Interjections as Tools for Sharing Mental States

Amanda Royka (amanda.royka@yale.edu)[1], Kate Choi (kate.choi@yale.edu)[1], Rory Schoenberger (rory.schoenberger@yale.edu)[1], Julian Jara-Ettinger (julian.jara-ettinger@yale.edu)[1]

[1]Department of Psychology, Yale University

Abstract

Humans are intuitive mindreaders. We use our Theory of Mind to infer other people's mental states based on how they behave. Yet, humans are also motivated to ensure that others can infer their mental states easily and accurately. However, to act on this motivation, we must have tools to help others efficiently understand our minds, particularly when our behavior could be misunderstood. We propose that interjections—simple vocalizations like *oh*, *oops*, and *ew*—are an important (but previously neglected) set of linguistic devices designed to reveal our mental states quickly and efficiently. We provide initial evidence for this account, showing that people believe that interjections ought to be used as if they were designed to broadcast mental states and that people spontaneously produce these interjections significantly more often in the presence of an observer. Our work sheds light on how humans are not only proficient *mindreaders*, but may also be adept *mindsharers*.

Keywords: Theory of Mind; Communication; Interjections

Introduction

Much of our social cognition is built on a basic capacity to make sense of other people's behaviors. Were the agent's actions intentional? Did she know the true identity of the object that she was interacting with? Was the outcome of her actions consistent with her desires? This ability to represent and infer other people's mental states, known as a *Theory of Mind* (Gopnik & Meltzoff, 1994), enables us to cooperate (Tsio et al., 2021), communicate (Royka et al., 2022; Sidera et al., 2018), and draw powerful inferences from pedagogical interactions (Gweon, 2021). Consistent with this, decades of research on Theory of Mind have focused on understanding how people use each other's observable behavior to infer underlying mental states like beliefs, desires, knowledge, and intentions.

While mental state inference is undoubtedly an important part of our social cognition, this picture neglects a major feature of our mental life. Humans are not just "mindreaders," making inferences about the unobservable mental states of some disinterested third party: We are also motivated to make our own mental states understood. Indeed, in real life, we are just as often observers trying to infer other agents' mental states as we are agents with mental states being inferred.

Making our mental states understood is often critical, and can become a major source of social conflict when failures arise. Naturally, using language to explain our mental states is a powerful solution that we commonly employ, but such an approach is also limited because it demands extended time and attention from observers. This can be a particular challenge for interactions where we may need to reveal our mental states quickly, such as when we're trying to navigate a crowd (and want to reveal where we intend to go), when we commit a social faux pas (and want to reveal that it was an accident), or when we are mid-conversation (and want to reveal that we were surprised by something that was said).

Here we suggest that, beyond being mindreaders, people are also natural *mindsharers* with a suite of cognitive tools that help us to help others make sense of our behavior. That is, we propose that people have tools for supporting other people's mindreading.

What would a system for broadcasting mental states look like? Ideally, this system would consist of specialized markers that allow agents to broadcast corrections over the types of mental states that others might be inferring. These markers should convey different dimensions of mental life (i.e., different markers to signal ignorance, positive preferences, intentionality, etc.), and they should be able to be deployed quickly and flexibly in order to efficiently reveal our mental states during any activity with little interference.

We propose that humans do in fact have a system that meets these specifications, which we call *mentalistic interjections*. Interjections are conventionalized vocalizations such as *oops*, *ouch*, *woah*, *ugh,* and *ew*, which are often conceptualized as capturing a reaction to an event (Bloomfield, 1933; Quirk et al., 1972). Interjections also have properties that make them ideal for mindsharing. They are typically short and not subject to traditional grammatical rules (i.e., an interjection can be vocalized independent of the surrounding grammatical context; Huddleston & Pullum, 2002; Wierzbicka, 1992), enabling people to use them quickly and efficiently in any context.

While linguists have previously noted that interjections are connected to mental states (Ameka, 1992; Goddard, 2014; Wierzbicka, 1992) and past empirical work suggests that people react to interjections beginning in early childhood (e.g., Carpenter et al., 1998; Warneken & Tomasello, 2007; Wellman et al., 2000; Butler & Tomasello, 2016; Wu et al., 2017), it is unclear whether (1) people represent interjections as mindsharing devices, and (2) people actually use them strategically as such.

Here, we dive further into the dynamics of interjections to systematically test people's intuitions about the relationship between these communicative signals and the dimensions of mental life commonly tested within Theory of Mind research. First, we examined people's expectations regarding interjection usage in a third-party vignette task. If people expect interjections to be used to broadcast mental states, then participants' acceptability judgements should change as a function of the agent's mental states. Then, we conducted an in-person production task to test how observer presence effects interjection use.

Experiment 1

In Experiment 1, our goal was to characterize people's intuitions about interjection usage and thus the plausibility of interjections as mindsharing tools. As an initial test, we focused on three interjections that capture a range of mental states that people may need to broadcast: *oops* (marking intentionality), *oh* (marking knowledge), and *ew* (marking desire). To evaluate people's intuitions about the use of mentalistic interjections, we designed a set of short vignettes that parametrically varied the protagonist's knowledge, intention, and desire, as well as the presence or absence of an observer. For each scenario, we asked people to judge how natural it felt to use each interjection (tested across participants).

Methods

Participants 120 participants (*M*=31.41 years, range 18-78) from the US (as indicated by their IP addresses) were recruited through the Prolific research platform.

Stimuli Our stimuli consisted of 14 simple vignettes, all following the same general structure (Table 1), where an agent always performed an action and produced a certain outcome.

The space of vignettes was designed by varying four parameters: 1) whether the agent was alone or being observed (*absent* vs. *present*), 2) whether the agent used the salt shaker intentionally or knocked it onto their food accidentally (*intentional* vs. *unintentional*), 3) whether the agent was already aware that the salt shaker contained sugar (*ignorant* vs. *knowledgeable*), and 4) whether the outcome was consistent with the agent's desires (*desired* vs. *undesired*; see https://osf.io/jxmy9/?view_only=ba750204039d42398c2cdc1460679535 for the full set of vignettes).

The full parametric space consisted of 16 possible vignettes. However, this set included two logically inconsistent events where a knowledgeable person intentionally produces an outcome that they do not desire (both when an observer is present and absent). We therefore omitted the two vignettes, resulting in a final stimuli set of 14 vignettes. The protagonist in each vignette was then given a different name to avoid spillover effects across trials.

Procedure Participants were randomly assigned to one of three conditions, determined by the interjection they would evaluate (*oh*, *oops*, or *ew*). Participants first completed a brief tutorial which explained that they would read a series of stories and judge how natural the characters' responses were (which always consisted of the character vocalizing an interjection). Given the similarity between the stimuli, participants were told to read carefully and to try to picture each scene in their head.

Today, someone refilled the salt shaker with sugar. Imagine that Aaron is in the break room and Charlotte is watching him. Aaron does not know that the salt shaker contains sugar. He accidentally knocks the contents from the salt shaker into his lunch. To Aaron's surprise, his lunch tastes sweet. Aaron is glad that this happened.

Parameter	Version 1	Version 2
Observer State	Present: *Aaron is in the break room and Charlotte is watching him*	Absent: *Aaron is alone in the break room*
Knowledge State	Knowledgeable: *Aaron knows the salt shaker contains sugar…As Aaron expected*	Ignorant: *Aaron does not know that the salt shaker contains sugar…To Aaron's surprise*
Intentionality	Intentional: *He intentionally adds the contents from the salt shaker into his lunch*	Unintentional: *He accidentally knocks the contents from the salt shaker into his lunch*
Desirability	Desired: *Aaron is glad that this happened*	Absent: *Aaron isn't glad that this happened*

Table 1: The topmost box shows an example of a full vignette from our study. In this vignette, the *observer state* is present, the *knowledge state* is ignorant, the *intentionality* is unintentional, and the *desirability* is desired. The rest of the table shows both dimensions of the four parameters which were combined into all internally consistent permutations to create our set of 14 vignettes. Note that in each vignette in the actual study, the names were different, but we keep them consistent here for clarity.

Results and Discussion

If people expect interjections to be used to broadcast mental states, then participant judgements should change as a function of the agent's mental states and the interjection. We begin by describing the qualitative structure of the results and then turn towards a more formal quantitative analysis. Figure 1 depicts the overall structure of the data.

As Panel 1a shows, participants found the use of *oh* to be more natural when the agent was ignorant about the outcome (M = 5.46) compared to when the agent was knowledgeable (M = 3.64). Panel 1b shows how the use of *oops* produced a notably different pattern: Here, participants gave similar ratings as a function of ignorance (M = 3.86) and knowledge (M = 3.41), but they thought that the interjection was more natural when the action was unintentional (M = 4.32) compared to when it was intentional (M = 2.79). Finally,

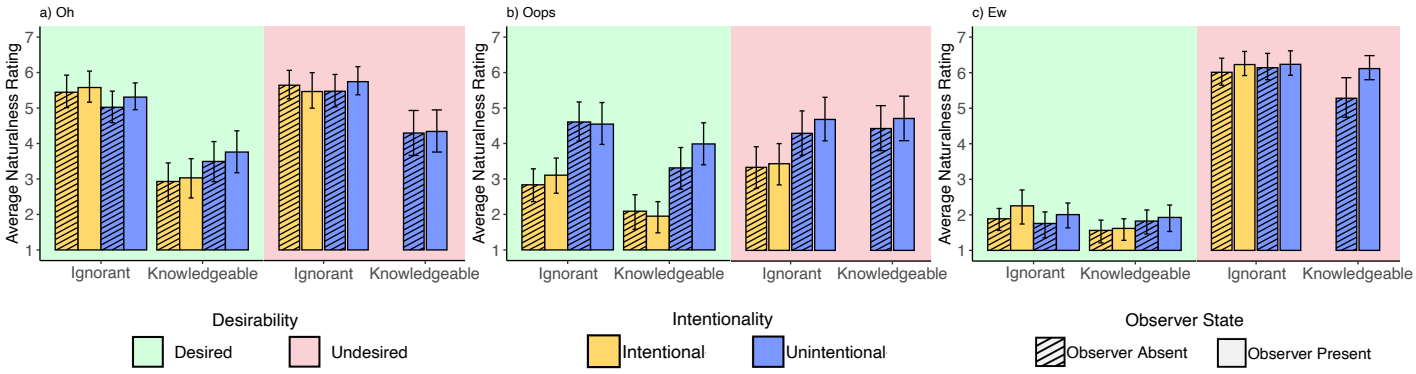

Figure 1: Full experiment results. Each bar shows the average naturalness ratings for interjection usage in each vignette, broken down by interjection: a) *oh*, b) *oops*, c) *ew*. Graphs with a green background represent vignettes where the final outcome was consistent with the agent's desires and graphs with a red background represent vignettes where the final outcome was inconsistent with the agent's desires (*desired* vs. *undesired*). The results are further broken down based on whether the agent knew that his action would bring about the final outcome (*ignorant* vs. *knowledgeable*). The color of the bars represents whether the agent performed the action accidentally (*intentional* vs. *unintentional*) and the pattern of the bars corresponds to whether the agent was alone or being observed (*observer absent* vs. *observer present*). Error bars show 95% bootstrapped confidence intervals.

Panel 1c shows how the use of *ew* produced a completely different pattern, such that participants found its use to be more natural in response to undesired actions (M = 6.00) compared to desired actions (M = 1.85).

We then analyzed each interjection using linear mixed-effects regressions predicting naturalness ratings as a function of the four story parameters as independent variables. Each model included the maximal random effects structure that allowed the model to converge (Barr et al., 2013). Regression results are shown in Figure 2.

***Oh* condition.** The random effects structure in the resulting maximal model for the *oh* condition included random slopes and intercepts for both knowledge state and desirability and random intercepts for participant. This regression revealed that participants found the use of *oh* to be more natural when the character was ignorant of the outcome (β = 1.77, p < .001), when the outcome was undesirable (β = 0.51, p < .05), and when their actions were unintentional (β = 0.21, p < .05; Figure 2a).

Intentionality, however, may also reflect an influence of knowledge state as well. People are sometimes unaware of having performed accidental actions. Thus, some participants may have interpreted the unintentional actions in our scenarios as a second type of ignorance (i.e., the agent was ignorant of performing the accidental action). This ambiguity may have motivated some people to expect the agent to say '*Oh*' regardless of whether the agent knew the abstract causal relationship between the outcome and the actions they unintentionally (and potentially unknowingly) took.

What do these effects reflect? One possibility is that these effects only reflect expectations about the content of the interjection, such that people expect *oh* to express information both about the speaker's knowledge state and also whether or not the outcome was in-line with the speaker's preferences and intentions. An alternative possibility is that these main effects reveal expectations about the content and usage of the interjection. That is, the effect of knowledge state and intentionality might reflect that people believe that *oh* encodes information related to ignorance, while the effect of undesirability might reflect that people believe that others are more likely to want to broadcast when something bad happens. Since our current study cannot differentiate between these two explanations, we return to this point in the discussion.

***Oops* condition.** For *oops*, the resulting maximal model for naturalness judgments included knowledge state, intentionality, desirability, and observer state with no additional random effects (Figure 2b). Participants gave significantly higher naturalness ratings when the actions were performed unintentionally (β = 1.55, p < .001), when the agent is ignorant (β = 0.62, p < .001), and when the outcome is inconsistent with the agent's desires (β = 0.48, p < .005; Figure 2b).

This pattern of effects may once again reflect participants' intuitions about both the mental state content of the interjection and the factors that may motivate a person to publicly mark an event. Crucially, *oops* can be interpreted as marking a lack of intention across both actions (e.g., when you accidentally hit a switch) and outcomes (e.g., you intentionally flip a switch that you don't know the function of and it causes all the power in your building to shut off).

In this way, one can say '*Oops*' either in response to an unintentional action or to an unintentional outcome – both are appropriate uses that may in actuality have more to do with intentionality than with knowledge state. Additionally, *oops* also tracks desire states, such that it is more natural in response to unwanted outcomes (e.g., it would be very unnatural to say '*Oops*' after accidentally dropping dice and they land on double sixes), which – like in the case of *oh* –

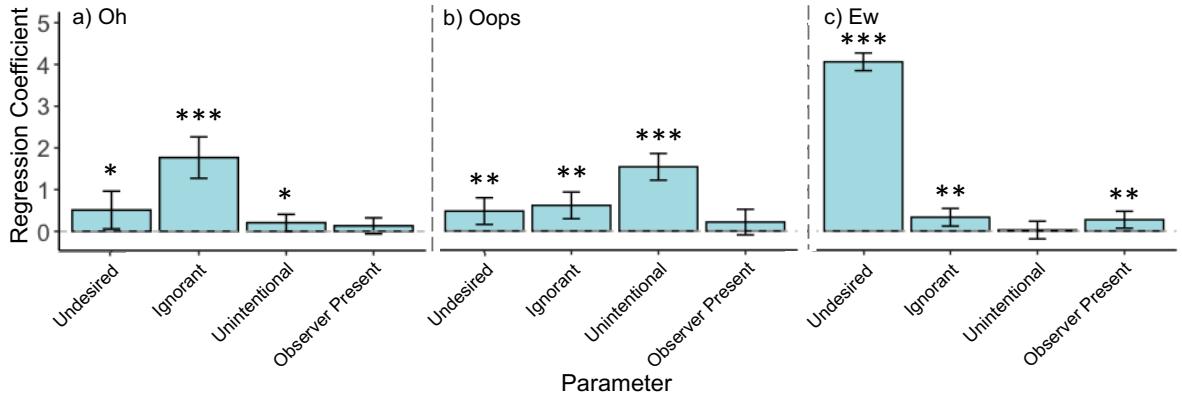

Figure 2: The bars show the regression coefficients for the four parameters (which are here referred on the x-axis by their reference level). a) *oh* was significantly predicted by desirability, knowledge state, and intentionality, b) *oops* was significantly predicted by desirability, knowledge state, and intentionality, and c) *ew* was significantly predicted by desirability, knowledge state, and observer state. Error bars show the 2.5% and 97.5% confidence intervals for the regression coefficients. * represents $p < .05$, ** represents $p < .01$, and *** represents $p < .001$.

may signal a more general intuition to broadcast an outcome that has a negative valence.

***Ew* condition.** For *ew*, participants' naturalness ratings were significantly predicted by three dimensions (Figure 2c). First, participants found the use of *ew* to be more natural when the outcome was undesired ($\beta = 4.09$, $p < .001$). This suggests that *ew* is used to reveal desires. Note that while *oh* and *oops* also had significant effects as a function of desire (possibly because agents are motivated to reveal that they did not know about or intend to produce the negative outcome), the strength of the effect for *ew* was orders of magnitude stronger (as can be seen in Figure 2).

Additionally, participants also rated *ew* as more natural when the agent was ignorant ($\beta = 0.33$, $p < .005$). This suggests that people expect *ew* to be more likely to be produced when the agent did not expect the outcome to occur; although note that this effect was also over 12 times weaker relative to desires. Finally, our regression also revealed a small effect of observer presence ($\beta = 0.27$, $p < .01$; Figure 2c). This is consistent with the idea that people might be more motivated to use interjections when there is an observer.

Role of observer status. Our results suggest that observer status affected people's intuitions only for the *ew* interjection, raising the question of why this was not the case for *oops* and *oh*. Interestingly, however, the effect size of observer status was comparable for *oh*, *oops*, and *ew* ($\beta = 0.13$, $\beta = 0.22$, and $\beta = 0.27$, respectively). If interjections broadcast mental state information to observers, then their use should be subject to audience effects. As such, the weak effect of observer status could provide evidence against the mindsharing view of interjections presented here.

However, one alternative explanation for this finding is that the production of mentalistic interjections may be contingent on the presence of an audience, but observers may interpret them as uncontrollable raw affective outbursts (similar to laughter or gasps; Goffman, 1981) that would be produced regardless of observer presence. Note, however, that this is not inconsistent with our account. It could, in fact, be a cognitively efficient way of interpreting interjections since assuming that everyone is more or less honestly broadcasting their reactions allows you to avoid having to track minds and figure out when the interjections are being used strategically.

Experiment 2

In our first study, we intentionally focused on people's intuitions about how interjections are used in order to assess whether people expect interjections to reveal mental states. Next, we shift our focus to the actual production of interjections.

Imagine that you are editing a paper on your computer as your advisor looks over your shoulder. In your haste, you accidentally make a typo. You realize your mistake and feel the urge to say "*oops*." Now consider, would you also produce that interjection if you were alone in your room? While Experiment 1 suggests that observers may treat mentalistic interjections as raw affective outbursts rather than intentional communicative signals, theorists have held both perspectives on interjection production (Ameka, 1992; Goffman, 1981; Scherer, 1994; Wharton, 2003). Thus, in Experiment 2, we directly tested the effects of observer presence on interjection production.

If interjections are meant to broadcast mental state information, then people should be more likely to use them when being observed. Given the difficulties inherent in reliably eliciting interjections, we chose to initially focus on one category of interjections that could be reliably evoked in a simple task. We specifically tested for audience effects on the production of interjections related to violations of intentionality (e.g., *oops*).

Methods

Participants 44 Yale undergraduates and New Haven community members (M=19.02 years, range 18-22) were

recruited and tested as pairs. All participants were recruited through the Yale Study Pool Participant system in exchange for either course credit or monetary compensation.

Stimuli In our task, participants played a modified version of the Simon Task (Simon & Wolf, 1963; Stoet, 2017). Each round began with a fixation cross centered on the screen. Then, the word "left" or "right" appeared either 200 pixels to the left or right of the fixation cross. Participants were instructed to press the "A" key (on the left side of the keyboard) when the word "left" appeared on the screen and to press the "L" key (on the right side of the keyboard) when the word "right" appeared on the screen. Originally, this task was designed to study the performance advantages conferred by the correspondence between the features of a stimulus (e.g., the word) and the required response (e.g., the side of the button press). We, however, chose this task because participants reliably make mistakes when completing it.

Our version of the Simon Task lasted for 30 rounds. If participants pressed the wrong key in response to a cue word, then then an error message was displayed for 1 second before the next round began. In order to increase the likelihood that participants would make mistakes, the experimenter instructed them to respond as quickly as possible.

Procedure After arriving at the lab, all participant pairs were told that they would alternate playing a series of games on a computer. Each pair was randomly assigned to one of two conditions. In the *together* condition, one participant completed the Simon Task in the presence of their partner (n=15 participants). In the *alone* condition, each of the two participants completed the Simon Task by themselves while their partner waited in another room (n=14 participants).

Data Coding The first author and one additional coder blind to our hypotheses independently reviewed recordings of participants completing the Simon Task. Each coder recorded 1) whether the participant completing the task produced any interjections at all and 2) how many interjections were produced in total. Agreement between the coders was high; both coders agreed on the number of interjections produced for 27 out of 29 sessions. In the two sessions where the coders disagreed, they disagreed about the production of a single interjection. In these two cases where there were disagreements, the blind coders' judgements were used.

Results and Discussion

If people use interjections to broadcast information to others, then they should produce interjections more often in the presence of an observer. Consistent with these predictions, participants in Experiment 2 were more likely to produce an interjection in the together condition relative to the alone condition ($\beta = 0.72$, $p < .001$; using a generalized linear model predicting whether or not an interjection was produced as a function of condition). Specifically, in the alone condition, only 2 out of 14 participants produced interjections while completing the task, whereas 13 out of 15 participants in the together condition produced at least one interjection (Figure 3). Crucially, however, there was no significant difference in the number of errors made during the task across the two conditions ($\beta = 0.68$, $p < .352$), suggesting that while participants made similar numbers of mistakes, only participants being observed marked those mistakes with interjections.

Taken together, Experiments 1 and 2 provide evidence of a gap between how people think interjections are produced (i.e., independent of observer presence) and their actual production characteristics. Although this initial investigation only provides evidence of this gap for interjections that are mistake-related (e.g., *oops*), it is possible that this gap exists for mentalistic interjections more broadly, regardless of the dimension of mental life that is being revealed. In future work, we plan to use more complex tasks to elicit a broader range of mentalistic interjections.

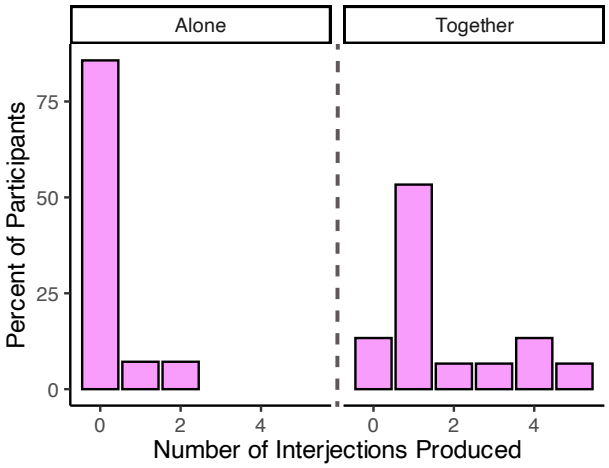

Figure 3: Each bar shows the percent of participants within each condition (y-axis) that produced a given number of interjections (x-axis) during the Simon Task.

General Discussion

Humans have to navigate a variety of complex social situations on a daily basis and our mutual understanding of each other's mental states forms the foundation that allows these interactions to proceed smoothly. Here we proposed that, because people are also motivated to make their minds understood, people may have rapid linguistic markers that are used to scaffold other people's interpretation of our behavior. Consistent with this, we found that mentalistic interjections satisfied the desiderata for such a system, and found initial evidence that they may be used as a tool for supporting the Theory of Mind inferences of others. People expect *oh*, *oops*, and *ew* to correspond to the mental states being experienced by the speaker and people are more likely to produce a mistake-related interjection in the presence of an observer.

In Experiment 1, we focused on three mentalistic interjections that intuitively track three main representations in Theory of Mind: *oh* for knowledge, *oops* for intention, and *ew* for desire. Our results suggest that people expect these

three interjections to be used to help reveal mental states that may be misunderstood. Specifically, *oh* most strongly expresses information about the speaker's knowledge state and is used to indicate that the speaker was ignorant of something. *Oops* marks a lack of intention that spans across both actions and outcomes. Finally, *ew* is treated as a marker of a mismatch between outcome and desire.

Here we focused on the isolated effects of each mental state, but it is possible that these interjections (or other interjections that we did not consider) also encode interactions between mental states. We focused on the main effects in part because our experiment was not well-equipped to find complex interactions: our full model would have 16 parameters and thus, model selection would be vulnerable to overfitting. While our current experiment does not have sufficient power to estimate such a large model, in future work, we hope to test more complex regressions with theoretically motivated interactions.

Critically, the naturalness judgements given in our task may tap into two distinct intuitions concerning interjection usage: 1) what mental state information does the interjection reveal and 2) what mental states motivate someone to reveal other mental states. For example, as we suggest earlier, the effect that negative desires have over all interjections may relate less to the type of information captured by a particular interjection and, instead, may capture a strong motivation to reveal that the negative outcome was unexpected, unwanted, or unintentional. It is possible that examining the interactions between parameters may reveal even more interesting patterns between interjections' mental state content and the motivation to reveal these mental states. For example, people may be more motivated to mark their ignorance with *oh* when the outcome was undesired. While our current study cannot differentiate between these two features that may motivate interjection usage, future work will incorporate additional manipulations to tease them apart.

Relatedly, the motivation to reveal that a negative outcome was unexpected, unwanted, or unintentional likely also impacted people's production of interjections in Experiment 2 where we found that people were more likely to produce mistake-related interjections ("*oops*", expletives, etc.) in the presence of an observer. Interestingly, participants produced these signals in a situation that was relatively low-stakes: the participant completed the task independently, the observer was not impacted by their score, and there was no reward for good performance. While further research is needed to explore why we produce these interjections (Do they mark our surprise at our own poor performance? Or perhaps our dissatisfaction with our errant response?), it is notable that, at least in this initial investigation, interjections are deployed as communicative signals even outside of joint activities that directly involve or effect our observer.

In future studies, we plan to further test the mindsharing characteristics of mentalistic interjections. For example, is the production of mentalistic interjections affected by how easy it would be for an observer to infer a given mental state? If it would be difficult for an observer to tell whether you accidentally or intentionally knocked over a pile of books, would you be more likely to say *oops* compared to when the action was obviously unintentional? This would not only provide additional evidence that interjections convey mental state information, but it would also provide further support for their role as tools to support others' mental state inferences as well.

So how many mentalistic interjections exist? In Experiment 1, we focused on three interjections as an initial case study, but the full space of this sub-category remains an open question. Relatedly, it is unclear the extent to which there is stability cross-culturally in the types of mental states that interjections express. The answers to these questions will give further insight to the potentially universal social inferential challenges that our communication systems have evolved tools to solve.

Taken together, our results show that mentalistic interjections are treated as markers of mental state information. Thus, people can use these interjections to refine their inferences about the knowledge, intentions, and desires of others. While our focus here is on understanding human social cognition, our work also makes a contribution to linguistic analyses. Our work highlights a sub-category of interjections that may be best understood as functionally distinct from other interjections. This grouping diverges from previous taxonomies that categorize interjections as cognitive (expressing information relating to the speaker's epistemic state; e.g., *woah*), emotive (expressing emotions, like irritation or disgust; e.g., *yuck*), or volitional (expressing a command; e.g., *shh*; Wierzbicka, 1992). Importantly, our work does not imply that these previous categorizations were wrong, but rather, our findings point towards different categorizations based on interjections' functional roles with relation to Theory of Mind and social cognition more broadly.

Mentalistic interjections are quick, but powerful expressions that reveal our mental states. However, we believe that they constitute just one instance of what is actually a suite of cognitive tools that help us to help others make sense of our behavior. In this way, Theory of Mind can be understood as an interactive process in which mental state inference does not solely rely on behaviors performed by a disinterested actor. Instead, people navigate their social world in a way that makes their own minds more legible to others.

Acknowledgements

We thank the members of the Computational Social Cognition Lab for helpful conversations. This work was supported by NSF award IIS-2106690.

References

Ameka, F. (1992). Interjections: The universal yet neglected part of speech. *Journal of pragmatics*, *18*(2-3), 101-118.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis

testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.
Bloomfield, L. (1933). *Language*. New York: Holt.
Butler, L. P., & Tomasello, M. (2016). Two-and 3-year-olds integrate linguistic and pedagogical cues in guiding inductive generalization and exploration. *Journal of Experimental Child Psychology*, 145, 64-78.
Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant behavior and development*, *21*(2), 315-330.
Goddard, C. (2014). Interjections and emotion (with special reference to "surprise" and "disgust"). *Emotion Review*, *6*(1), 53-63.
Goffman, E. (1981). Forms of talk. Oxford, UK: Blackwell.
Gopnik, A., & Meltzoff, A. N. (1994). Minds, bodies, and persons: Young children's understanding of the self and others as reflected in imitation and theory of mind research.
Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, *25*(10), 896-910.
Huddleston, R., & Pullum, G. K. (2002). The Cambridge grammar of English. Cambridge, UK: Cambridge University Press.
Quirk, R., Greenbaum, S., Leech, G. N., & Svartvik, J. (1972). A grammar of contemporary English.
Royka, A., Schouwstra, M., Kirby, S., & Jara-Ettinger, J. (2021). I Know You Know I'm Signaling: Novel gestures are designed to guide observers' inferences about communicative goals. *In CogSci Proceedings*.
Scherer, K. R. (1994). Affect bursts. In S. H. M. van Goozen, N. E. van de Poll, & J. A. Sergeant (Eds.), *Emotions: Essays on emotion theory* (pp. 161–193). Hillsdale, NJ: Lawrence Erlbaum.
Sidera, F., Perpiñà, G., Serrano, J., & Rostan, C. (2018). Why is theory of mind important for referential communication?. *Current Psychology*, *37*(1), 82-97.
Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics*, 6(1), 99-105.
Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31.
Tsoi, L., Hamlin, J. K., Waytz, A., Baron, A. S., & Young, L. L. (2021). A cooperation advantage for theory of mind in children and adults. *Social Cognition*, *39*(1), 19-40.
Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11(3), 271-294.
Wellman, H. M., Phillips, A. T., & Rodriguez, T. (2000). Young children's understanding of perception, desire, and emotion. *Child development*, 71(4), 895-912.
Wharton, T. (2003). Interjections, language, and the 'showing/saying' continuum. *Pragmatics & Cognition*, *11*(1), 39-91.
Wierzbicka, A. (1992). The semantics of interjection. *Journal of pragmatics*, *18*(2-3), 159-192.
Wu, Y., Muentener, P., & Schulz, L. E. (2017). One-to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *PNAS*, 114(45), 11896-11901.