

# Dogwhistles, Trust and Ideology

Robert Henderson and Elin McCready

## Abstract

Given that someone is consistently untruthful, why should we ever trust them? The question is not academic. Consider politicians and others who are known to consistently lie, but who are still voted back into office. This talk addresses this puzzle via three mechanisms: (i) a theory of source evaluation based on interactional histories and heuristics for judgments of reliability (McCready, 2015), (ii) a game-theoretic view of how speaker ideologies and political positions are communicated by linguistic acts (Burnett, 2018; Henderson and McCready, 2018), and (iii) a theory of how ideological considerations are valued alongside truth-conditional content. In the process, an analysis of *fake news* claims is provided.

## 1 Opening

In politics (and beyond), it is common to see people trusting others who are known to consistently not care much about the truth.<sup>1</sup> Donald Trump is the most prominent case. Many have devoted a great deal of time to exposing the false claims he routinely makes in his speeches and comments, but his supporters seem to trust him nonetheless. On standard views of the evaluation of information sources in epistemology, which usually involve the tracking of truth, this is a surprise. Given that someone is consistently untruthful, why should we ever trust them? This paper addresses this puzzle via three mechanisms. First is a theory of source evaluation based on interactional histories and heuristics for judgments of reliability (McCready, 2015), which is used to spell out the puzzle and as an exemplar of truth-based views on trust. The second is a game-theoretic analysis of how speaker ideologies and political positions are communicated by linguistic acts (Burnett, 2018), and an extension of that analysis to dogwhistles in political speech (Henderson and McCready, 2018). This latter contains the seeds of the third mechanism, one new to this paper: a theory of how ideological considerations are valued alongside truth-conditional content. We claim that such considerations are what makes people trust testimonial agents like Trump; in the process, an analysis of *fake news* claims is provided. We close with several conclusions, first the descriptive claim that trust in Trump-like agents can be rational depending on one's preferences, and second a set of prescriptions for how political discourse needs to proceed given the foundations of ideologically based trust.

## 2 Evaluating Information Sources

How can one determine whether content obtained from a particular agent, or other information source, should be believed? This is a longstanding and controversial issue in epistemology, and has been extensively addressed in thousands of works; our goal here is not to present a solution to this question, but simply to make it more precise. In order to do so, we make use of the work of McCready (2015), who proposes a two-factor theory of source reliability for both agent testimony and other evidence sources (as applied to dynamic update with evidential constructions in formal semantics/pragmatics). Here we consider only the case of testimony, as the more general theory is not required to clarify the main question we are interested in: the relationship between reliability/truthfulness and trust.

---

<sup>1</sup>Acknowledgments: Audiences at Expressing Evidence and LENLS17, especially Regine Eckardt, Natasha Korotkova, Mari-bel Romero, Stephanie Solt.

On the view of [McCready \(2015\)](#), an agent first uses a set of heuristics to determine whether a particular testimonial source has properties associated with reliability. Sample properties that might be taken into account by these heuristics include the source’s profession, personal presentation, gender, race, etc. Of course, some of these properties are genuinely useful for determination of reliability and some are not. For example, an agent’s profession is relevant to deciding whether that agent is going to be reliably truth-tracking for assertions about content domains relevant to that profession: we might believe a linguist on issues of linguistic practice, for example. However, some choices of property lead to *epistemic injustice* ([Fricker, 2007](#)), such as gender and race, which aren’t in general relevant to evaluation of reliability, but are nonetheless taken into account by (some) agents in making judgements about whether to believe a piece of testimony (see [McCready and Winterstein 2017, 2019](#) for a detailed investigation of the case of gender, which shows via experimental methods that at least for some cases men are assigned higher baseline degrees of reliability than women). Still, application of these heuristics lead to a probability that the source is reliable. This is an initial determinant of whether it is a rational move to accept content proffered by the source.

The initial probability assigned using the heuristics is only a starting point which is subsequently modified by interaction with that source. Each observation of a discourse move by the source, together with verification of whether its content is truth-tracking, alters, via conditionalization, the probability the source is reliable. In particular, following [McCready 2015](#), suppose that each information-transmitting discourse move by a source of testimony contributes to a history of discourse moves, understood as iterations of a repeated game of information transmission. Such moves are modeled, simplifying slightly from the original system, as tuples  $\langle \varphi, V \rangle$ , where  $V$  is a value in the set  $\{T, F, ?\}$  indicating the truth value of the proffered content ( $?$  indicating that the move cannot be evaluated or that its value is otherwise indeterminate). The basic probability of reliability across this discourse sequence is then interpreted as the frequency of  $T$ -valued moves, so the result is a real number in  $[0,1]$  where each  $T$ -valued move induces an uptick in value (unless the frequency is already 1). The result is that each truth-tracking move raises (possibly very slightly) the perceived likelihood that the source is reliable, and each move that fails to track truth lowers it.

But how should reliability relate to belief? That is, how should an agent incorporate information from a variety of sources of varying reliability? Furthermore, when should an agent be willing to say they believe a proposition based on information from these various sources. Once we understand how reliability affects belief, we can model the fact that, on par, people discount unreliable sources, and begin to address the puzzle that some unreliable sources continue to be believed.

We link reliability and belief using a flavor of dynamic semantics ([Groenendijk and Stokhof, 1991](#)), following work on dynamic updates in Plausibility Models (e.g., [Baltag and Smets 2008](#)). In this literature, slightly simplified, a frame  $\sigma$  is a set of worlds ordered with a ‘plausibility ranking’ reflecting epistemic preferences on states, which is enriched into models in the usual way. In Reliability Dynamic Logic ([McCready, 2015](#)) the situation is slightly complicated. While we have a global model  $\sigma$ , we care more about indexed submodels (hereafter *information states*), each of which represents a source whose reliability we mean to track.

- (1)  $j \in \text{Source} \cup \mathcal{A}$ , where **Source** is the set of evidence sources and  $\mathcal{A}$  the set of agents.

These information states are ordered by the total ordering  $\leq_a$  satisfying (2) for agent  $a$  who is tracking reliability.

- (2)  $i <_a j$  iff  $P(\text{Rel}(i)) < P(\text{Rel}(j))$ , where  $P(\text{Rel}(i))$  is the probability that source  $i$  yields reliable information according to  $a$ .

Updates on information states are of the following form, where the subscript  $i$  marks the source of the evidence of  $\varphi$ . For cases of testimony, as we discuss here,  $i$  ranges over  $\mathcal{A}$ , the set of agents.

$$(3) \quad \sigma [E_i\varphi] = \sigma' \text{ where, for all } \sigma_j \in \sigma, \begin{cases} \sigma'_j = \sigma_j[\varphi] & \text{if } i = j \\ \sigma'_j = \sigma_j & \text{if } i \neq j \end{cases}$$

Note that a sentence  $E_i\varphi$  always induces the standard update of state  $\sigma_i$ , namely  $\sigma_i[\varphi] = \{s \in \sigma_i | s \in \varphi\}$ . At this level, update with  $\varphi$  always takes place — but this is *not* the same as coming to believe  $\varphi$  at a global level.

To determine global beliefs, we unify the information of all tracked substates  $\sigma_i$  via lexicographic merge in (4).

$$(4) \quad \text{Lexicographic merge.} \\ R_{a \circledast b} := R_a^< \cup (R_a^= \cap R_b) = R_a^< \cup (R_a \cap R_b) = R_a \cap (R_a^< \cup R_b)$$

The core idea is that merging  $\sigma_i \circledast \sigma_j$  will result in state where all non-contradictory content survives, and in case of conflict, information from the higher-ranked source overrides the lower-ranked source— i.e.,  $i <_a j$  will privilege content from  $j$ . Recursively applied, the global state  $\sigma_T$  on which belief is defined will almost never exhibit conflicts.

With the link finally established between reliability (encoded in  $<_a$ ) and belief (propositions that follow from  $\sigma_T$ ), we are now in a position to state the puzzle. Each information-transmitting discourse move affects the perceived reliability of a source in a way dependent on the properties of the source and whether it is truth-tracking. Thus, an agent that is consistently non-truth-tracking will rank lower on  $<_a$  with each non-truth-tracking discourse move. As these agents fall on  $<_a$ -rank, the content this agent provides will be less likely to survive lexicographic merge and appear in  $\sigma_T$ . Thus, the content such an agent provides will not be believed, as belief is defined on the global information state. This seems right in the general case, but now the puzzle is clear: what about politicians who are known to consistently ignore the truth, but are still trusted? For the specific and obvious case, what about Trump? The remainder of this paper will address this puzzle, claiming in short that, in certain cases, ideology trumps reliability.<sup>2</sup>

### 3 Dogwhistles and Ideologies

Henderson and McCready (2017, 2018, 2019) build a novel theory of so-called *dogwhistle* communication, a communicative strategy familiar from the literature on political speech. This work is relevant to the question at hand because, as we will see, dogwhistle communication involves simultaneously sending truth-conditional information while at the same time sending information about the speaker's persona, which can include ideological orientation. The formal tools used to model dogwhistle communication can be, essentially, immediately extended to solve the puzzle developed in the previous section. The larger result is that we begin to see the emergence of a general theory of social meaning and how it interacts with truth-conditional meaning, one that can handle prima facie disparate linguistic phenomena like dogwhistles and trust.

Dogwhistles can be roughly defined as messages which communicate aspects of the speaker's ideology to an ingroup in a way which is not accessible to an outgroup. Consider this line from George Bush's 2003 State of the Union address; the term 'wonder-working power' is used in an evangelical Christian hymn, and thus has associations that will be recognized only if the interpreter has the requisite background.

$$(5) \quad \text{Yet there's power—wonder-working power—in the goodness and idealism and faith of the American people.}$$

---

<sup>2</sup>Painful pun very much intended.

What is the covert message sent to listeners who recognize the phrase ‘wonder-working power’? [Henderson and McCready \(2017, 2018, 2019\)](#) argue that it is only information about the speaker’s conversational persona, and possible broader identity, as an evangelical Christian. This is backed by experimental literature in political science (e.g., [Albertson 2015](#)), where it is shown that evangelical listeners are more likely (than non-religious listeners) to recognize the phrase, more likely to approve of religious rhetoric in religious appeals, and more likely to identify the speaker as a religious conservative politician.

These *identifying* dogwhistles can be contrasted with *enriching* dogwhistles. With the latter sort of expression, the hidden message sent to savvy listeners involves the enrichment of truth-conditional content. We see a clear example in the following example from a 2014 radio program, where Representative Paul Ryan said the following.

- (6) We have got this tailspin of culture, in our inner cities in particular, of men not working and just generations of men not even thinking about working or learning the value and the culture of work.

He was criticized shortly after by fellow Representative Barbara Lee for making a “thinly veiled racial attack”. This is because the phrase *inner-city* is code or euphemism for African American neighborhoods (especially stereotypically racialized views of such neighborhoods). We see, then, that this kind of dogwhistle does not only send information about the speaker’s persona, but involves altering the truth-conditional content of the message. That is, one is committed to radically different propositional content when saying there is “a tailspin of culture, in our inner-cities”, versus the enriched message that there is “a tailspin of culture, in our inner-city, African American neighborhoods.”

The fact that dogwhistles have these two components—(i) identifying a speaker persona, and (ii) enriching truth-conditional meaning—motivates the account developed in ([Henderson and McCready, 2017, 2018, 2019](#)). We argue that identifying dogwhistles, which only do (i), are a kind of special case of enriching dogwhistle, which involve both (i) and (ii). In particular, we propose that recovering the enriched meaning of the dogwhistle is recovered *in virtue of* recognizing the persona of the speaker.

Given that the persona-signaling aspect of dogwhistles is core, we use the *social meaning games* of Burnett (2017) as the foundation for our analysis of dogwhistles. Such games involve utility assignment based on recovery of speaker persona/ideology and hearer evaluation of that persona/ideology. In this previous work, we modified Burnett’s original definition of ‘social utility’ for use with dogwhistles in (7).<sup>3</sup> Here  $P_L$  gives the probability the speaker has persona  $p$  given the observed message  $m$ , which is just the RSA-style sociolinguistic literal listener (e.g., [Goodman and Frank 2016](#); [Franke and Jäger 2016](#); [Franke and Degen 2016](#)), while  $v_S$  and  $v_L$  (borrowed from [Yoon et al. 2016](#)), give the speaker and listener’s affective values of various personas weighted by the likelihood the speaker will be assigned that persona given their message.

$$(7) \quad U_S^{So}(m, L) = \sum_{p \in [m]} P_L(p|m) + v_S(p)P_L(p|m) + v_L(p)P_L(p|m)$$

Utility as in (7) is enough to model identifying dogwhistles. Speakers pick messages in the context of a listener (or group of listeners) weighing what persona they want to have and how listeners will judge that persona weighted by how likely it is that message will signal that persona. Dogwhistles arise when the speaker wants a persona  $p$ , listeners who don’t like  $p$  don’t realize that the  $m$  signals  $p$ , and listeners that do like  $p$  realize that  $m$  signals  $p$ .

To extend the account to handle enriching dogwhistles, which involve truth-conditional meaning, the account must allow for speakers and listeners to extract information from two dimensions. The core idea is to assuming signals with two possible meanings, one an enriched version of the other. Then,

<sup>3</sup>Slightly simplified. We, for instance, are not using log probabilities here because we don’t consider the issue of bounded rationality, which can modeled with a temperature parameter.

we let recovery of the enriched version be tied to recognition of the relevant persona. Thus, messages now denote pairs of truth-conditional meanings and social meanings:  $\langle [[m]], [m] \rangle$ . Given  $T$ , a set of states  $t$  (worlds), speaker strategies  $S_\sigma$  are now functions from pairs of states and personas to messages, and listener strategies  $L_\rho$  are functions from messages to such pairs. We then have the following utility function for information retrieval.

$$(8) \quad \begin{aligned} US(m, L) &= US_{\text{soc}}(m, L) + EU(m, L), \text{ where} \\ EU(m, L) &= \sum_{t \in T} Pr'(t|p) \times U(t, m, L), \text{ where} \\ U(t, m, L) &> 0 \text{ if } t \in L_\rho(m) \text{ and else } = 0 \text{ (cf. van Rooij (2008)).} \end{aligned}$$

There are three critical aspects. First, the top line gives the general formula for the utility of a message given a listener—it is computed as a linear function of its social utility and its truth-conditional utility. This is critical for the puzzle at hand about truth and reliability, and we return to this in the following section. Second, note that in computing the utility of a message for signaling a state, we consider the (posterior) probability of the state given the speaker's persona— $Pr'(t|p)$ . That is, we compute the speaker's persona and then use that information to determine the probability that some meaning was transmitted given that the speaker has the persona they do. For most types of truth conditional meaning  $Pr'(t|p) = Pr'(t)$ . That is, we gain no information from knowing the speaker's persona. When using enriching dogwhistles, though, the speaker's persona is critical for catching the enriched meaning. Once we identify Paul Ryan's persona, we are more likely to interpret *inner-city* as *African American neighborhoods*, rather than merely neighborhoods in the city. Finally, the third line (8) asserts that there is a payoff in the truth-conditional domain just in case the listener picks the correct state of the world. The bigger picture is that the social meaning is always recovered, but if the listener fails to recover the proper truth-conditional meaning, no value is extracted from this aspect of the communication.

With (8) we have a unified account of both identifying and enriching dogwhistles. The next section shows how a consequence of this analysis provides a solution to the puzzle we began with. That is, how is it that certain speakers are able to consistently speak falsely, while not being punished by certain listeners in terms of trust? The core insight that the account of dogwhistles provides is that messages package both social/ideological messages alongside truth-conditional content. This will allow a listener to track the utility of a speaker's message along either dimension, which we will refer to as reliability and trust respectively.

## 4 Ideology and Trust

The analysis of dogwhistles and of social personas in the previous section (and also in Burnett 2018) assumes that we have a way to assign affective values to personas, as cast in the function  $v_{S/L}$ . But on what basis? Agents can have various reasons to assign positive (or negative) value to a given persona. There are obviously many options, including relatively high-level criteria, for instance the degree to which the persona instantiates some value independently held by the agent such as originality, rebelliousness, safety, or conformity to some social norm. Here, though, we will consider only a very simple metric, namely *similarity*: 'I like people who are like me,' or, in this context, 'I like people who have social personas which are like mine.' The intuition behind this criterion is that there is intrinsic value in discovering other individuals who share our core values, ideologies, and personal styles. Assuming this criterion means we can assign affective values on the basis of similarity metrics between speaker and hearer personas.

With even this simple metric we can enrich the analysis of reliability in McCready 2015, which was predicated on truth-tracking alone. The core idea is to let hearers determine trust through a mix of truth-conditional and social meaning, with different mixes leading to different kinds of listener utilities.

Fleshed out, speaker payoffs rest on two fundamental aspects of meaning: first, for truth-conditional meaning, the successful transmission of true information about the world, and, second, for social meaning, successfully transmitting information about speaker personas together with the hearer's evaluation of that persona. We treat hearer payoffs identically because (for truth-conditional meaning) clearly the hearer cares about learning the truth, and (for social meaning) as social agents, we value finding individuals who share our values and social groupings, as already indicated in the previous paragraph.<sup>4</sup>

We follow Henderson and McCready (2018) and weight the two components of the utilities with values  $\delta, \gamma$ , giving (where  $U_S^{soc}(m, L)$  is the utility of the social meaning and  $EU(L, Pr)$  is the value of the truth-conditional meaning, following van Rooij 2008):

$$U_S(m, L) = \delta U_S^{soc}(m, L) + \gamma EU(L, Pr)$$

This  $\delta$  indexes the value placed on the social meaning, while  $\gamma$  indexes the value of the truth-conditional meaning. Setting  $\delta = 0$  gives a style of communication where social meaning is disregarded, for instance “science” as traditionally construed, i.e. as a completely objective enterprise where social aspects of the agents (scientists) involved are irrelevant. At the other extreme, setting  $\gamma = 0$  gives “post-truth”, a style of communication in which facts are irrelevant and only social persona matters, on the assumption that relevance is dependent on utility assignments.

This mechanism can be exploited for an analysis of reliability in the face of countervailing evidence as in the case of Donald Trump. When  $\delta > \gamma$ ,<sup>5</sup> we end up with Trump-voter-style confidence and trust, because considerations of truth are vastly undervalued compared to persona signaling. Indeed, enriching the system of McCready (2015) with social meanings means that perceived reliability can increase in repeated game settings in virtue of social signaling alone, given that social meaning is allowed to play a role in the ‘vetting’ process.<sup>6</sup> This kind of strategy, while prima facie aberrant, can be rational in terms of maximizing utility in a variety of scenarios. One case would be in scenarios where social signals are more easily interpreted. In a political campaign where voters may not know much about certain policy domains, sending a true message about those policies may not actually help the listener pick the true state of the world. If these message are additionally uninformative in terms of social meaning, then it would be rational in terms of payoffs to pick an alternative message that is false, but which clearly sends a message about the speaker's persona / ideology. In the extreme case of this strategy, which we've been taking Trump to exemplify, over a conversational history interlocutors become ever more sure of each other's sociolinguistic personas and attendant ideologies, but do so by discounting truth-conditional considerations.

While  $\delta > \gamma$  discounts truth conditional information, note that this setting does not necessarily lead to perverse truth-conditional outcomes. Remember that on first interaction with a speaker, a listener must fix a prior for reliability based on, in part, the speaker's persona (indicated by non-linguistic social cues). We expect that one viable strategy to evaluate a speaker's reliability is to, early in the conversation, set  $\delta > \gamma$  so to better alight on their persona and likely ideological commitments. Crucially, this can be truth-conditionally virtuous. Recall, for instance, that certain dogwhistles only get their enriched meaning once a listener has identified the persona of the speaker. This means that social meaning can

<sup>4</sup>Stephanie Solt (p.c.) points out to us that this assumption is too simple: it can be useful for hearers to learn speaker personas even when they do not share them, for instance when the persona in question is highly objectionable or when it might otherwise impact judgements about the reliability of the speaker in a truth-conditional sense. The interaction between social and truth-conditional aspects of utility is complex and this paper is not the place to explore it; we will do this in future work.

<sup>5</sup>Where ‘>’ indicates a vast difference in value.

<sup>6</sup>Space considerations preclude spelling this additional mechanism out in the present paper; but the basic idea is to allow, instead of simple truth-tracking as a vetting mechanism for discourse moves, either checking whether accurate information about social persona is transferred or whether the total utility of the discourse move exceeds a certain level, i.e. the communication is sufficiently useful that it's worth keeping on paying attention to the communicating agent. These two views give subtly different results which we plan to explore further in Henderson and McCready (2019).

affect the truth conditions of messages, and so having an accurate picture of the speaker is crucial. Focusing social meaning by setting  $\delta > \gamma$  early, and then switching to  $\gamma > \delta$  in order to monitor the truth-tracking of an agent whose persona/ideology has been sussed out is thus a viable strategy to maximize payoffs in both communicative domains.

We have shown how focusing on the social meaning domain is a viable strategy for maintaining high payoffs in repeated communication, and thus can be used to establish trust. We now turn to two communicative moves that can be used by speakers who are employing this strategy. The first, signals that, indeed they are pursuing this strategy. The second kneecaps alternative strategies.

In discussing Trump's communication style, multiple pundits and partisans have said that we must "take Trump seriously, not literally". What could this mean? In our account, the meaning is clear. It is an exhortation to set  $\delta > \gamma$ , or even to set  $\gamma = 0$ . The idea is that we should not pay attention to the literal truth-conditional content of his utterances, but to instead take him seriously. But take him seriously as what? We think it means to take his social persona as a billionaire, as a businessman, as a fighter, as a nationalist, etc. seriously. This rhetorical move immediately makes sense in the framework we have developed, and shows that social meaning and truth conditional meaning is separable, and that we can prioritize one or the other.

Moreover, with this setup, we are in a position to give an analysis of the *fake news* effect. In some conservative/right-wing discourse, taking a statement and calling it *fake news* is a signal to remove it from truth-relevant consideration. We see several different ways to model this within the present setup. The first is to view it as a call to set  $\gamma = 0$  in its payoff evaluation. The interpreter then considers only similarity of persona (ie. politics) in the utility calculation. With this, if the fake news move is accepted, considerations of truth become completely (though temporarily) irrelevant to political discourse. This is initially plausible, but it has several shortcomings. First, since it eliminates any utility stemming from the recovery of truth-conditional information, it leads to generally lower utilities for 'fake news' content; while this may be part of the intended effect, it is not clear to us that this way of computing utility properly tracks the way in which we value such content. Even if something is deemed to be fake news, there is a sense in which the truth-conditional content still provides some value to the interpreter, specifically that she comes to believe that the content is *false*, which is already useful in a way that should be reflected in utility calculation.

We thus think that *fake news* is best viewed as a kind of hedge which functions as a denial operator. Simultaneously, it has a pragmatic effect as an invitation to the hearer to ascribe a persona in ideological opposition to whatever agent made the original claim. Given the way in which we tie opposition and thus negative affect to lack of similarity, the ascription of low similarity will result in a negatively viewed persona. We use the following definition.<sup>7</sup>

$$(9) \quad FN(\varphi) = \begin{array}{l} (i) \text{ Deny}(\varphi) \\ (ii) \neg Sim(P_s, P_h) \end{array}$$

The utility of this operator can be motivated in various ways, but here is our preferred take. Imagine you're playing a Trump strategy in which the goal is to gain favor via the use of social meaning rather than being a reliable communicator in a truth-conditional sense. You would want to tell people that other news is fake so that they can't play a strategy which maximizes truth-conditional meaning: the explicit denial of a bit of content means that one won't be able to extract utility from it on the truth-conditional side of the utility calculation, so it will only be sensible to try to extract value from the social meaning side. As a result, the hearer (who accepts this discourse move) has to play the social meaning strategy-maximizing game you're playing, which is your preferred outcome, whether because you are better at the social meaning side in general, or because it is better for you to avoid strategies which pay attention

<sup>7</sup>As with the work of McCready (2015), the precise way in which hedges behave compositionally is left underdetermined here.

to truth-conditional content in the particular case in question.

## 5 Conclusions

Trust is distinct from reliability. We can trust the unreliable, and not trust the reliable. Moreover, contrary to first impressions, this is even a rational way to behave, if we value social matching over truth—which we might in the case of political speech, if we care most about what kind of policies a politician might want to implement. This insight leads to a major takeaway. If we want to re-imbue politics with truth, pointing out the falsehoods of politicians is not a productive method in general. Instead, our only way forward is to show that the ideological presentation of those politicians is insincere, or that the ideology that they uphold is itself flawed.

## References

- Bethany L Albertson. Dog-whistle politics: Multivocal communication and religious appeals. *Political Behavior*, 37(1):3–26, 2015.
- Alexandru Baltag and Sonja Smets. A qualitative theory of dynamic belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory*, number 3 in Texts in Logic and Games, pages 13–60. Amsterdam University Press, 2008.
- Heather Burnett. Signalling games, sociolinguistic variation and the construction of style. *Linguistics and Philosophy*, 2018.
- Michael Franke and Judith Degen. Reasoning in reference games: Individual-vs. population-level probabilistic modeling. *PloS one*, 11(5):e0154854, 2016.
- Michael Franke and Gerhard Jäger. Probabilistic pragmatics, or why bayes??? rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft*, 35(1):3–44, 2016.
- Miranda Fricker. *Epistemic Injustice*. Oxford University Press, 2007.
- Noah D Goodman and Michael C Frank. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11):818–829, 2016.
- Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14: 39–100, 1991.
- Robert Henderson and Elin McCready. How dogwhistles work. In *Proceedings of LENLS 14*. JSAI, 2017.
- Robert Henderson and Elin McCready. Dog-whistles and the at-issue/not-at-issue distinction. In Daniel Gutzmann and Katherine Turgay, editors, *Secondary Content*, pages 191–210. Brill, 2018.
- Robert Henderson and Elin McCready. *Signaling without Saying: The Semantics and Pragmatics of Dogwhistles*. 2019. To appear from Oxford University Press.
- E. McCready. *Reliability in Pragmatics*. Oxford University Press, 2015.



E. McCready and Grégoire Winterstein. Negotiating epistemic authority. In Kurahashi, Setsuya and Ohta, Yuiko and Arai, Sachiyo and Satoh, Ken and Bekki, Daisuke, editor, *New Frontiers in Artificial Intelligence (JSAI-isAI 2016 work-shops, LENLS, HAT-MASH, AI-Biz, JURISIN and SKL revised selected papers)*, volume LNAI 10247, pages 74–89. Springer, Berlin, 2017.

Elin McCready and Grégoire Winterstein. Testing epistemic injustice. *Investigationes Linguisticae*, XXXVII, 2019.

Robert van Rooij. Game theory and quantity implicatures. *Journal of Economic Methodology*, pages 261–274, 2008.

Erica J Yoon, Michael Henry Tessler, Noah D Goodman, and Michael C Frank. Talking with tact: Polite language as a balance between kindness and informativity. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2016.