goodman & Frank, 2016), have also provided a unifying framework, characterizing speech acts as reflecting a balance of informational and social goals (Yoon et al., 2020).

Yet, modality of communication is not just constrained to language. As illustrated by the example above, linguistic utterances are often accompanied by a range of nonverbal signals. Among various kinds of nonverbal signals (e.g., pointing, gestures, touch; Kita, 2003; Bohn & Frank, 2019; Tomasello, 2010; Gweon, 2019; Hertenstein et al., 2009), one particularly rich source of information is our emotional expressions[2]: we smile, frown, or widen our eyes, and we exclaim "wow," "ohh," or "aww" when talking. These expressions change dynamically in real time and are pervasive in face-to-face conversations. What role do these emotional expressions play in human communication?

Humans draw rich inferences from perceived emotional expressions (Wu et al., in press). Beyond using others' expressions to infer their internal affective states (see Cowen et al. 2019; Barrett et al. 2019 for review), even young children readily use others' emotional expressions to recover unknown aspects of the world: e.g., inferring that something is yucky from a disgusted emotional response and something is cool from an excited vocal burst (Wu et al., 2017; Egyed et al., 2013). By reasoning about what others know, want, and think, older children can also differentiate between emotional expressions that are genuine and emotional expressions that are displayed for social concerns (e.g., making others feel good; Wu & Schulz, 2020). Recent advances in computational cognitive science provide an integrative framework to formalize these inferences; by assuming that people have an intuitive theory of how true states of the world and others' internal mental states give rise to their emotional expressions (Ong et al., 2015; Wu et al., 2018; Saxe & Houlihan, 2017), these models can capture how people use others' emotional

[1]These authors contributed equally to this work.

[2]We use "emotional expressions," "emotional displays," or "emotional signals" to refer to the facial, vocal, and bodily features that are commonly associated with emotions.

expressions to infer hidden world states (Ong et al., 2015) and others' unobservable mental states, such as their beliefs and desires (Wu et al., 2018).

These recent computational and empirical advances suggest that emotional expressions may play similar roles as language in human communication. Like linguistic information, emotional expressions can be considered communicative signals that either convey information about the world or serve social functions. Just as how people can infer the meaning of utterances by reasoning about the speakers' mental states, the same can be done with emotional expressions.

However, there might be important differences between language and emotional expressions. While speech is almost always a communicative act, an emotional response to states of the world may not always be communicative; rather, they might sometimes reflect non-communicative, automatic manifestations of their internal states. Indeed, a longstanding theoretical debate in emotion research concerns how emotional expressions are generated; while some theories argue that emotional expressions are automatic displays of emotions governed by the autonomous nervous system (e.g., Darwin, 1965; Ekman & Friesen, 1971; Izard, 1994), others emphasize people's strategic use of affective displays as communicative devices (e.g., Kraut & Johnston, 1979; Janney & Arndt, 1992; Shariff & Tracy, 2011).

Both the potential similarities and differences between language and emotion pose interesting questions about human communicative reasoning. When both sources of information are available, how do listeners make pragmatic inferences? Using Bayesian models of pragmatic inference to formalize the integration process (see also Bohn et al., 2019, 2020), our current study investigates how people combine others' emotional expressions with their utterances in communicative reasoning. The study has three specific goals. First, we ask if we can replicate recent findings (Yoon et al., 2016, 2020) that listeners interpret utterances in socially sensitive contexts as satisfying a balance of informational and social goals. Second, we ask whether there is a role of the speaker's emotional expressions in listeners' utterance interpretation and communicative reasoning. If so, our third goal is to investigate how listeners incorporate emotional expressions: do they consider emotional expressions as spontaneous reactions to the external world, or as communicative acts? If the latter, do listeners interpret those emotional expressions as conveying an informational goal, a social goal, or a balance of the two?

Behavioral Experiment

We conducted an experiment in which participants inferred either the likely state of the world or a speaker's communicative goals based on a speaker's utterance and emotional expression. Experimental details, including sample size, exclusion criteria, design and analyses, are pre-registered[3].

[3]Links to pre-registrations, raw data, analyses, and models can be found at `https://github.com/yang-wu-github/emo-rsa`

Method

Participants We recruited 120 participants living in the United States via Prolific ($n = 60$ received true-state inference questions; $n = 60$ received communicative-goal inference questions). Following our predetermined exclusion criterion, we excluded 5 participants who answered our attention-check questions incorrectly.

Stimuli and procedure We selected four sets of emotional expressions. Each set includes a neutral, happy, and disgusted emotional expression from the same actor (see Figure 1). Three of the four sets were from the IASLab Face Set[4] and one was from the IMPA-FACES3D database: `http://app.visgraf.impa.br/database/faces/`.

Figure 1: Emotional expression stimuli

Participants read stories in which a person made food and another person tasted it. Then the taster's utterance ("It tasted [good/bad]."') and facial expression (happy/disgusted) were revealed. See Figure 2. The two possible utterances and emotional expressions created four distinct conditions, and each participant received one trial for each condition. Then, one group of participants received the true-state question and the other group received the communicative-goal questions. For the true-state question, participants were asked the food's taste (e.g., "How do you think Casey's salad tasted from 1 to 6 stars?"). Participants responded by filling in 1 to 6 stars on a scale. See Figure 2A. For the communicative-goal questions, participants were asked two questions: one about the speaker's informational goal ("How likely do you think Alex's goal was to provide accurate feedback?") and the other about the speaker's social goal ("How likely do you think Alex's goal was to be nice?"). They answered on a four-point scale ranging from "Not At All Likely" to "Extremely Likely". See Figure 2B. The experiment for the true-state question can be viewed at `https://tinyurl.com/`

[4]Development of the Interdisciplinary Affective Science Laboratory (IASLab) Face Set was supported by the National Institutes of Health Director's Pioneer Award (DP1OD003312) to Lisa Feldman Barrett.

y2ccb5ag and the experiment for communicative-goal questions can be viewed at: https://tinyurl.com/y6y58cy3.

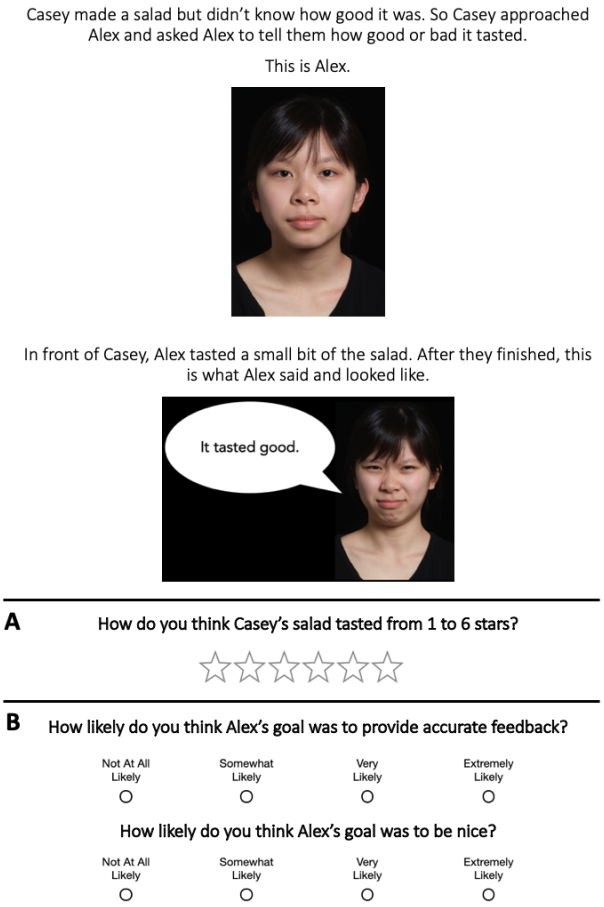

Figure 2: Example experimental trials for (A) true-state inference and (B) communicative-goal inference. Participants read identical cover stories for both true state and communicative goal inference trials.

Four tasting scenarios (in which Alex, Sam, Avery, and Taylor tasted a salad, soup, pasta, and a sandwich, respectively) were used. The mapping between conditions and scenarios was counterbalanced across participants and the order of conditions was randomized within participants. An attention-check question was included at the end of each trial.

Results and discussion

Data are shown in Figure 3. Although our primary goal was to understand variation in our measures using cognitive models, we also analyzed the data directly using Bayesian mixed-effects models (brms package in R; Bürkner 2017).

True-state inference As shown in Figure 3A (first column), both utterance (β = 18.39, 95% CI [5.28, 35.17]) and emotional expression (β = 29.64 [14.90, 49.76]) contributed to participants' state predictions, but there was no interaction between the two (β = 4.40 [-8.69, 18.17]). These results suggest that participants considered both cues to infer true states.

Communicative-goal inference As shown in Figure 3B, (first column), both utterance (β = -22.81 [-44.38, -8.69]) and emotional expression (β = -9.82 [0.65, 22.75]) contributed to participants' inferences of informational goal, with an utterance-by-emotion interaction (β = 68.70 [33.44, 124.33]). For social goals, however, only utterance contributed to participants' inferences (β = 29.52 [14.68, 51.42]); there was neither an effect of emotional expression (β = 6.75 [-2.56, 18.22]) nor an utterance-by-emotion interaction (β = -10.39 [-28.34, 3.99]). These results are consistent with prior findings that people consider both informational and social goals when interpreting others' utterances (Yoon et al., 2020), and motivate our hypothesis that people also consider these goals when interpreting others' emotional expressions. To explore this idea, below we formalize listeners' inferences based on verbal utterances and emotional expressions and compare these behavioral data against model predictions.

Computational Models

To better understand how listeners combine language and emotions in pragmatic inference, we formalize a space of hypotheses about how emotional expressions are integrated with linguistic utterances, defining a series of seven primary models that instantiate increasingly complex relationships between utterances and emotional expressions. Our models are couched in the Rational Speech Act modeling framework (Frank & Goodman, 2012; Goodman & Frank, 2016), which views language understanding as a recursive, social reasoning process whereby a listener reasons about why a speaker chose to say what they said. Specifically, we build on the model of Yoon, Tessler et al. (2020) that formalizes how social goals—in particular, politeness considerations—impact language understanding. We extend this model by considering a space of hypotheses about the communicative goals that emotional expressions and linguistic utterances aim to achieve.

Background: Politeness in rational communication

Our behavioral task involves interpreting a speaker's utterance and emotional expression to make state and goal inferences. We first define a pragmatic listener L_1 who interprets an utterance u by combining their prior expectations about the state of the world $P(x)$ and a speaker's goals $P(\vec{\phi})$ with their internal model of how the speaker came to produce the utterance $S_1(u \mid x, \vec{\phi})$.

$$L_1(x, \vec{\phi} \mid u) \propto S_1(u \mid x, \vec{\phi}) \cdot P(x) \cdot P(\vec{\phi}) \tag{1}$$

The speaker chooses an utterance based on its utility $U_{utt}(u; x; \vec{\phi})$, which depends on the state of the world x about which the speaker is trying to communicate as well as the speaker's goals $\vec{\phi}$. The speaker selects utterances soft-max optimally (with softmax parameter α) according to this utility function.

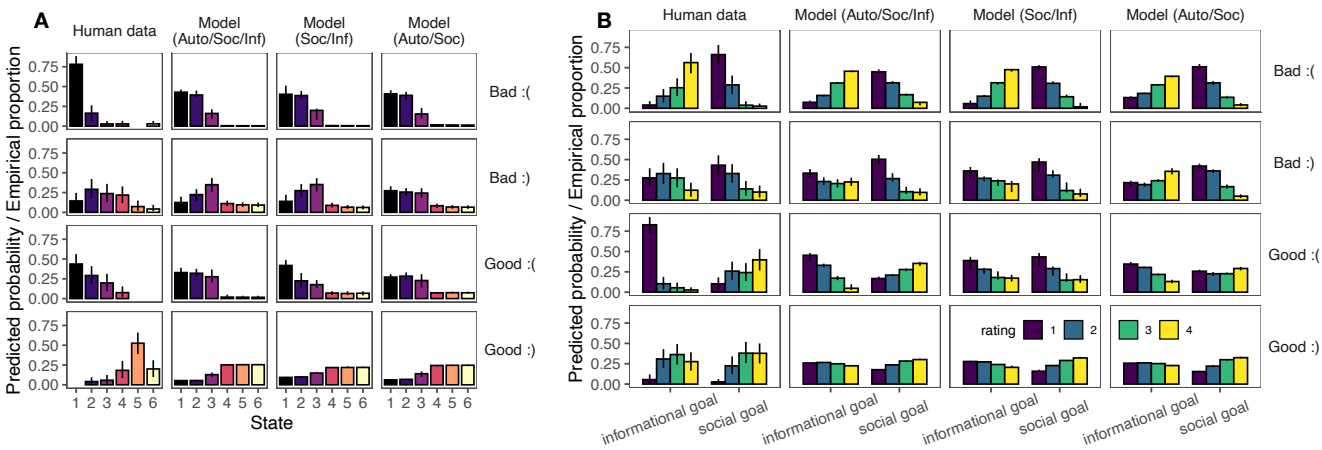

Figure 3: Results of behavioral experiments and model predictions for three of the seven models. A: Empirical and predicted distributions on states inferred by participants in the state inference task. B: Empirical and predicted distributions on likert scale ratings given by participants in the goal inference task. Error bars denote bootstrapped 95% confidence intervals for the data and 95% Bayesian credible intervals for the model. Emotional expression is denoted by emoticons :) and :(

$$S_1(u \mid x, \vec{\phi}) \propto \exp\left[\alpha \cdot U_{utt}(u; x; \vec{\phi})\right] \tag{2}$$

In the model of Yoon, Tessler et al. (2020), the utility of an utterance has informational and social components. Informational utility is defined by the negative surprisal of the utterance u for state x under a model of literal interpretation L_0 (i.e., a listener who interprets utterances according to their literal meaning; defined in Eq. 3 below): $U_{utt}^{inf}(x;u) = \ln L_0(x \mid u)$ (Goodman & Stuhlmüller, 2013). Social utility is defined as the expected subjective value V of the state implied by the utterance, under the model of literal interpretation L_0: $U_{utt}^{soc}(u) = \mathbb{E}_{L_0(x|u)}[V(x)]$. Yoon, Tessler et al. (2020) define the speaker's overall utility of an utterance as a weighted combination of the social and informational utilities: weight parameters $\vec{\phi} = [\phi_{inf}, \phi_{soc}]$ modulate the contribution of the two sources of utility: $U_{utt}(u; x; \vec{\phi}) = \phi_{inf} \cdot U_{utt}^{inf}(x;u) + \phi_{soc} \cdot U_{utt}^{soc}(u)$.

The model of literal interpretation L_0 is of an agent who updates its prior beliefs about a world state x through the literal meaning of an utterance u, encoded in the lexicon $\mathcal{L}_{utt}$:

$$L_0(x \mid u) \propto \mathcal{L}_{utt}(x,u) \cdot P(x) \tag{3}$$

The lexicon $\mathcal{L}_{utt}(x,u)$ is a mapping from utterances and states of the world to truth values (i.e., whether or not the utterance is literally true in a state). Rather than assume categorical truth values, we empirically measure the degree to which participants believe an utterance is literally true in a state (described in the Model Evaluation section below).

Integrating emotional expressions into pragmatics

We extend the model of Yoon, Tessler et al. (2020) by assuming that emotional expressions are generated by speakers according to some utility function U_{emo}, and we consider different hypotheses about that function. Specifically, we consider three possibilities about how emotions are generated (as well as combinations thereof): emotional expressions could be (a) automatic expressions that reflect the speaker's genuine feelings (*Auto*), (b) communicative acts chosen intentionally to convey information to the listener about the speaker's genuine feelings (*Inf*), or (c) communicative acts chosen intentionally to make the listener feel good (*Soc*). The automatic generation of emotional expressions (*Auto*) has a utility function that operates like a "literal speaker" model, where a truth-value function $\mathcal{L}_{emo}$ defines the utility of the emotional expressions e for state x.

$$U_{emo}^{auto}(e;x) = \ln \mathcal{L}_{emo}(e,x) \tag{4}$$

Like the lexicon function $\mathcal{L}_{utt}(x,u)$, the emotion lexicon $\mathcal{L}_{emo}(x,e)$ is measured empirically (described in the Model Evaluation section below). The utilities of emotional expressions that are generated as communicative acts (*Inf* or *Soc*) mirror those for the utterance utilities for the original Politeness model, where $L_0(x \mid e)$ (Eq. 3) describes a literal listener who updates its prior beliefs based on the true expression of the emotion $\mathcal{L}_{emo}$:

$$U_{emo}^{inf}(e;x) = \ln L_0(x \mid e) \tag{5}$$

$$U_{emo}^{soc}(e) = \mathbb{E}_{L_0(x|e)}[V(x)] \tag{6}$$

Note that these three possibilities are not mutually exclusive; thus, in addition to the three simple hypotheses about the utility of emotions, we formulate more complex hypotheses by considering their combinations. The simplest combination follows the politeness utility function, where the utility of an emotional expression is a weighted combination of informational and social goals (Soc/Inf):[5].

$$U_{emo}^{both}(e;x;\vec{\phi}) = \phi_{inf} \cdot U_{emo}^{inf}(e;x) + \phi_{soc} \cdot U_{emo}^{soc}(e) \tag{7}$$

[5]While, in principle, utterances and emotions could be used to differentially satisfy informational vs. social goals, for simplicity we assume that the speaker uses both utterances and emotional expres-

Next, we consider a space of possibilities wherein the listener is uncertain about whether the emotional expression produced by the speaker was intentional or not. These models enrich the pragmatic listener function by allowing them to reason about speakers with different utility functions.

$$L_1(x, \vec{\phi}, \chi \mid u) \propto S_1(u \mid x, \vec{\phi}, \chi) \cdot P(x) \cdot P(\vec{\phi}) \cdot P(\chi) \quad (8)$$

In this set of models, χ is a Boolean random variable that gates between different utility functions (e.g., U^{auto} and U^{both}).[6] We consider three such models in this vein, where the listener is uncertain whether or not the emotional expression was an automatic expression or a communicative one, where the communicative utility is either social (Auto/Soc; Eq.6), informational (Auto/Inf; Eq.5), or a combination (Auto/Soc/Inf; Eq.7).

We assume the prior distribution over world states $P(x)$ is uniform over the discrete state space $[1, 2, ..., 6]$, representing the star-rating given in the experiment. The subjective value V of a state here is its numeric value on the 1-to-6 star scale (i.e., V is the identity function). The pragmatic listener's priors over goal weights are independent, uniform distributions: $P(\phi_{inf}) = \text{Unif}(0,1)$, $P(\phi_{soc}) = \text{Unif}(0,1)$, and for models with uncertainty about the automatic vs. communicative status of emotions: $P(\chi) = \text{Bern}(0.5)$.

Model Evaluation

We evaluated our seven computational models based on their ability to predict state inference and goal inference data sets simultaneously.[7] To determine the literal meanings of utterances $\mathcal{L}_{utt}$ and emotions $\mathcal{L}_{emo}$ used in the literal listener equations, we performed a "literal judgments" task with an independent group of participants ($N = 60$), where participants evaluated if the utterance or emotion was literally compatible with the star-rating. We integrated all three data sources (literal semantics, state inference, goal inference) in a Bayesian data analysis model, where the global speaker optimality parameter α (Eq.2) is also fit (see Yoon, Tessler, et al. (2020) for a related analysis). We include two additional global parameters to the data analysis, in which we model the goal-weight scale as an ordinal, but not necessarily linear, scale (e.g., the psychological distance between goal ratings of 1 and 2 may be larger than the distance between 2 and 3).

We put uninformative priors over these parameters and inferred the model parameters and generated posterior predictions from each model by running 3 MCMC chains for 40,000 iterations, discarding the first 20,000 for burn-in. We additionally estimated the marginal likelihood of the data under each model (in order to calculate Bayes Factors) by running an Annealed Importance Sampling algorithm for 10,000 steps, collecting between 20-100 samples per model. All models were implemented in the probabilistic programming language WebPPL (Goodman & Stuhlmüller, 2014).

Model	r_{state}	r_{inf}	r_{soc}	$\log BF_{state}$	$\log BF_{all}$
Auto/Soc/Inf	0.78	**0.77**	0.88	–	–
Auto/Soc	0.76	0.57	**0.91**	-7.4	-20
Auto/Inf	0.75	0.72	0.77	-4.0	-20
Auto	0.79	0.69	0.79	-6.3	-19
Soc/Inf	0.75	0.73	0.70	-2.4	-27
Soc	0.39	0.26	0.75	-64	-117
Inf	**0.81**	0.73	0.77	-2.1	-20

Table 1: Summary statistics and model comparison for models varying in how emotional expressions are understood by the listener. r values are Pearson correlations; BFs are Bayes Factors (in log scale) quantifying evidence in support of a model in comparison to the *Auto/Soc/Inf* model (values less than 0 indicate support in favor of the *Auto/Soc/Inf* model).

We found that the most parsimonious explanation of the combination of data sets was the model of a listener who is uncertain about the generative process of the emotional expression they observed (i.e., whether it was generated intentionally by the speaker or whether it was an automatic expression conveying the true beliefs). Specifically, when the listener believed the emotional expression was intentionally produced, they assumed it was produced in order to satisfy a combination of informational and social goals (Auto/Soc/Inf; Eq.7; Table 1). This model performs 8x better (log BF = 2.1) than the next best model (*Inf*) when evaluated on the state inference data only and roughly 10^8x times better when evaluated on the full data set. The model also does the best job at capturing the key qualitative patterns in the data set, most notably in the conditions where the utterance and emotion conflict (Fig. 3).

General Discussion

How do language users integrate emotional expressions into the process of interpreting language? Combining behavioral experiments with computational modeling, we found evidence that emotional expressions were integrated rationally into language comprehension, supporting inferences about both the true state of the world and the speaker's communicative goals (i.e., being informative or social). These inferences were best captured by a model that formalized listeners' interpretations of utterances as satisfying a balance of informational and social goals, and emotional expressions as either emerging from the same goal balance or as an automatic manifestation of internal states. Notably, this model assumed that listeners can flexibly interpret whether a particular emotional expression is a communicative gesture (a smile to encourage

sions as a means to achieve the same (weighted combination of) goals.

[6] Although we start with this binary gating for simplicity, a more general formulation could consider a continuous weighting between intentional/communicative and unintentional/automatic production of emotional expressions.

[7] In addition to the models discussed in the main text, we formalized a parallel set of seven models that differ in that they treat utterances as purely informational (as opposed to satisfying both informational and social goals). As none of these seven models performed as well as the best-fitting models reported in the main text, and most were ranked lowest in the comparison, we omit these results for brevity.

a struggling colleague) or an unintended gaffe (an inadvertent look of disgust after eating a friend's cooking).

These results replicate recent findings that listeners consider speakers' informational and social goals in utterance interpretation (Yoon et al., 2016, 2020). They are also in line with a growing body of work showing that people use observed emotional expressions to draw rich inferences about the world and other people (e.g., Wu et al., 2017, 2018; Wu & Schulz, 2018, 2020; see Wu et al., in press for review). The current study goes beyond the prior work, however, by synthesizing these recent advances and providing a formal model of how we integrate linguistic and emotional cues in communicative reasoning.

One noteworthy finding is that our best fitting model is one in which there are similarities and differences in people's interpretation of linguistic and emotional signals. Importantly, while people believe both utterances and emotional expressions can reflect a balance of informational and social goals, they think emotional expressions can also be generated unintentionally by internal affective states. Paralleling the debate on whether emotions are *actually* communicative or reflective of internal affective states (e.g., Darwin 1965; Ekman & Friesen 1971; Shariff & Tracy 2011; Keltner et al. 2019; Cowen et al. 2021; Barrett et al. 2019), our work provides insights into laypeople's *intuitive theories* of how emotional expressions are generated: In the mind of the observer, emotional expressions can be either automatically generated by internal states or deliberately displayed for communicative purposes. By integrating the emotional expression with the broader context, the observer can flexibly infer the generative process of an emotional expression.

We formalized a large space of possible models concerning how emotional expressions enter into pragmatic inference. Each of these models is a complex set of interconnected parts, the workings of which we have not been able to do justice in this short paper. One of the core assumptions of the modeling approach here is that the informational contribution of an emotional expression grounds out in some representation analogous to a context-invariant, literal semantics of an utterance. While this assumption captures the similarities between linguistic utterances and emotional expressions, we acknowledge that the mapping from states of the world to emotional expressions can be complex and nuanced; thus, this "semantic" approach to emotional expressions may not be viable across all contexts. Indeed, the current work uses only two relatively simple emotional expressions (happiness/disgust) that are relatively predictive (more diagnostic) of world states and subjective values. Despite these limitations, we hope that this space of models provides a starting point for integrating more complex emotional expressions into theories of pragmatic inference; furthermore, pragmatics, in turn, can provide a window into how people read each other's emotions.

The current work raises a number of directions for future research. First, our current study tests people's inferences from only two categories of emotional expressions (i.e., happiness and disgust) in a well-constrained context (i.e., tasting food). Future work should explore the generalizability of these findings to other emotions and contexts. In particular, disgust has been proposed to be a basic, evolutionarily important emotion that keeps us away from poisons, parasites, and contaminants (Curtis et al., 2011; Chapman & Anderson, 2012). Would our findings (in particular, that people consider emotional expressions to be automatic responses in addition to being communicative acts) hold only within the domain of food-related disgust? Or do they reflect a more general, intuitive belief that all emotional expressions could be generated by either internal affective states or communicative intents? Second, whereas this study presents utterances and static emotional expressions simultaneously, temporal offsets and dynamic changes in emotional expressions might change the integration process. For instance, would people consider an immediate emotional response to an event to be more genuine or automatic, whereas a later emotional expression might be more likely to have a communicative (in particular, social) goal? Last, the current study focuses on two types of communicative goals: being informative and being nice; it remains an important task for future research to understand other goals, such as joking, being mean, or being sarcastic.

In sum, the current work synthesizes emotion understanding and language comprehension to investigate how people combine the two sources of information to make pragmatic inferences. When emotional expressions are available in addition to a speaker's utterance, people use both cues to draw flexible inferences about what the true state of the world is, and whether the speaker is trying to be informative or nice. By bridging emotion research, language studies, and formal modeling, the current work brings us closer to a more unified formal account of human communication.

References

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68.

Bohn, M., & Frank, M. C. (2019). The pervasive role of pragmatics in early language. *Annual Review of Developmental Psychology*, *1*, 223–249.

Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2019). Predicting pragmatic cue integration in adults' and children's inferences about novel word meanings. *PsyArXiv*. doi: 10.31234/osf.io/xma4f

Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2020). How young children integrate information sources to infer the meaning of words. *PsyArXiv*. doi: 10.31234/osf.io/2wgfb

Brown, P., Levinson, S. C., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge university press.

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, *80*(1), 1–28.

Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. *Annals of the New York Academy of Sciences*, *1251*(1), 62–76.

Cowen, A., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature*, *589*(7841), 251–257.

Cowen, A., Sauter, D., Tracy, J. L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychological Science in the Public Interest*, *20*(1), 69–90.

Curtis, V., De Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1563), 389–401.

Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press. (Original work published in 1872).

Egyed, K., Király, I., & Gergely, G. (2013, May). Communicating Shared Knowledge in Infancy. *Psychological Science*, *24*(7), 1348–1353.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, *17*(2), 124.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998.

Goffman, E. (1967). On face-work. *Interaction ritual*, 5–45.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, *20*(11), 818–829.

Goodman, N. D., & Stuhlmüller, A. (2013, January). Knowledge and implicature: modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*(1), 173–184.

Goodman, N. D., & Stuhlmüller, A. (2014). *The Design and Implementation of Probabilistic Programming Languages.* `http://dippl.org`. (Accessed: 2021-2-3)

Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.

Gweon, H. (2019). Understanding others to learn and help others learn: Inferences, evaluation, and communication in early childhood. *Varieties of Understanding*, 167–190.

Hertenstein, M. J., Holmes, R., McCullough, M., & Keltner, D. (2009). The communication of emotion via touch. *Emotion*, *9*(4), 566.

Ide, S. (1989). Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. *Multilingua*, *8*(2-3), 223–248.

Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research.

Janney, R. W., & Arndt, H. (1992). Intracultural tact versus intercultural tact. *Politeness in language: Studies in its history, theory and practice*, 21–41.

Keltner, D., Tracy, J. L., Sauter, D., & Cowen, A. (2019). What basic emotion theory really says for the twenty-first century study of emotion. *Journal of nonverbal behavior*, *43*(2), 195–201.

Kita, S. (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: an ethological approach. *Journal of personality and social psychology*, *37*(9), 1539.

Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, *143*(C), 141–162.

Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a Bayesian model of theory of mind. *Current Opinion in Psychology*, 1–21.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395–399.

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Wu, Y., Baker, C. L., Tenenbaum, J. B., & Schulz, L. E. (2018). Rational inference of beliefs and desires from emotional expressions. *Cognitive science*, *42*(3), 850–884.

Wu, Y., Muentener, P., & Schulz, L. E. (2017). One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(45), 11896–11901.

Wu, Y., Schulz, L., Frank, M. C., & Gweon, H. (in press). Emotion as information in early social learning. *Current Directions in Psychological Science*.

Wu, Y., & Schulz, L. E. (2018). Inferring beliefs and desires from emotional reactions to anticipated and observed events. *Child development*, *89*(2), 649–662.

Wu, Y., & Schulz, L. E. (2020). Understanding social display rules: Using one person's emotional expressions to infer the desires of another. *Child development*, *91*(5), 1786–1799.

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2016). Talking with tact: Polite language as a balance between kindness and informativity. In *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2771–2776).

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2020). Polite speech emerges from competing social goals. *Open Mind*, *4*, 71–87.