5-2022

Cognitive Tribalism: A Social Doxastic Model

Robert Ragsdale
University of Arkansas, Fayetteville

Cognitive Tribalism: A Social Doxastic Model

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Philosophy

by

Robert Ragsdale
University of Arkansas
Bachelor of Arts in English, and Philosophy, 2014

May 2022
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____
Eric Funkhouser, Ph.D.
Thesis Director

_____
David Barrett, Ph.D.
Committee Member

_____
Barry Ward, Ph.D.
Committee Member

Abstract

How are facemasks – seemingly innocuous artifacts of the biomedical industry – currently embroiled in cultural wars? What motivates popular rejections of scientific consensus and messaging about the reality and consequences of anthropogenic climate change or the COVID-19 virus and vaccine? The puzzle is that (a) despite its being in everyone's rational interests to have a well-informed public and body politic about collective threats, and (b) despite the public availability of accurate and reliable information, scientific messaging and public discourse surrounding climate change, COVID-19, and vaccine hesitancy, nevertheless, tend to be hijacked by political interest. Yet, if belief is essentially truth-directed or truth-governed as many philosophers have supposed, then ameliorating anti-scientific attitudes should be a simple matter of explication: effectively communicating the relevant evidence and reasoning that supports the judgments about which there is scientific consensus I propose a solution to this challenge: anti-scientific beliefs gain traction *not* in spite of their perceived epistemic cost but *because of* it. This is at least one piece of the puzzle in understanding the reasons causally responsible for the proliferation of anti-scientific beliefs. Many popular rejections of scientific messaging are motivated by social – not epistemic – aims and enact social rather than navigational functions. My aim in this paper is twofold: (1) to advance a Disjunctive Signaling Model (DSM) of social beliefs and explore how such a model explains motivated rejections of scientific messaging, specifically the motivated rejection of messaging about Climate Change and COVID-19. And (2), to motivate positing the category of *tribal belief* into our psychological ontology. If DSM is accurate, then this motivates opening up the psychological space to allow for a subspecies of social beliefs, plausibly delineated by their selected-for proper functions, such as signaling socially strategic information, independently of any veridical functions, and which are neither constituted nor governed by a norm of truth. Motivating the addition of such a theoretical postulate is the aim of adequately describing socially and culturally oriented beliefs, which

function apart from and are valuable independently of any navigational or epistemic functions or utility.

Acknowledgements

Eric Funkhouser and Trip Glazer

Dedication

Kit-Kat and Edelweiss

Epigraph

It is wrong always, everywhere, and for anyone to believe anything on insufficient evidence" (Clifford 1877)

Suppose that [a man] is presented with evidence, unequivocal and undeniable evidence, that his belief is wrong: what will happen? The individual will frequently emerge, not only unshaken, but even more convinced of the truth of his beliefs than ever before (Festinger, Riecken, & Schachter 1957: 3).

Table of Contents

Introduction

How are facemasks – seemingly innocuous artifacts of the biomedical industry – currently embroiled in cultural wars? What motivates popular rejections of scientific consensus and messaging about the reality and consequences of anthropogenic climate change (Hornsey & Fielding 2017; Kahan et al. 2012, Kahan 2015, 2016, 2017; Hornsey et al. 2016) or the COVID-19 virus and vaccine (Shepherd, MacKendrick, & Mora 2020; Pivetti et al. 2021; Hornsey et al 2020)? Given that we all have an interest, *ceteris paribus*, in forming true climate and health beliefs – since climate change and COVID-19 present collective threats, irrespective of group membership[1] – we might reasonably expect a corollary motive for people to form true, epistemically rational beliefs based on the most accurate models available, irrespective of political affiliation or cultural worldviews. Or we would expect people at least to desire experts to accurately assess the reality of collective threats like climate change and COVID-19 for the sake of finding effective solutions (since, for most, their individual beliefs about climate change or COVID-19 are of little import, either to research or policy making).[2] In the face of collective crises, we all have "skin in the game" and are all ultimately vulnerable to bearing the costs. It is therefore in everyone's rational interest, in the face of commons-style threats, to get things right and strive to form true climate and health beliefs about climate change, COVID-19, and vaccine safety and efficacy.

The puzzle, however, is that (a) despite its being in everyone's rational interests to have a well-informed public and body politic about collective threats, and (b) despite the public availability of accurate and reliable information, scientific messaging and public discourse surrounding climate change, COVID-19, and vaccine hesitancy, nevertheless, tend to be hijacked by political interests,

[1] I am bracketing issues of environmental and medical inequality wherein those who are most vulnerable to the costs of climate change and global pandemics are people of color and lower socioeconomic status.
[2] I am sidestepping issues about how we determine who the relevant experts are, which is influenced by ingroup biases.

with scientific messaging often drowned out by the noise of rhetorical blustering. As a result, individuals often succumb to political and cultural influences – influences which neither aim at nor emphasize truth – in determining what to believe.

Unsurprisingly perhaps, research suggests that the average person's views about climate change or COVID-19 are explainable in terms of tribal sensibilities (Hornsey et al. 2016; Kahan 2015, 2016, 2017; Shepherd, MacKendrick, & Mora 2020; Pivetti et al. 2021; Greene 2013), depending on ingroup and outgroup messaging. Regardless of political ideology or affiliation, individuals are motivated to form beliefs that are congruent with their ingroup's messaging, even when these beliefs are about seemingly consequential issues (e.g., vaccines, climate change, facemasks, war).

Given their stability among the population, the epistemic and practical costs of socially motivated beliefs, therefore, must compensate by some other means, such as by securing access to the benefits of group membership and procuring other positive payoffs, leading to their adaptive stability. Given, for instance, the small cost of wearing a mask in public or getting a COVID-19 vaccine compared to the potentially high costs of failing to do so, what explains anti-masker attitudes, such as beliefs that facemasks are ineffective at mitigating the spread of COVID-19 or that they are dangerous to the wearer (Goodman & Carmichael 2020; Taylor & Asmundson 2021; Romer & Jamieson 2020)**?** More puzzling is the question: why is it that such beliefs, even while remaining to some extent maladaptive, gain significant traction among the public? A plausible answer is that the utility of such belief counteracts their potential maladaptiveness. Perhaps in certain social contexts, such anti-masker and anti-vaxxer beliefs are not maladaptive, but pro-adaptive.

Despite this live hypothesis, many in the media, the scientific community, politics, and the public health sector continue to express perplexity at the significant resistance to scientific and public health messaging from serious media outlets and from governmental and public institutions

(e.g., the CDC, WHO) about COVID-19 preventions, such as wearing facemasks, social distancing, and the safety and efficacy of COVID-19 vaccines, or the transmissibility and mortality rate of the COVID-19 virus. This perplexity is further compounded by the pervasiveness of misinformation despite the public having access to reliable information about the safety and efficacy of COVID-19 vaccines or the dangers of the virus.

The traction of anti-scientific attitudes among the general population presents a challenge not only to contemporary democracies but to traditional philosophical accounts of *belief* as well. Motivated rejections of scientific messaging (and anti-scientific beliefs generally) are, like delusions, self-deception, positive illusions, and conspiracy theories, characteristically arational. They similarly tend to be recalcitrant to counterevidence and reasoning, remaining highly entrenched and encapsulated, even when such rejections can have serious negative consequences for individuals.

Yet, if belief is essentially truth-directed or truth-governed as many philosophers have supposed (Velleman 2000; Shah 2002; Steglich-Petersen 2009; Gendler 2007; Engel 2013; Williams 1973), then ameliorating anti-scientific attitudes should be a simple matter of explication: effectively communicating the relevant evidence and reasoning that supports the judgments about which there is scientific consensus (Hornsey & Fielding 2017: 465). This is the ameliorative strategy the *principle of explication* recommends and which the *deficit model* of scientific knowledge and messaging presupposes. The problem, however, is that the principle of explication often fails to induce attitude change, proving ineffectual in contexts of motivated cognition, and even occasionally producing a *backfire effect,* leading to a rejection of the counterevidence and a strengthening of their confidence in the target belief (Nyhan & Reifler 2010: 307; Hornsey & Fielding 2017). Research on motivated rejections of scientific messaging and anti-scientific beliefs make it clear that the deficit model of knowledge and scientific communication fails as an empirical model and an ameliorative strategy in

the face of masse motivated rejections of scientific messaging (Hornsey & Fielding: 459; Simis et al. 2016).

The empirical inadequacy of the deficit model highlights the inadequacy of traditional philosophical models of belief. Specifically, traditional philosophical models are ill-equipped in accounting for or including beliefs that are psychologically motivated by social and/or personal reasons and which systematically deviate from epistemic or veridical norms. This elicits the question: are these models incomplete or fundamentally wrong? Consequently, traditional accounts face a challenge in explaining how either (a) anti-scientific attitudes are not genuine beliefs (but some other kind of proximal psychological attitude) *or* (b) anti-scientific attitudes are, after all, essentially governed or constituted by a truth norm or perform the exclusive proper function of aiming at truth.

I propose a solution to this challenge: anti-scientific beliefs gain traction *not* in spite of their perceived epistemic cost but *because of* it. This is at least one piece of the puzzle in understanding the reasons causally responsible for the proliferation of anti-scientific beliefs. Many popular rejections of scientific messaging are motivated by social – not epistemic – aims and enact social rather than navigational functions (Funkhouser 2017; Ramsey 1990). Paradoxically, I presume, many anti-maskers, anti-vaxxers, climate change deniers and others who display motivated rejections of scientific messaging, nevertheless accept scientific judgements regarding other quotidian facets of reality: weather forecasts, health-risks of ingesting lead, the risks of diabetes, the benefits of exercise, predictions of comets and eclipses, and the theory of DNA, the latter being presupposed by conspiracy theories surrounding COVID-19 and MMR vaccines. Despite their incredulity toward scientific consensus when it comes to protective masks, anthropogenic climate change, evolution, GMOs, etc., people who reject scientific consensus still place confidence in technologies advanced by scientific theories. This overwhelming acceptance of scientific messaging in daily life

circumstantially suggests that the rejection of specific scientific messaging programs, like those embroiled in public controversies and defining of cultural and political boundaries, are motivated attitudes, sustained by motivated reasoning. But motivated by what?

While anti-maskers and anti-vaxxers often posture as defenders of constitutional rights, or as scientific skeptics, data suggests that anti-mask attitudes, like other related COVID-19 attitudes, are tracking political affiliation and ideology – or at least they are strong positive correlates. An Axios-Ipsos Coronavirus poll (Bryan 2020), for instance, shows that the percentage of Democrats who always wore masks when leaving home rose from 49% to 65% from April to June 2020, whereas for Republicans, that number only increased from 29% to 35%. Despite the surrounding noise of "constitutional rights," a more plausible explanation is that rejections of scientific and public health messaging regarding social-distancing, facemasks, and the COVID-19 virus and vaccine, are in fact a means – a social opportunity – for individuals to enhance, maintain, or rehabilitate their social reputation relative to, in this case, their partisan social group, by signaling (via the rejection) both where their political allegiance lies and how strong that allegiance is, i.e., how committed they are to the partisan group (Funkhouser 2020), by indicating they are willing to stick with their cultural and political tribe despite the costs (though the costs may be perceived by ingroup members to be less than they objectively are (cf. Sosis 2003)). Such social aims, I propose, explain the motivated rejection of COVID-19 messaging and cultivate anti-COVID, anti-masker, and anti-vaxxer health beliefs.

Similarly, Kahan's (2016, 2017) claims that in polls, when participants either assert or deny beliefs in anthropogenic climate change or in evolution, they are attempting to express and protect their social identity – their commitment to a cultural worldview – rather than trying to express a scientifically accurate belief.

The rejection (or acceptance) of scientific messaging, which is culturally entangled, provide individuals an opportunity for signaling not only where their tribal allegiances lie but the strength of their commitment to that tribal allegiance, by displaying a willingness to incur certain kinds of costs, such as a cost to their social reputation's standing among the general public (Williams 2021), i.e., non-ingroup members, forgoing future opportunities to coordinate and cooperate with non-group members or the opportunity to migrate to other groups. They may also display a willingness to incur direct costs to their fitness such as risking health and well-being. Such characteristically social beliefs belong to a doxastic category which I denominate *tribal beliefs* – a determinate subspecies within the more general *belief* taxonomy.

My aim in this paper is twofold: (1) to advance a Disjunctive Signaling Model (DSM) of social beliefs and explore how such a model explains motivated rejections of scientific messaging, specifically the motivated rejection of messaging about Climate Change and COVID-19. To this extent, DSM explanatorily interfaces with many conspiracy theories insofar as they motivate the rejection of scientific messaging (including nanotechnology, GMO foods and other "chemophobia" beliefs – scientific issues about which the cultural and political Left tend to display socially motivated rejections (Hornsey & Fielding 2017; Beck 2017)). And (2), to motivate positing the category of *tribal belief* into our psychological ontology – a category which subsumes those beliefs explained by DSM. If DSM is accurate, then this motivates opening up the psychological space to allow for a subspecies of social beliefs, plausibly delineated by their selected-for proper functions, such as signaling socially strategic information, independently of any veridical functions, and which are neither constituted nor governed by a norm of truth.

Additionally, a sub-hypothesis of this paper is to argue that group ideology is often determined post-hoc, informed by the content of the group signal. Such a hypothesis runs counter to our usual way of conceiving of causal direction between prior commitments of groups and the

position they take on an issue, i.e., signal. Typically, however, the position a group takes on an issue is underdetermined by prior commitments or values of the group's ideological suite (Levy 2021a; Cohen 2003). Instead, the position is largely determined by which views or beliefs come to mark tribal affiliation. From the content of these belief signals, I claim, ideology is subsequently established as an after-the-fact means of rationalizing the position of the group.

Given the independent research tracts in social and cultural cognition – from *group polarization* to motivated climate and health beliefs – suggesting many of our evaluative and even factive beliefs are a function of what groups we socially-identify with (Kahan et al. 2010, Kahan et al. 2012; Kahan 2016, 2017; Hornsey & Fielding 2017), there is a clear need for developing a model or theoretical framework for discussing socially-oriented beliefs that are motivated by social aims and utility. The aim of this paper is to introduce a model of tribal beliefs. Under such a framework, tribal beliefs are a determinate species of the more generic intentional attitude *belief* and thus the tribal belief model postulates additional doxastic kinds into our psychological ontology. Motivating the addition of such a theoretical postulate is the aim of adequately describing socially and culturally oriented beliefs, which function apart from and are valuable independently of any navigational or epistemic functions or utility.

The applied value of such a framework is its relevance to important empirical questions addressing how pervasive tribal attitudes are among our general web of beliefs and how do we induce attitude-change for beliefs that are potentially, all else being equal, socially and personally deleterious. These questions inform policy-level decisions and bear on important descriptive and normative questions about democratic nations. The theoretical questions I take up here, however, are organized around unpacking the conceptual richness of the category, *tribal belief,* and attempting an inchoate theoretical outline of such a category. This involves seeking to explain why it would be adaptively valuable for the selection of tribal cognitivism in which social and epistemic beliefs

diverge in their normative structures. Why would distinctly social beliefs be selected for, despite their being epistemically and even pragmatically costly, and what are the mechanisms that make such attitudes possible?[3]

The following is divided into three sections. In section one, I survey the background of traditional doxastic theories, particularly Doxastic Veridicalism, against which the social model of belief is developed. In section two, I propose a Disjunctive Signaling Model of belief, where I define the two kinds of signaling and discuss their adaptive value and consider some objections Veridicalists might offer. Lastly, in section three, I briefly sketch a more general preview of tribal beliefs, drawing parallels with other kinds of arational beliefs, and developing a cost-benefit schema of social beliefs.

§1. Background: Doxastic Veridicalism

1.1 The Norm of Truth

Before laying out a theory of tribal beliefs, it will help to situate the discussion within the larger philosophical literature on belief to get a sense of the theoretical background against which the tribal belief model is developed. Motivating an account of tribal beliefs are the problems that arise from conflating *epistemic/veridical* beliefs with belief *simpliciter* – a conflation which is empirically problematic. In particular, the upshot of this conflation is the predominant supposition that a norm of truth is a necessary or essential feature of belief. The received wisdom is that belief is essentially a normative attitude and concept, with differing views as to what the normative content is specifically.

[3] For now, *social pressure* delineates two distinct loci of causal mechanisms: *endogenous* and *exogenous*. The former refers to the mechanisms of our mental architecture whereas the latter refers to external social forces, enculturation processes, and social pressures which effect conformity by means of norm-enforcement by individuals and institutions within a group (Matthews 2013: 109; see also Zawidzki's 2013 *Mindshaping*).

Typically, the normativity of belief is cashed out in terms of truth, whereby belief is essentially governed or constituted by a norm of truth (Boghossian 2008; Shah & Velleman 2005),

Truth Norm: S ought to believe *that p* only if *p*[4,5]
Truth-Criterion: belief *b* is correct if and only if *b* is true

Others, however, have construed the essential norms in terms of epistemic norms – evidential (Adler 2006) or knowledge norms (Engel 2004; Williamson 2000) – or in terms of rationality norms, given the essentially normative nature of rationality (Zangwill 2005).[6]

Since the most common construal of the normativity of belief is that it is essentially governed by a norm of truth, that is the view I will assume here, though the specific normative framework is ultimately irrelevant to the present project, which broadly resists any accounts that characterize belief as essentially and exclusively governed by epistemic, rationality, or veridicality norms.

Views which conceptually privilege the truth-criterion and hold that a norm of truth is a necessary or essential feature of belief fall within the more general theoretical tradition of *Doxastic Veridicalism*.[7] The Veridicalist tradition comprises two main theoretical strands which track different explanatory approaches as to why belief is governed by a truth-norm: *Normativism* and *Teleologism*. On the one hand, Normativists claim that a norm of truth is essential to or constitutive of belief as a basic or "brute fact" about belief and the kind of mental state it is, or at least is constitutive of our concept of belief (McHugh & Whiting 2014: 706; Shah 2002). A natural view is that the norm which

[4] As is often noted, the sufficient condition, if *that p* is true then S ought to believe *that p,* implies that we ought to know all true propositions, including trivial and useless propositions such as a true belief about how many blades of grass are in my yard (example borrowed from Papineau 2013). The problem is not only does it imply that we ought to believe truths that seem to have no real epistemic or pragmatic merit – that are simply not worth believing – but it also saddles us with an impossible task given the kind of creatures we are with limited epistemic access and cognitive capacities; thus, it also violates the maxim, "*ought* implies *can.*"
[5] Alternative formulations construe the norm in terms of rationality's essentiality to belief and the evidentiary conditions, e.g., S ought to believe P iff S has sufficient evidence P (Glüer & Wikforss 2013: 83; Engel 2013; Papineau 2013).
[6] See also Davidson (1973) and Dennett (1987) for *Interpretationist* accounts of belief which take the rationality constraint to be constitutive of belief, as a necessary constraint on our belief-attribution practices.
[7] I borrow the terms *Doxastic Veridicalism* and *Doxastic Absolutism* from Zalabardo (2010).

essentially governs belief is a *sui generis* doxastic norm, unique to the belief attitude and grounded in the nature of the attitude itself or in our concept of it – the former being a metaphysical claim and the latter being a conceptual one (Papineau 2013). It follows from the Normativist view then that a given psychological attitude, *x*, is a belief only if it is constituted or governed by a doxastic norm, i.e., a norm of truth. This is the *Normativism*[8] thesis (Engel 2013: Shah 2002).

On the other hand, the *Teleological* view claims that a psychological attitude *x* is a belief if and only if *x* aims at truth or instantiates the property of "truth-directedness," in the sense that belief-formation processes are regulated by this teleological constraint, i.e., "regulated for truth"(Velleman 2000: 277-8; Steglich-Petersen 2009, 2013, 2017; Gendler 2007; Williams 1973).[9] The precursor to the Teleological view is Bernard Williams, who famously claimed that belief, as a conceptual necessity, aims at truth (1973).[10] For the Teleologist, it is an aim – rather than a norm *per se* – that is essential to or constitutive of belief. Essential to individuating *belief* from other propositional attitudes is its "*telos* of truth*"* which aims at truth or the property of *being true* (Gendler 2007: 236). It is in virtue of this aim, which regulates how we form, update, revise, and reject beliefs, that a psychological attitude is uniquely a belief (McHugh & Whiting 2014: 703; Velleman 2000).

Importantly, as McHugh and Whiting (2014) note, while Normativism and the Teleologism are often "presented as a foil" to one another, they are not in fact incompatible but differ with respect to how they explain the normativity of belief:

> In fact, the claim that belief has an aim can be and often is drawn on to *explain* the truth of some version of Normativism: it is *because* belief has this aim that belief is essentially governed by the norms by which it is governed (McHugh & Whiting 2014: 706)

For Velleman (2000: 17) it is the fact that belief aims at truth "that generates this norm for its correctness." A virtue of the Teleological view is that it can offer a "deeper explanation"

[8] Referred to as "*Doxasticism"* in Papineau (2013)
[9] Another way to express this view is that belief has an essential "truth *telos*" (Gendler 2007).
[10] Perhaps it is accurate to attribute descriptor to William James (1896) who claimed that the aim of belief is truth.

(McHugh & Whiting 2014: 706; Steglich-Petersen 2017) of belief's normativity in terms of its essential aim or telos rather than claiming that it is a basic, fundamental fact about belief that it is essentially governed by certain norms. Moreover, for Teleologists, who spell out the aim of belief in terms of a biological design or proper function which has been selected for in virtue of its adaptive value (Steglich-Petersen 2013, 2017; Velleman 2000), Teleologism has the additional theoretical virtue of being capable of offering a *naturalistic* account to explain the essential normativity of belief. Proper functions are often deployed in naturalistic explanations, especially biology, to explain the nature of some kind. Such a truth-directed proper function is comparable to that of the heart, which is to pump blood, or the pancreas, which is to produce insulin, or sperm cells, which is to fertilize an ovum (Millikan 1984, 1993). For Teleologists, attributing a proper function to belief-producing mechanisms in our cognitive system is just an extension of this explanatory strategy. Plausibly, we might think that an account of belief that can give a naturalistic explanation of its normativity is a theoretical desideratum since belief is a psychological kind, within the purview of psychological sciences, which traffic in naturalistic or empirical explanations of phenomena rather than making appeal to *sui generis* normative facts or notions.

Teleological accounts (Velleman 2000; Steglich-Petersen 2006, 2009, 2017; Gendler 2007) ground the "essential truth-tie" of belief in its proper function of aiming at the truth (Steglich-Petersen 2017: 177). Belief is responsive to truth *because* it is the proper function – the essential *telos* (Gendler 2007: 236) – of belief or the regulatory cognitive systems responsible for producing beliefs (Velleman 2000; Steglich-Petersen 2017; McHugh & Whiting 2014) to get the truth-value of their propositional content right, that is, to believe truly, or to produce true beliefs. The normativity of belief – the truth norm – thus derives its connection to belief as a natural corollary of belief's essential aim or *telos* (structurally analogous to the instrumental value of true beliefs and the navigational function (Funkhouser 2017: 817)). This corollary specifically is explainable in terms of

the truth norm's *instrumental* value in relation to the aim of belief. Citing again McHugh and Whiting (2014: 707),

> …basing a belief on sufficient evidence is instrumental to satisfying the aim of believing truly, since beliefs based on sufficient evidence have a greater tendency to be true than those not so based. This…is why beliefs based on sufficient evidence have a positive normative status.

Thus, epistemic norms, including a norm of truth, are crucial to belief insofar as beliefs formed in accordance with such norms tend to be true and thus satisfy the aim of belief.

As just previewed, Teleologists locate the aim of belief at either the *personal* or *subpersonal* level (Velleman 2000; Steglich-Petersen 2017; McHugh & Whiting 2014). At the personal level, the function of aiming at truth is construed in terms of the aims or intentions of the agent, S, who intends to believe truly and forms beliefs through acts of judgement and deliberation. At the subpersonal level, the function of aiming at truth is spelled out in terms of the regulation of the cognitive system which subsumes belief-producing mechanisms, presumably selected for, and responsible for outputting beliefs which are subsequently taken up at the personal level or are consciously accessible for rational control, such as reasoning and decision-making (Block 1995).

Normativism, on the other hand, is the view that it is "part of the essence of belief to be subject to" the truth norm (McHugh & Whiting 2014: 698). According to this view, *belief* is a normative concept and "beliefs have normative essences – they are *constitutively normative*" (Steglich-Petersen 2017: 179). Normativists construe the truth norm as an essential, basic, and non-reducible feature that is constitutive of the belief attitude and is neither grounded in nor explainable in terms of any more fundamental or external facts, such as its adaptive value. The implication of such a view is that

> …defenders of distinctive doxastic norms...think that the very state of believing is *constituted* by distinctive doxastic norms. Subjects who are not governed by the relevant norms do not qualify as believers…if this is right, the very existence of believers implies there are distinctive doxastic norms…we cannot deny they exist without denying beliefs (Papineau 2013: 70-1).

Accordingly, for Normativists, subjects may be in belief-like states (e.g., pragmatic acceptance, stipulation, pretense), but unless these intentional attitudes are constituted or governed by the norm of truth ("doxastic norms"), they fail to count as beliefs, and therefore individuals in such psychological states are *ipso facto* not believers.

1.2 Belief Demarcation: Motivation for Doxastic Veridicalism

A central theoretical motivation for Doxastic Veridicalism is the explanatory value in accounting for what distinguishes beliefs from other kinds of propositional attitudes. This strategy proceeds by observing what relations, ***Φ***, to truth are too weak and hence too inclusive for delineating belief. Conative attitudes, for instance, like desire as well as other kinds of cognitive attitudes consist in some relation between the concept of truth and the attitude's propositional content. Normativists and Teleologists alike, for instance, assume that among belief's general relations to truth, the following are insufficient for individuating belief from other intentional attitudes:

(1) ***Φ***-ing *that p* to be true (e.g., desire, hope)

(2) ***Φ***-ing *that p as* true: representing or regarding *p* as true (e.g., cognitive attitudes including accepting, supposing, imagining, stipulating)

Conditions (1) and (2), Doxastic Veridicalists assume, are too weak to serve as criteria for singularly picking out belief from other intentional attitudes like desire, acceptance, and imagining, which also "involve an attitude toward the truth of [their] propositional content" (Steglich-Petersen 2017: 178). When we desire *that p,* we desire that *p* represents a true state of affairs. Yet, there is nothing defective nor has anything gone wrong when we desire *that p* and *p* is false. In fact, we tend to desire states of affairs to be true which are not so at the time of our desiring (though we also desire states of affairs to remain the case as well).

Similarly, for other kinds of cognitive attitudes: when I imagine or pretend that I am a 13th century peasant or that Tupac is still alive, neither I nor the cognitive attitudes involved in representing such states of affairs are blameworthy or defective if in fact Tupac is dead or I am not a 13th century peasant. Thus, there must be a stronger, more exclusive relation that holds between beliefs and truth that distinguishes belief from other intentional attitudes.

The point of difference then among Veridicalists such as Teleologists on the one hand and Normativists on the other is over the fundamentality of the truth norm and the source of belief's normativity. For Teleologists, the necessary condition for *x's* being a belief is captured by the following relation to truth:

(3) (TEL) Aiming (or intending) to ***Φ*** *that p* truly

where *aiming* refers to either (a) personal level intentions, judgements, and deliberations *or* to (b) the regulation of subpersonal belief-producing systems, presumably selected for in virtue of their fitness-enhancing properties.

Normativists, on the other hand, formulate the individuating condition of belief in terms of a norm of truth,

(3) (NORM) S *ought* to ***Φ*** *that p* only if *p*

with the normative operator *ought* identifying a relation to truth which belief uniquely bears. NORM is supposed to identify an appropriately stringent condition necessary for classifying some mental state as a belief.

Of course, for the Teleologists too, NORM picks out an individuating condition of belief. But for the Teleologist, the individuating condition is derivative, following from TEL, which is the fundamental relation sufficient for individuating belief from other intentional attitudes. TEL entails

that S ought to believe *that p* only if *p* is true – or so Teleologists commonly claim – in virtue of its instrumental value.

1.3 Problems with Doxastic Veridicalism

One worry, however, is that the truth-criterion – either in its TEL or NORM construal – is insufficient for individuating beliefs from other intentional attitudes. Other cognitive attitudes, e.g., *guesses,* are subject to a norm and/or aim of truth; a guess is correct, for instance, if and only if it is true (cf. Owen 2003).

Despite the theoretical virtues of Veridicalist accounts of belief, especially in terms of explanatory power and parsimony, they nevertheless face key empirical and theoretical issues.[11]

For Teleologists, Normativists' explanations of belief's essential connection to truth in terms of *sui generis* doxastic norms leaves much to be desired; identifying the norm of truth as a conceptually and metaphysically primitive feature of belief is, for the Teleologists, explanatorily unsatisfactory and inadequate in a theory of psychological attitudes. The doxastic norm, i.e., the truth-norm, is supposed to be the *explanandum,* not the *explanans*. Quoting Steglich-Petersen,

> Just as the causal roles played by belief call for explanation, the normative roles of belief call for explanation too—we should be unsatisfied with an account that bottoms out in certain constitutive norms, the force of which themselves cry out for explanation (2017: 179)

Despite the ability of Teleologism to naturalistically explain belief's normativity, it should be noted that Veridicalism is not the only game in town, however. By the Teleologist's own naturalist criterion, if there are proper functions of belief besides that of aiming at truth – functions which have been selected for in virtue of their fitness-enhancing properties or historical role in the evolution of belief – then such functions would also constitute an aim of belief. Moreover, such

[11] See McHugh & Whiting's (2014) "The Normativity of Belief" for an overview of various Normativists accounts and critiques of them.

proper functions or aims cannot be specificed *a priori*; they must be empirically discovered. This is the case, at least, if we are employing a naturalistic criterion for determining the aim of belief, as Teleologists do, which looks to the selected-for proper function of belief to determine its aim. Furthermore, the proper function of belief can only be determined *a posteriori,* through empirical discovery. And whatever the proper functions are that constitute the aims of belief, which our best empirical theories of belief identify, these will be the functions which shed further light on the applications of belief. To see a key issue then which Veridicalism faces, I would like to note a presupposition shared by Normativists and Teleologists: *doxastic absolutism*. This is the view broadly that

> …there is a criterion such that whether a belief satisfies it will determine whether the belief is right or wrong in an absolute sense (Zalabardo 2010: 2).[12]

Traditional Veridicalist theories presuppose doxastic absolutism in connection to the truth-criterion, assessing the correctness of beliefs based on whether they are true or false. When talking about doxastic absolutism, we mean absolutist conceptions of belief which privilege the truth-criterion. Doxastic absolutism contrasts with the views of *doxastic relativism*, according to which,

> Beliefs count as right or wrong relative to the criterion employed in each assessment, and no particular criterion enjoys a privileged status that justifies speaking of beliefs as right or wrong *simpliciter*, according to whether they satisfy this criterion (ibid.).

Thus, for instance, beliefs may be evaluated for their capacity to produce or promote subjective and psychological well-being (Bortolotti 2020; Papineau 2013; Taylor & Brown 1988), to contribute to subjective utility, or for their conformity with ingroup norms (Zalabardo 2010). Beliefs can also be evaluated for their capacity to enhance the fitness or utility of the believer according to an array of diverse standards – e.g., sociocultural, biological, personal – which at times may pull in

[12] In principle, there are doxastic absolutists who do not privilege the truth criterion but some other criterion by which beliefs are evaluated or assessed for their correctness. For pragmatists, they may privilege the utility criterion and thus maintain that we ought to assess the correctness of belief based on its contribution to subjective utility.

opposite directions, such that one and the same belief may be positively and negatively evaluated according to different criteria. So, according to doxastic relativism, it is possible (and in fact the case) that beliefs are evaluated according to different criteria, such as various pragmatic standards of utility or social criteria, which people employ to assess the correctness of a belief depending on the context and on the interests and purpose of those evaluating a given belief. So, while there are no *a priori* privileged criteria of assessment, are there *a posteriori* privileged criteria?

An important distinction, as Zalabardo notes, is that classifying a theory of belief along the absolutism-relativism divide "is independent of the specific criterion to which it ascribes this privileged status" (Zalabardo 2010: 2). So, for instance, a theory may recognize the undeniable empirical fact that people often form and evaluate beliefs with respect to the truth criterion while also denying that the truth criterion is conceptually or metaphysically privileged; it is just that it is in virtue of our purposes or our practices or some other motivation that we tend to assess beliefs in terms of the truth criterion. Furthermore, given the view that the truth criterion holds no privileged status and is just one of many possible criteria by which beliefs may be assessed, this stretches the evaluative space of belief and makes room for advancing other evaluative criteria, which may be grounded in natural functions of belief. It is then coherent to evaluate beliefs not for their truth-value but for their conduciveness to, say, cultural success by enhancing the social reputation of S.

As may be anticipated, doxastic relativism is an important thesis in relation to the tribal belief account generally and a social signaling model of belief specifically, which evaluates beliefs based on their communicative capacity, their ability to modify the behavior of the receiver, and their contribution to the fitness of the receiver and/or signaler. It is a thesis like doxastic relativism which such a social account of belief needs to open up the psychological space for proposing finer-grained ways of individuating subspecies of belief based on their proper functions or evaluative criteria.

Importantly, the theory of tribal beliefs I advance presupposes the principle of doxastic relativism, which allows for a plurality of criteria by which we may legitimately evaluate and analyze the belief attitude, and for some of which, the truth criterion may in fact be orthogonal or even detrimental at times. Determining which criteria people actually evaluate beliefs by is an empirically informed and open-ended enterprise. Postulating the category of tribal beliefs is partly motivated by empirical evidence suggesting that many of our beliefs are sensitive to social contexts as well as by theories which attempt to explain this evidence in terms of beliefs functioning as social tools for expressions of *group* or social identity (Funkhouser 2020; Boyd & Richerson 2005; Kahan 2016, 2017; Tajfel & Turner 1986) and which contribute to coordination and cooperation and cultivating prosocial behavior (Boyd & Richerson 2005; Zawidzki 2013; Sterelny 2012).

A relativistic framework, however, is not doxastic anarchy. We can, perhaps, organize doxastic criteria into *a posteriori* hierarchies based, for instance, on those functions which have figured prominently in the selective history of the belief attitude or which have crucially contributed to our species' adaptive success; or, in terms of frequency, those criteria by which beliefs are typically assessed. Such a move does not privilege certain criteria *per se,* that is, conceptually or metaphysically, but allows us to organize the diverse criteria pragmatically and empirically along some metric, which is determined by our theoretical and practical interests (e.g., for the mental health clinician and the epistemologist, the criteria salient to assessing belief will be very different presumably). The question becomes which meta-criteria or which framework should we use to organize and order the different evaluative criteria of belief, which, again, may be determined solely by pragmatic considerations.

Proper functions offer a theoretically viable strategy for determining, within a relativistic framework, which criteria are appropriate for doxastic evaluation, based on which criteria have historically played an important role in the development of the belief attitude or the psychological agents who traffic in them, explaining the seemingly intimate connection of some evaluative criteria

to the belief attitude (e.g., truth criterion, utility criterion, social expression criterion) in comparison to others, say, the criterion of beauty (believing *that p* just because of its aesthetic value, say). Proper functions also provide a naturalistic basis for determining which criteria are salient to belief evaluation in terms of actual human psychology.

Lastly, if we are inclined to think that belief's proper or natural function is constitutive of the attitude, such that it is necessary to *x's* being a belief that it performs function *F*, then such a function will ground which criteria are relevant to the assessment of belief based on its performance of its constitutive functions, which may form a disjunctive functional profile. While such a model might not strictly be a relativist one, it suggests a more liberal view than standard absolutist ones, with the criteria that are privileged (a) being determined empirically based on the natural or proper function of belief, rather than being a primitive normative fact about the concept of belief and the meaning of the concept (though Teleologists also pride themselves on grounding the truth-criterion in a naturalistic account of belief); and, more importantly, (b) it allows for a multiplicity of privileged criteria by which we assess beliefs, since beliefs can and seemingly do have multiple proper functions, as is empirically suggested. This at least runs counter to the spirit of Normativist and Teleologist accounts since it allows for the privileging of criteria besides that of the truth-criterion, grounded in the natural functions of belief which have been selected for in virtue of their adaptive value, and among the set of empirically privileged criteria, one criterion is not privileged among any other and a belief may still be evaluated as correct even if false in virtue of its utility or fitness-enhancing properties.

1.4 Doxastic Veridicalism: The Dilemma

To be clear, the dilemma which Veridicalist theories face is the dilemma as to whether seemingly arational attitudes like politically motivated and conspiracy beliefs such as those that

traffic in misinformation about COVID-19, COVID vaccines, and anthropogenic climate change, are genuine beliefs. Either Veridicalists must deny that the operative attitudes are genuine beliefs, effectively denying a doxastic account of conspiracy theories, even though conspiracy theorists often seem to be making sincere assertions and even act on conspiratorial beliefs – actions which can lead to violence and incarceration for the believer (Pizzagate!). Or, if Veridicalists accept that these are genuine beliefs and accept a doxastic account of conspiracy theories, then it seems they are committed to demoting the status of the truth-criterion in their conceptions of belief. In other words, they are committed to accepting that beliefs do not essentially aim at truth or that beliefs are not essentially constituted or governed by a norm of truth or epistemic rationality; the rationality constraint is at most eliminable and at least negotiable.

Structurally, this dilemma parallels that over the belief status of *delusions* (Bortolotti 2009; Bayne & Pacherie 2005; Campbell 2001). This dilemma frames the debate for theorists whether to embrace a doxastic or anti-doxastic account of delusions. This parallel stems from the similarity between delusions and conspiracy theory beliefs along the following dimensions: (1) the implausibility of their propositional content, (2) their fit with or responsiveness to evidence, or the failure to be based on *good* reasons rather than *mere* reasons, or their lack of evidence entirely, (3) they are both pragmatically self-defeating, and (4) they often fail to issue in the appropriate actions or inferences (Bayne & Pacherie 2005: 163-165).[13]

Yet, the defender of doxastic accounts of delusions can reasonably parry each of these objections. For instance, we ordinarily believe things that are logically and nomically impossible (e.g., many religious and new age beliefs), incoherent, and even inconsistent both locally and globally. We believe highly implausible things that are in direct contradistinction with our prior beliefs and

[13] Although an objection against doxastic accounts of delusions is that they fail to issue the appropriate affective response to the content of the delusion as a genuine belief would, for the conspiracy theorist, there often seems to be a strong affective dimension.

concepts. Moreover, many of our ordinary beliefs are based on motivated reasoning, such as positive illusions, wishful thinking, and the beliefs underpinning self-deception. We form beliefs based on mere reasons rather than epistemic reasons; and we often construe *q* as evidence for *p* even if considerations of *q* have no rational connection to the truth-value of *p*. To be clear, *mere reasons* are causal reasons which fix a belief; they are not, in distinction, rationally responsible reasons based on the appropriate fit between a belief and a body of evidence or its consistency with other beliefs. Moreover, we habitually and persistently arrive at beliefs through fallible and unreliable processes and commit various formal and informal logical and argumentative fallacies, from fallacies of inductive and probabilistic reasoning to formal fallacies such as affirming the consequent. We pervasively form ordinary beliefs which lack evidence or do not fit the evidence in the right kind of ways or which violate the rules of reasoning in the formation-process.

Lastly, like delusions, conspiracy beliefs often figure as theoretical reasons in inference-making or as practical reasons in action-guiding – people exhibit commitment to their political and conspiratorial attitudes. Conspiracy theorists about COVID-19 ransack testing sites (Browning 2021), refuse to get vaccinated, wear masks, or even social distance; they also write their representatives and government officials (even threatening them), attend protest rallies, etc. Yet, it is not always clear what action can be guided by the conspiracy theory attitude: if one believes that anthropogenic climate change is a hoax or that COVID-19 is a hoax, it is not clear what practical actions such beliefs can issue in, except perhaps consuming information which sustains or reinforces the belief, sharing information, and attending occasional protests or political conferences. Yet some COVID-19 conspiracy theorists have threatened to kill public officials or even been discovered en route to kill public officials, including Dr. Fauci (Li & Lee 2021). Thus, conspiratorial beliefs, like delusions, often do guide action, supporting the view that such attitudes are genuine beliefs. Even when, however, it seems that delusions or conspiratorial beliefs fail to elicit or guide action or

function as a practical reason, they are, as Bayne and Pacherie note, "no less capable of guiding one's action than is the thought <there are no irrational numbers>" (2005: 182). In general, it is questionable whether the criteria that the separate objections presuppose are in fact constitutive of belief, or rather normative constraints that prescribe how we rationally ought to believe or behave if we believe *that p*. Or they may be stereotypical dispositions or features we associate with belief, as a cluster concept (Schwitzgebel 2010), such that it is possible for there to be give-and-take in the characteristic features of belief for a psychological attitude to still count as a belief – *x* can still maintain its belief status even if it fails to instantiate the full list of stereotypical dispositions or associations deployed in belief-attribution. Moreover, the extent to which ordinary beliefs or paradigmatic cases of belief satisfy these conditions admits of degrees, with different beliefs more or less issuing in the appropriate commitments (e.g., practical and theoretical reasoning or affective response), being more or less based on evidence, being more or less consistent with other beliefs, and being more or less plausible in their content. Given these facts about ordinary beliefs, the irrationality or arationality exhibited by delusions and conspiratorial beliefs is not a difference in kind from ordinary beliefs, but a difference in degree.

In fact, the debate over the belief-status of delusions has been forceful in decentralizing or demoting the rationality constraints of belief and decentralizing the role of truth in the normative content of belief, showing that the irrationality of delusional beliefs is on a continuum with ordinary beliefs, being merely a difference in degree, not in kinds of irrationality (Bortolotti 2010). Given the qualitative similarity between delusional beliefs and conspiratorial beliefs, this would seem to both further support the demotion of truth and rationality as core normative principles in our conception of belief and to make room for conspiratorial attitudes to be accommodated by revised and empirically informed conceptions of belief.

If the foregoing is correct, then the Veridicalist, despite her theoretical inclination to the contrary, cannot endorse the first disjunct of the dilemma and deny that such beliefs as those that underpin COVID-19 and climate change conspiracy theories are genuine beliefs. At the very least, there are *prima facie* forceful reasons for thinking that a moderate doxastic account of conspiracy theories is correct, covering many but perhaps not all cases. And it is by no means an obvious or settled fact that such attitudes are not beliefs. This then leaves the Veridicalist saddled with the second disjunct, which is to essentially admit that is incorrect and that belief is not essentially governed or constituted by a norm of truth, nor does belief essentially and exclusively aim at truth, which would be a fundamental reorientation in traditional accounts of belief in philosophy of mind and epistemology.

For the doxastic relativist (or *doxastic pluralist* we might say) a potential "solution"[14] – besides happily accepting the second disjunct – to the dilemma which Veridicalist and Rationalist theories of belief face, given the existence of conspiratorial belief, positive illusions, and other seemingly non-truth-centered beliefs which stand in need of explanation, is to propose alternative, extra-veridical, extra-epistemic functions for belief besides that of aiming at truth or knowledge – functions which generate their own normative constraints and success criteria and for which the truth-criterion is not essential nor relevant. If one is a pluralist about the functions of belief (specifically the proper functions of belief) and their corresponding criteria of evaluation, then there is no parallel dilemma – the dilemma dissolves. We can instead attempt to explain conspiratorial beliefs in terms of alternative functions which the belief attitude performs and explain them by appeal to the utility they engender in virtue of performing these different functions. The functional suite of the belief attitude, as previously suggested, might be a disjunction of various and disparate functions which

[14] Of course, the doxastic relativist or pluralist doesn't face the same challenge and can happily accept the second disjunct of the dilemma.

have been selected for in virtue of their utility, say, adaptive utility, cultural utility, or personal and psychological utility.

§2. Social Signals: A Cultural Function of Belief

A plausible alternative function which has begun to gain traction recently is the proposal that beliefs function as social signals (Funkhouser 2017, 2020; Williams 2021; Boyd & Richerson 2005). According to signaling accounts generally, the function of some selected for signal (e.g., a belief) is to communicate information to a receiver in order to modify the receiver's behavior, which can enhance the fitness of both the signaler and the receiver, as when they share common interests or goals, or may enhance the fitness of the signaler or receiver at the expense of the other (Sterelny 2012; Boyer 2018). In terms of sociality specifically, the function of beliefs *qua* signals is to communicate important or strategic social information to others in the signaler's social environment, be they within or outside of the signaler's own ingroup (Boyd & Richerson 2005; Sterelny 2012; Boyer 2018).

Generally, there seems to be two kinds of social information which beliefs signal: (i) information about which ingroup the signaler is a member of, or, group affiliation, and (ii) how committed the signaler is to the affiliated ingroup, that is, how reliable, trustworthy, and loyal a member she is in her capacity as a group member (Funkhouser 2020; Sterelny 2012). This distinction delineating the two different kinds of social information which social signals communicate is the basis of a *Disjunctive Model of Social Belief Signaling*. The disjunctive model not only tracks a distinction in the kinds of information communicated but also tracks differences in the structure of the signal.

2.1 The Structure of Signals

Before presenting a disjunctive model of belief signaling, it will be useful to explicate a general theory of signals to get a sense of the basic structural features of a signaling framework. Adapting Funkhouser's (2017: 811-12) general triad, we formulate the condition for something's being a signal as,

> *x* is a signal if and only if:
> (1) *x* is designed or selected-for
> (2) for the sake of being detected by some receiver R
> (3) in order to modify R's behavior

(1) establishes the *selection* condition necessary for something's being a signal. This condition distinguishes signals from mere *cues* and objects of natural meaning (Funkhouser 2017; Maynard Smith & Harper 2003; Grice 1957). While cues communicate information that can guide behavior, this is not what they are selected or designed for; thus, it is not a proper function of cues to communicate information in order to modify behavior. Furthermore, since they are not selected for, their communicating features or mechanisms are not enhanced nor optimized for the purposes of more effectively communicating information. Examples of cues include rain clouds, smoke, deer tracks, or red spots on the skin caused by a disease, e.g., measles (Grice 1957: 377). While the presence of each conveys information (e.g., smoke conveys that there is fire), they are not designed nor selected for communicating information. In contrast, a peacock's plumage functions as a signal since it has been selected for via natural selection to communicate information about the differential genetic quality of individual specimens and thus the desirability of an individual as a potential mating partner. As a result, the plumage of peacocks tends to be exaggerated or enhanced for the purposes of communicating information about genetic quality, which, upon successful detection by a peahen, results in modifying her behavior toward the signaling peacock.

An important consequence of (1), previewed just now, is that the evolution of a signal tends to favor the development of the signal in such a way that it optimizes the signal's ability to communicate information and be detected: signals tend to evolve so as to enhance their *ease of detection* for the receivers. Ambiguity in the signaling medium not only obstructs the proper function of communicating information but it also wastes resources (e.g., time, energy) if the signaler S undertakes the act of signaling, yet receivers are unable to detect or disambiguate the signal. Instead, receivers will tend to modify their behavior towards those signals that are easier to detect, have clearer resolution, and are less ambiguous (though this may be maladaptive if the interests of the receiver and the signaler diverge and the signaler communicates harmful or misleading information through deceptive signaling, in which case, the specific signaling enterprise will become unstable as receivers adapt to distrust signals of a given type and from a specific source). Ambiguity in the signal, besides underdetermining a medium's signaling status, might also correspond to ambiguity in the informational content, leaving the receiver unsure what exactly is being communicated and thus unsure how to appropriately modify her behavior accordingly.

Lastly, the selection process of signals is not just through "natural selection" but can also include *intentional choice*, *individual learning*, and *cultural transmission* (Funkhouser 2017: 812). Especially relevant to our current project is the means of selection through intentional choice and cultural transmission. This is because, as social signals, beliefs are often selected for through the intentional choice of group leaders, elites, and other group models, such as politicians, media personalities, influencers, or "conspiracy entrepreneurs" (Sunstein & Vermeule 2008), etc., wherein the signal is subsequently taken up by common group members for signaling social information. Beliefs as social signals are also selected for through cultural transmission, such as through institutions of cultural pedagogy and group narratives (Zawidzki 2013), through one's community or family, or through social media and other word-of-mouth analogues. Cultural transmission of signals depends

importantly on our motivation to imitate the behavior and beliefs others within our group, especially group models, and to conform to social norms, with such conformity often encouraged through mechanisms of norm enforcement and punishment of norm violators.

Another important feature of the signaling framework is the honesty mechanism which ensures that the information the signal communicates about the signaler is accurate. The costs of undertaking certain signals ensures that they are an honest means of communication since the cost of the signal makes it untenable to fake and to deceive the receivers. *Conspicuous consumption* is an example of a type of costly signaling which helps ensure the honesty in the quality of the signal (Veblen 1899): buying an expensive watch, a luxurious sports car, or luxurious foods, for instance, is taken to be an honest signal of one's wealth and social prestige given the financial cost of these items since only a person who actually has the expendable wealth could afford the loss of paying for a Rolex or a Ferrari. A person who is trying to appear rich and thus fake their wealth presumably cannot afford to spend money on a Rolex or Ferrari since this will incur too great of a financial cost. The relative cost suffered by the imposture would be much greater than the cost for the person who in fact has expendable wealth so that engaging in such signaling practice would be maladaptive and counterproductive in terms of the payoff matrix, costing more than any benefit the fake signal might incur. Thus, the costliness of signaling is a mechanism which keeps the signal honest, ensuring that the information which the signal communicates about the signaler is accurate.

To be clear, not all signals are costly. Some beliefs and behaviors are merely *credibility enhancing displays* ("CREDs") (Henrich 2009). If you're trying to get others to drink absinthe but they are unsure if it's safe to drink, you can enhance the credibility of your claim that it is safe by taking the first swig. If you have the prior belief that absinthe is safe to drink and you want to drink absinthe, then this is not a costly display – it only seems costly to those with different prior beliefs and values about absinthe, drinking alcohol, etc. Furthermore, there are different means by which a

signal's cost are implemented, that is, in the structure of the cost mechanisms. For our purposes, we will be concerned with only two: *handicapped* (Zahavian) costly signals, which communicate the differential quality of an individual, and *public investment* costly signals, whereby the signal modulates the payoff of group defection so that it becomes too costly and hence maladaptive (Sterelny 2012).

2.2 Disjunctive Model of Social Belief Signals

With a general theory of signals in hand, we can begin to sketch a disjunctive model of belief signals based on the two principal kinds of social information people communicate through the medium of beliefs. As a social signal, beliefs specifically signal at least the following disjunctive content:

(1) S is a member of ingroup *g* (*Shibboleth signaling*)
Or,
(2) S's degree of commitment to ingroup *g* (*Commitment signaling*)

2.2.1 Shibboleth Signaling

I begin with a formulation of shibboleth signaling (SHIB) and then discuss how it fits within the larger empirical and theoretical literature. Afterwards, I discuss why beliefs which function as shibboleth signals might be selected for in terms of what benefits they incur for either the signaler, the receiver, or both. It should be noted that the term *adaptive* is used broadly, as an umbrella term for different standards of fitness, including biological, cultural, individual, psychological, and hedonic fitness, with the different standards often interrelated to each other.

> **(SHIB)**: *x* is a *shibboleth signal* whenever *x* has been *selected for* so as to communicate to some receiver R who detects *x* that signaler S is a member of ingroup *g, and* in fact *x* communicates (typically) to some receiver R who detects *x* that signaler S is a member of ingroup *g*

The first conjunct states the purpose of the shibboleth signal, which is grounded in the kind of information that it communicates: tribal affiliation. Shibboleth signals are selected for in virtue of

their capacity to index which ingroup (or cluster thereof) the signaler is a member of. In essence, shibboleth signals communicate where the signaler's tribal allegiances lie.

To get an intuitive feel of shibboleth signals, let's examine certain climate and health beliefs:

(1) Anthropogenic climate change is a hoax (e.g., Climategate)
(2) COVID-19 is a Democrat hoax
(3) COVID-19 vaccines cause impotency/magnetism, etc.
(4) Wearing masks is dangerous/ineffective

Empirical data from researchers and professional pollsters confirm what we already suspect, which is, believers of anti-masker and anti-vaxxer rhetoric overwhelmingly self-identify as and are members of the cultural and political Right.

Recently, researchers have looked at the effects of cultural and political worldviews (partisanship, political ideology, Trump approval, and conspiracy ideation) on "forming novel health beliefs" (Shepherd, MacKendrick, & Mora 2020: 2), specifically beliefs about the existence of COVID-19, its perceived risk, and the effectiveness of medical and behavioral preventative measures, including wearing a mask, social distancing, and getting vaccinated (Pivetti et al. 2021; Shepherd, MacKendrick, & Mora 2020; Hornsey et al. 2020). Shepherd, MacKendrick, and Mora (2020: 2), for instance, specifically seek to answer, "how the polarized political context and the politicized nature of COVID-19 public discourse has shaped American's COVID-19-relevant beliefs and everyday actions." Analyzing data collected from 8,800 registered voters in California a month into the state's lockdown, they studied the relation between health beliefs about COVID-19 (its existence, its perceived risk, and medical and behavioral prevention) and political worldviews. Like public attitudes towards climate change, health beliefs about COVID-19 and trust in public health institutions vary along tribal lines of political and cultural worldviews. Their analysis indicates a "significant association between political worldviews, regardless of the measure, and Americans' COVID-19-related health beliefs and behaviors," concluding that on average, health beliefs about

COVID-19 track social identity and categorization, that is, tracks an individual's tribal alignments. Summarizing their findings, they write,

> Republicans, ideological conservatives, and Trump supporters express less trust in public health institutions; perceive less risk of contracting COVID-19; believe individualist preventive strategies, such as vitamin use and diet, to be more effective and public health preventions measures, such as staying at home and wearing a mask, to be less effective; and leave the house more frequently than…Democrats, ideological liberals, and those who disapprove of Trump.
>
> (Shepherd, MacKendrick, & Mora 2020: 14)

Nor are such findings outliers but are supported by other studies mapping the demographics of anti-maskers, anti-vaxxer, and anti-COVID attitudes (Pivetti et al. 2021; Hornsey et al. 2020). These studies suggest that beliefs and behaviors which express anti-masker, anti-vaxxer, and anti-COVID attitudes indicate a person's political and cultural alignment, communicating where in the social space their tribal allegiances lie.

Climate beliefs too are an effective predictor and indicator of a person's tribal memberships. Strikingly, a person's measure of scientific literacy and numeracy (i.e., scientific knowledge and quantitative reasoning) is not the most reliable indicator of that person's beliefs about the anthropogenic status or dangers of climate change. To infer a person's climate beliefs accurately and reliably, we must also consider the variable of tribal alignment, specifically whether the person is a conservative Republican or a liberal Democrat (the two dimensions of *ideology* and *partisanship*). While for someone who self-identifies as a progressive, a liberal, or a Democrat, the trend runs in the direction we might expect, with greater scientific knowledge and reasoning correlating with a greater belief in anthropogenic climate change and its dangers, the opposite trend is true of someone who self-identifies as a conservative Republican (Greene 2013; Kahan et al. 2012; Kahan et al. 2010). A conservative Republican who scores high on measures of scientific literacy (knowledge) and scientific numeracy (quantitative reasoning) is, on average, more likely to be skeptical of climate change than her liberal Democrat counterpart.

This of course is not to say that everyone on the cultural and political left believes in or expresses true beliefs about anthropogenic climate change for epistemically virtuous or legitimate reasons – no political or cultural group or tribe is exempt from social biases and motivated reasoning. These motivated reasons are to express their tribal alignment, signaling to both ingroup (including themselves) and outgroup members, which political and cultural worldviews they are committed to – which cultural and political tribe they are members of – rather than being motivated to believe truly or get the truth-value of their beliefs right. It also signals how embedded they are within the respective ingroup, how much they conform to the prototypical or ideal group member, and how reliable or trustworthy of a group member they are. These underlying motivational reasons modulate the expressive function of their views and beliefs about climate change (or COVID-19) for the sake of *identity protection* (Kahan 2016, 2017). When prompted to express their beliefs or views about climate change or COVID-19 policies, people are motivated in their reasoning to protect their social identity by expressing their cultural commitments and worldviews, which presumably aligns with that of their ingroup, in this case, their cultural and political tribe. Failing to express views or beliefs that align with the messaging of one's cultural and political tribe might be taken as evidence by others and by oneself that one is a defective group member. Thus, in order to manage and maintain our self-conception and the conceptions other have of us as a definitive group member, we are motivated to update our beliefs such that they conform with our group's messaging. In other words, the social cost of failing to believe in alignment with ingroup messaging outweighs that of the epistemic benefit of believing truly or epistemic cost of believing falsely; determining what to believe is filtered through a social cost-benefit analysis (Funkhouser 2020).

This trend suggests at least two things: (i) we can account for the average American's views about climate change based on their tribal sensibilities (Greene 2013: 91-93) – how they culturally and politically self-identify – and (ii) that people's views about anthropogenic climate change are

acquired and maintained by socially motivated reasoning. When it comes to conservative Republicans, a common explanation of the negative correlation between scientific knowledge and reasoning on the one hand and their disbeliefs in anthropogenic climate change and its dangers on the other is that those who score higher on measures of scientific knowledge and reasoning are more adept at rationalizing their desired conclusion, which is established by tribal norms and group messaging. With a greater ability to rationalize their beliefs, say, that the threat of climate change is exaggerated by the media, the scientific community, and the political Left, or the belief that climate change is just part of a natural cycle, explainable primarily in terms of natural causes, so too their confidence in such beliefs increases. Thus, their disbelief in anthropogenic climate change is socially motivated, though this socially motivated reasoning is modulated by abilities in individual cognition, such that the performance or execution of motivated reasoning is enhanced or optimized by increased abilities in individual cognitive style (cf. Stanovich 2011).

Further evidence that this trend is the result of tribal influences is the fact that skepticism on the Right of anthropogenic climate change is on the rise, suggesting that it is not the result of scientific ignorance, but rather, the result of a social trend. If the rejection of anthropogenic climate change was the result of sincere debate over whether climate change is anthropogenically induced and the existential threat it poses, then why would a smaller percentage of conservative Republicans disbelieve in anthropogenic climate change now than they did at the beginning of the 21st Century (Greene 2013: 93)? This polarization in climate attitudes is not tracking any discrepancy or split in the opinion of climate scientists for whom there is an overwhelming consensus (97%) (Hornsey & Fielding 2017) that climate change is anthropogenically caused and poses a vital threat. And if anything, the availability and access to climate change evidence has only increased, not dwindled. Rather than tracking any differences of opinions or developments among climate scientists, this polarization between those on the Left and those on the Right in their disbelief of anthropogenic

climate change seems instead to be motivated by tribal trends, such that these skeptical attitudes are performing some social function, which is, defining themselves in reaction to their Leftist outgroups through the mechanism of social comparison and outgroup enmity.

The shibboleth signaling model predicts this trend to the extent that its effect is the divergence or shift in views between the two sides, i.e., the two political and cultural tribes. Given that views on climate change have come to signal one's tribal affiliation, the more motivated are those on both sides to embrace views that contrast themselves against rival outgroups and to enhance or accentuate the differences between "us" and "them." This is analogous to and follows the general trend of polarizing attitudes as a means of optimizing the ease of detection of shibboleth signals such that the more polarized or extreme the belief becomes, the less ambiguous it is as a signal and the easier it is to detect as a social signal, both of which enhances its capacity to communicate socially strategic information.

2.2.2 Signaling & Polarization

Further support for the claim that health beliefs with respect to masks, vaccines, and the COVID-19 virus itself, or climate beliefs about anthropogenic climate change, potentially function among the population as a means of signaling group affiliation is that there is nothing inherit in the political ideology of the Right or Left, conservatives or liberals, from which the associated attitudes follow. Levy (2021a: 1-2) is right to point out that polarized responses to COVID-19 (climate change, etc.) and policies which institutionalize preventive measures are not explained by an

> …obvious linkage between divergent *values* and the position each side has taken. Of course Republicans can oppose vaccine mandates and lockdown in the name of freedom…but Republicans could just as easily support them in the name of ingroup loyalty and respect for authority (Graham et al., 2009) they equally value. Democrats might have opposed lockdown on account of the way in which they impose significantly greater burden on the vulnerable and the poor than on those who can easily work from home…The actual positions each side

> has taken seem significantly underdetermined by their prior commitments, suggesting that political polarization is explained by some very different kind of cause.

I propose that the polarized responses are explained, at least to a significant degree, in virtue of their signaling function. Such an explanation fits well with the fact the positions or attitudes of each side are "significantly underdetermined" since what attitudes signal membership within group will be partly determined by what views or attitudes competing outgroup endorses, presumably giving rise to a subsequent feedback loop between the two, causing views to further diverge and become more polarized. And what position or on which side of an issue a group finds itself can be informed by a myriad of historical contingencies, such as what views political leaders happen to hold, which can be motivated by personal interests rather than based on any prior ideological or evaluative commitments of the political group; what the scientific messaging about an issue is; what the position of scholars and academics happens to mostly be, and so on. The attitudes which signal group affiliation may be selected for in part because of their contradistinction with rival or competing outgroups or because such attitudes distinguish the group from the rest of society.

We mentioned that the ease-of-detection criterion is a selective pressure for signals, influencing which aspects of the signaling medium tend to be enhanced so as to optimize the signaling capacity. Given the selection for ease-of-detection or resolution, this also helps explain polarized responses and how underdetermined positions are subsequently determined. The more polarized the beliefs associated with rival groups become, i.e., the more divergent from each other they are, the less ambiguous they are in communicating the signaler's group affiliation. Thus, the correlative relation is roughly that the greater the divergence, i.e., polarization, between beliefs of competing or rival groups, the greater is the resolution and ease-of-detection of the beliefs *qua* signals and thus the more efficiently they perform their social signaling function.

What ensues once beliefs assume the shibboleth signaling function is an "arms race" with respect to the content of beliefs both within groups as members compete for the position of being

the most prototypical group member (McGarty et al. 1992; Tosi & Warmke 2016), to conform to the norms of the group and occupy the space most central to the group identity rather than letting their membership ambiguously reside on the margins. This is done with the aim of enhancing, maintaining, or rehabilitating their social reputation among their ingroup, which increases their social prestige and, concomitantly, the access and securement of group benefits, e.g., access to group resources, fellow cooperators, etc. In other words, enhancing their social reputation enhances their cultural success and social utility, which in turn can enhance their personal and biological utility.

Within the literature on virtue signaling and moral grandstanding, this arms race in ratcheting up the content of belief signals – leading to polarized viewpoints between rival groups – is known as *ramping up* and is characterized as one of the negative consequences or symptoms of virtue signaling (Tosi & Warmke 2016, 2020). Ramping up refers to the act of "making increasingly strong claims about the matter under discussion" which leads to a "moral arms race" (Tosi & Warmke 2016: 205). In the context of virtue signaling and the *public moral discourse* – the space in which people virtue signal – group members ramp up the content of their moral claims and beliefs so as to signal that they are "more attuned to matters of justice" (ibid.) in comparison with others in the moral space or their moral tribe, thus potentially enhancing their moral reputation by signaling we are the most morally respectable and attuned. Tosi and Warmke explain this symptom of virtue signaling in terms of the mechanism of *social comparison* (2020: 177; also cf. Festinger 1954):

> our self-conceptions partly depend on how we think we measure up against others. Other peoples' behavior affects both how we think of ourselves and how we appear to others, since those judgements are so often based on comparisons.

And in the arms race, to appear as the most morally respectable person within the public moral discourse,

> …people…try to outdo others to stand out, gain attention, and look like moral paragons either to an in-group or an out-group. This competitive attention seeking leads people to adopt (or at least express) radical views.

The thought is that, as part of their self-conception "people often imagine themselves as occupying a certain position in comparison to others" (Tosi & Warmke 2016: 205). Specifically, within the context of virtue signaling, this position is as someone more morally respectable and attuned than others – either than members within one's own moral tribe or members of an outgroup. More generally, this self-conception may be the image of ourselves as occupying a position as a prototypical or ideal group member. Yet, once someone else expresses a view suggesting they are more morally respectable, or more fully exemplify the features of an ideal group member, then to enhance or maintain our social/moral reputation and maintain and manage our self-conception and the conception of ourselves among others as the most morally respectable or the ideal group member, we must shift our moral, political, cultural, etc. beliefs to more extreme positions. Hence, the model of an arms race between ourselves and others competing to occupy the same position, to be seen as the moral paragon or the ideal group member.

Yet, the process of ramping up the content of our beliefs is not determined solely in response to what others in our own ingroup or moral tribe express; we must also consider the views and beliefs expressed by members of competing or rival outgroups or tribes. Revealing oneself to be morally competent or respectable means not only standing out among one's fellow ingroup members but also contrasting oneself against the views expressed by rival moral tribes, who presumably are not as morally attuned or oriented as one's own moral tribe. Again, generally, to represent oneself as a more prototypical or ideal group member, or group exemplar, means not only socially comparing oneself with one's fellow ingroup members and trying to out-do them but also contrasting oneself with competing outgroups, that is, contrasting one's views and beliefs against the views and beliefs expressed by such outgroups. This leads to a polarization among responses to an issue or among the views and beliefs expressed by competing groups as a consequence of both intra-

group social comparison and inter-group social comparison, causing views and beliefs on both sides to shift to more extreme or radical positions in their content.

Through the mechanism of social comparison, when this involves the dynamics of intra- and inter-group comparison (rather than mere intragroup social comparison) and thus "there is pressure to ramp up, the positions of the opposing sides of an issue will tend to move away from each other, and the middle ground will be seen by each side as requiring unacceptable moral compromise" (Tosi & Warmke 2020: 177) or ideological compromise or messaging/rhetorical compromise or compromise in general of the positions competing groups have staked out. Hence, polarization emerges.

Empirical support for this explanation of polarization through the mechanism of social comparison is supported by the research on *group polarization* of McGarty et al. (1992). Whereas studies on group polarization have traditionally focused on attitude shifts of the pre-discussion view, which is enhanced post-discussion, McGarty et al. studied the effect that perceived outgroups had on the attitude shift in group polarization paradigms – not just the shifts induced by deliberation and discussion within a group. McGarty et al. (1992) points to evidence showing that the shift of the group belief can be predicted by averaging the attitudes of *both* members in the in-group and the out-group, with the more extreme the outgroup's view is in one direction, the more extreme will the ingroup's view go in the opposite direction. That is, the view of the out-group is an additional variable in predicting the belief shift of the ingroup. This suggests that sometimes the shift in the group's pre-discussion view is as much about imitating and comparing oneself with the members of one's ingroup as it is with contrasting it against members of the outgroup.

Situating the signaling model within Boyd and Richerson's (2005) discussion of *ethnic markers:* if a belief is a signal *and* if signals are constrained by the need for greater resolution/clarity as an informational medium, then we would expect these beliefs to become more distorted in order to

better signal a person's tribal allegiance. Assuming a need for resolution as a selected for feature in signaling beliefs, we should expect on our models that signals ("ethnic markers") are more pronounced "along boundaries" (Boyd & Richerson 2005: 126) of interaction, which we need not restrict to geographic spatial boundaries. The logic runs that, since boundaries are the loci of the most interactions with people of other groups with distinct markers, this is where there is the greatest need for differentiation among the group markers or signals. Thus, we would expect group signals or ethnic markers, such as dialect or beliefs, to shift in one direction more radically, becoming more easily detectable as they become more pronounced.

If Boyd and Richerson's assumption is correct – that people prefer to interact with others who share "easily observable traits" (2005: 119) – we would expect to see exactly the phenomenon of group polarization, realized through the process of social comparison, presumably to enhance the resolution or disambiguation of the ethnic marker or signal. This assumption thus identifies a possible explanation for group polarization and phenomena of its ilk. It helps explain why people are motivated to contrast their social identities against the identity of the outgroup. It also explains and locates a more general purpose for theoretical posits like the *meta-contrast principle*[15] and *meta-contrast ratio*[16], which quantitatively predicts the value-point to which the group view will shift (McGarty *et al.* 1992: 3).

2.2.3 Shibboleth Signaling: Motivation & Adaptive Value

The motivation for the signaling function is definitionally to modify the behavior of others – the receivers. Specifically, beliefs function as shibboleth signals in order to modify R's behavior in a way that is fitness-enhancing for the signaler, S, and for the receiver(s), R.

---

[15] McGarty *et al.* 1992, p. 3: "the less a person differs from in-group members and the more he or she differs from out-group members, the more representative is he or she of the in-group…[and] defines what the group has in common *in contrast to other relevant out-groups*."

[16] (Ibid. 3), "obtained by dividing a person's average difference by his or her average difference from in-group member."

Perhaps we can begin to get some grip on understanding the fitness-enhancing virtues of shibboleth signaling by considering the puzzles which such health and climate beliefs present. Specifically, we consider the tragedy of commons style problems that COVID-19 and anthropogenic climate change present to society at large. What is puzzling about skepticism or disbelief in climate change or COVID-19 and scientifically-sanctioned preventive measures is that they both present a collective crisis common to all – liberals and conservatives, Republicans and Democrats alike. As Greene (2013: 91) describes the problem,

> You would think that, in the case of climate change, both Democrats and Republicans would be strongly motivated to get things right. After all, Republicans' stake in the earth's future habitability is no smaller than that of Democrats…Why then, do so many American political conservatives deny the facts on climate change, in apparent defiance of their own interest?

After all, in a commons-style problem, all parties involved have "skin in the game" or have a stake in the issue and an interest in resolving the problem, since all parties alike will bear a cost if the problem is unmitigated or unresolved (though that cost, depending on the social structure, may be distributed unevenly). The question then is, how is such cognitive behavior, which effectively denies there is a problem or downplays its danger, nevertheless sustainable, adaptive, or fitness-enhancing, as opposed to maladaptive and fitness-degrading? What is the payoff? What is the benefit of such epistemically and practically costly beliefs that would, at first blush, seem to go against an individual's self-interest? Commenting on the research of Kahan et al. (2010, 2012), Greene continues,

> …the key is to recognize that the problem of getting the facts right is actually a commons problem of its own, involving a tension between individual and collective self-interest. It is definitely in our collective self-interest to face the facts on climate change and act accordingly. But for some of us, as individuals, the payoff matrix is more complicated. Suppose you live in a community in which people are more skeptical about climate change—and skeptical about people who *aren't* skeptical. Are you better off as a believer or a skeptic? (2013: 91-2)

Similarly, we would think conservative Republicans and liberal Democrats would be equally motivated to get things right about COVID-19 (e.g., its transmissibility and mortality rate), the

effectiveness of recommended preventive measures (e.g., social-distancing, mask-wearing), and the safety and efficacy of COVID-19 vaccines. Yet, for one cultural and political tribe, the pandemic occasions another social space wherein shibboleth signals arise and flourish, with the social functions and benefits of these anti-COVID beliefs outweighing the costs of such beliefs and the benefits of believing truly – which can be substantial pragmatically.

Comparable to the shibboleth signaling function is Kahan's (2015, 2017) thesis of *Identity-Protective Cognition* (IPC). According to the IPC thesis, when people are stating their beliefs or views about anthropogenic climate change, evolution, and I assume COVID-19, what they are doing is not trying to state accurate beliefs that conform with the consensus of the scientific community and experts – their aim is not to believe truly or to express true beliefs. Rather, their aim is to express their social identity or cultural worldview. As Kahan (2017: 1-2) explicitly states, the "goal [of IPC] is protection of one's status within an affinity group whose members share defining cultural commitments." In other words, the goal of such motivated reasoning is to protect one's status as a member of a specific in-group *g*. The motivation for protecting this in-group membership, or one's status as an in-group member, greatly outweighs the motivation for holding epistemically respectable or virtuous beliefs about such subject matters. This differential in motivational sources is due to the simple fact that "most people have no reason to have a position on climate change aside from expression of their identity" (Kahan qtd. in Beck 2017).

As a signal, the payoff of such beliefs is their ability to modify the behavior of the receivers who detect them. While this includes benefiting the signaler, crucial to the selection for shibboleth signals is the benefit of the receivers as well. In many cases and contexts, it will be the case that the receiver does benefit by detecting the signal and modifying her behavior accordingly. Most notably, it will modify the receiver's behavior in terms of decision-making, specifically, informing with whom the receiver decides to coordinate and cooperate.

Beliefs as social signals provide a cheap and reliable social heuristic for population sorting in terms of "us" and "them" and determining who belongs to the same tribe or ingroup and who belongs to rival outgroups (Boyd & Richerson 2005). This heuristic modifies the receiver's behavior by helping her determine with whom to coordinate and cooperate. If she coordinates and cooperates with members of the same tribe or ingroup, her interaction will be more beneficial, i.e., have a higher net gain or payoff, than if she were to interact with people from other groups, say, from some neutral tribe or, even worse, from some rival tribe or outgroup, where either one participant or both would be motivated to defect and cheat, resulting in a sub-optimal interactions – at least in comparison with the net gain that would be likely if she were to interact with someone belonging to the same ingroup.

Moreover, such a heuristic for population sorting is an especially valuable adaptive tool in a socially complex environment where people must interact with strangers in order to attain their goals and achieve their ends. In a specialized and complex society, getting what we want or need, achieving our goals, often requires coordinating and interacting with others, usually through cooperation. This is the condition of *social dependency* (Sterelny 2012). As a result of this social dependency, we must engage with our social environment. Navigating our social space or social environment is facilitated by means of a sorting heuristic, which allows us to navigate our social environment more reliably, efficiently, and effectively. By means of the sorting heuristic, we can reliably discern or judge who is most likely to cooperate with us and is most optimal to interact with and thus with whom we ought to coordinate.

Underpinning and motivating this explanation is the assumption that social norms regulate social behavior such that people who share beliefs about "how one should behave" or what one should believe "yield higher payoffs than interactions among people with discordant beliefs" (Boyd & Richerson 2005: 119). When people share norms about how to act and what to believe, they are

less vulnerable to situations of social costs, less likely to make "social blunders," and thus are more likely to incur positive benefits through social interaction (ibid. 120). They are also able to discern who is less likely to defect or cheat and thus can judge with whom to coordinate and cooperate, enabling a general strategy for managing prisoner dilemma style interactions. There is also the fact that shared norms about coordination increase the overall optimization of the cooperation since both parties come to the table with similar views about the norms for divisions of labor and how to carry out the shared labor (Zawidzki 2013). Thus, having a clear, resolute, unambiguous signal or marker that is easier to detect functions as a general strategy or heuristic for social coordination and cooperation that helps to ensure optimal coordination for different parties as they can more easily discern both who are fellow group members and who are the more reliable and trustworthy group members and therefore least likely to defect.

Thus, the upshot of such a reliable and cheap sorting heuristic, which shibboleth signaling enables, is that it helps individuals manage prisoner-dilemma style interactions (Boyd & Richerson 2005, Ch. 7), both in one-off interactions (which may be less certain and pose greater risk) and in sustained interactions over time. It also is conducive towards judging with whom to interact and ensuring coordinated and cooperative successes.

If the foregoing is correct, then the shibboleth signaling function of belief is fitness-enhancing, both for receivers and signalers. By signaling one's tribal allegiances, one is more likely to attract interactants of the same ingroup, meaning they are more likely to cooperate and successfully coordinate with the signaler and thus the signaler and receiver will likely have an interaction with a higher net payoff. Importantly, the shibboleth signaling function of belief is fitness-enhancing for the receivers since it enables them to judge with whom they should interact, that is, with whom their interaction would result in the highest net payoff. Of course, the shibboleth signal also has the function of letting receivers know whom to avoid interacting with, who might be most likely to

defect or cheat in an interaction, or who would be less efficient or less coordinated in any interaction. This is turn benefits not only the receiver but also the signalers since their shibboleth signal discourages or staves off potential defectors or costly interactions and attracts those with whom interactions are likely to result in cooperative and coordinated success. In this way, we see how the selection for beliefs as shibboleth signals benefits or enhances the fitness of both the signaler and the receiver and thus has been selected for as a culturally and biologically adaptive function.

The shibboleth signaling function also takes place within the space of *coalitional psychology* (Cosmides & Tooby 1989). Contributing to group fitness, the shibboleth signaling function is adaptively valuable within a social environment comprising inter-group competition and coalition building. First, signaling group affiliation enables the formation and sustainment of large coalitions whose size in membership and protracted social and support networks allow the given coalition to outcompete other coalitions within the same social environment (Boyer 2018: 49). Second, signaling group affiliation makes is virtually impossible for other coalitions to recruit the signaler (ibid. 50), making it impossible, or at least extremely difficult, to defect or emigrate from one's own ingroup. The upshot is that a system of sophisticated signaling systems gives a coalition an advantage over others by allowing the emergence of increasingly complex coalitions, further satisfying the condition of social dependency.

2.2.4 Commitment Signaling

The second kind of signaling function which beliefs perform is the *commitment signaling* function. Whereas shibboleth signaling is selected for so as to communicate information about where the signaler's tribal allegiances lie, commitment signaling on the other hand is selected for in

virtue of its capacity to signal the strength of that tribal allegiance. In other words, commitment signaling communicates a person's loyalty to and dependency on her ingroup.

> **(COMM):** *x* is a commitment signal whenever *x* signals a willingness by S to incur potential costs (say, the standing of S's social reputation), often in relation to outgroup *o,* for the sake signaling commitment to ingroup *g*.

As is clear from the formulation of COMM, there are intra- and inter-sourced costly mechanisms that underpin commitment signals. First, something may be a commitment signal by simply displaying a willingness to incur a cost to oneself for the sake of or on behalf of the group. Ritualized actions, such as sacrificing one's resources, or undergoing some kind of privation (fasting, abstinence from sex and alcohol), or even bodily mutilations and tattoos may all be taken as signals of one's personal commitment to the group and which directly contribute to the group's wellbeing (e.g., sharing of resources) or indirectly, by showing solidarity, contributing to group cohesiveness, etc. and preventing free riders. This latter benefit is perhaps the more important contribution to group fitness: costly commitment signaling prevents Machiavellian freeloaders from infiltrating the group and procuring all the benefits of the group without paying any of the costs (Sosis 2003; Henrich 2009; Sterelny 2012; Zawidzki 2013). Thus, commitment signaling may enhance the fitness not only of the individual signaler but the group as a whole, as an aspect of coalitional psychology.

Second, while "commitment devices" or "commitment mechanisms" like sacrificing one's resources for the common good, fasting or abstaining from alcohol or sex, and engaging in costly, dangerous behavior such as group hunts (Sterelny 2012: 110-111, 116), do not require for their commitment signaling or commitment building the existence of rival tribes or outgroups, other commitment devices or mechanisms do presuppose the existence of a rival outgroups. Tattoos and bodily scars can function not only to signal tribal allegiance but also to block mobility between groups, helping to ensure that one is committed and loyal to her current ingroup or "social network" (Sterelny 2012: 120). For such markings to have any purchase as commitment signals, they require a

social environment divided between "us" and "them" and is subject to the forces of intergroup dynamics.

The claim is that beliefs too can function as commitment signals, that beliefs are another kind of commitment device or mechanism (Williams 2021). Absurd, extreme, contentious or inflammatory beliefs, in particular, such as those held by conspiracy theorist or extreme partisans, function in the same way as other commitment devices such as tattoos and bodily mutilation that result from group initiation rituals: "the insignia of membership and exclusion are public investment that change payoffs in the triggering context that everyone will recognize" (Sterelny 2012: 120). Here, Sterelny is specifically commenting on tattoos and initiation rituals that increase the cost of defecting from the gang (e.g., the "triggering context), making it possible for its members "to succeed only in the gang, and with its cooperation" (119). This is because "initiation rituals often involve high-price-defection from the norms of the larger society" or because membership insignia "mark [one] publicly and indelibly as a target of other gangs" (119-120). The logic then of such commitment devices is that they effectively cut one off from other social networks or groups by making the cost of defection too high to be fitness-enhancing, ensuring members' commitment to and dependency on their current ingroup.

Similarly, when we signal extreme partisan beliefs, or conspiracy beliefs, we are effectively making a public investment to our current ingroup by signaling beliefs that make it more costly in the future to defect from the current ingroup since such beliefs incur social reputational costs from outgroups or the general population. Outgroups or the general population may, as a result, perceive the signaler to be irrational, unreliable, and the signaler may become the object of public ridicule and scorn whom others outside the ingroup grow to distrust and view with suspicion. This in effect cuts ties with the social world beyond the ingroup, since those outside the ingroup are hesitant or

unwilling to cooperate with the signaler in light of her extreme beliefs. In effect, beliefs function as commitment signals like the moko tattoos of the Maori:

> By making the migration between groups more difficult, such signals make individuals within groups more trustworthy; they must hang together or die separately...The moko and similar examples work by decreasing mobility and thus increasing the agent's stake in his or her current network (Sterelny 2012: 120)

Extreme and conspiratorial beliefs likewise increase the signaler's being perceived as a trustworthy, reliable, and loyal ingroup member by making group defection or migration unfeasible or unviable and fitness-degrading in virtue of its high cost. Conspiracy theorists and cult members in particular, whose absurd beliefs make them highly visible membership insignia, mark themselves as "belonging to a…hated group, binding their fate to that of their group" (120).

The logic of such commitment devices is that "*costs are investments"* (Sterelny 2012: 110): an agent alters her environment so that adhering to her commitment to the group is the less costly and thus the optimal course of action. This investment in changing the environment ("niche-alterating") is a public investment that functions as a public signal and makes it known to other ingroup and outgroup members that the cost of defecting from the group for S is higher than the cost of remaining committed to the group. Again, this is often accomplished by some signal which makes group migration more difficult or unfeasibly costly, enhancing one's trustworthiness as a reliable and loyal group member. As Sterelny (2012: 111) writes,

> In advertising your origins and affiliations, you inherit your allies' enemies, whether or not you keep your allies' support. You are trustworthy because you have no other choice. So these signals constrain choice, and constrain it publicly…

In signaling beliefs that invite public censure and alienation and cutting ties with the outside social world, we enhance our trustworthiness as a group member by increasing the cost of group defection – we have no other option but to remain committed to the group and depend for our fitness and success upon cooperation with group. Presumably, this is a willingness to incur cost and sanctions such as ridicule and mockery from the general public or members of the relevant

outgroup, to invite their anger and scorn (such as achieved in trolling), to alienate and estrange themselves from the general public and the relevant outgroup as a symbolic gesture of "cutting ties"[17] with the outside world – the social world beyond the in-group. This signaling of severing ties with the outside social world shows the signaler's willingness to depend on the ingroup for any and all needs – a gesture of turning inwards toward the group, shedding the potential for any future engagement with the outside world, indicating a wholesale dependency on and commitment to one's ingroup. Just to enumerate a few of the potential costs commitment signaling communicates that the signaler S is willing to incur: cost to one's social reputation and prestige among outgroups, loss of professional opportunities, institutional sanctions and punishment, loss of family or friends, or even personal danger

Importantly, while commitment signals are costly, the cost is an investment, not a handicap. That is, the logic of the costly mechanisms of commitment signals is different from the cost of Zahavian handicap costs, such as the cost involved in signals of conspicuous consumption. Handicap costly signals communicate information about the signaler's quality, such as his genetic quality or wealth. Costly handicap signals "fall differentially on individuals of different intrinsic quality" (Sterelny 2012: 120). Since handicap signals are honest, hard-to-fake signals which indicate an agent's superior quality, it would make no sense for each group member, assuming most do, to engage in costly commitment signals since such signaling would fail to reveal information about a signaler's superior quality rather than information about the average group member. Instead of communicating information about the signaler's differential quality, costly commitment signals are communicating information about a signaler's trustworthiness and reliability as a group member, which is not a zero-sum dynamic like handicap signaling (e.g., peahens choosing the superior

[17] This is comparable to the "burning one's bridges" strategy, which makes group migration a foregone choice (Boyer 2018: 50; Mercier 2020).

peacock A as a mating partner means less or no mating partners for the inferior peacock B). Yet, everyone can commitment signal without loss in the quality of the signal because increasing the trust in signaler A does not deprive or mean less trust for signaler B.

A caveat, however, is that some extreme or conspiratorial beliefs may be the result of trying to signal differential quality in the signaler's quality as a group member. I have in mind here beliefs that are formed as the result of the ramping up phenomenon, where group members are in an arms race to outcompete each other to see who can hold the most moral beliefs or the beliefs that most represent the group, or who signal that the one is the most outstanding and unimpeachable ideal group member. In such a case, belief signals may be used to try to signal one's superior ingroup qualities. Yet the strategic costliness of such beliefs seems to be the same as public investments in the ingroup, which make it unfeasible to defect or migrate in the future. We might, however, think that such public investments are a kind of social handicap – handicapping our prospects of social mobility or migration.

As always, while some belief or action may enhance one's social reputation relative to one social group, it may also simultaneously diminish one's social reputation relative to another group. Organic food consumption, for instance, provides a case study for how costly signaling may enhance the signaler's social reputation yet also harm or diminish the signaler's reputation, depending on whom the audience or the receiver of such a signal is. Relative to an audience of receivers who are progressive and green-minded, are sensitive to environmental issues and food health, the costs of buying organic food can purchase social prestige and enhance one's social reputation (Puska 2019; Luomala et al. 2020). To others however who are not as green minded or environmentally conscious, such costly behavior may only serve to diminish one's social reputation, as the receivers of such a signal judge the person to be self-righteous and morally-posturing, to be insincere, and impugn other negative character traits to the signaler (van de Grint, Evans, Stavrova 2021). Thus,

whether the cost of consuming organic foods pays off for a person socially will depend on who one's audience is, that is, who is detecting, interpreting, and evaluating the costly signal and who S's ingroup is.[18]

2.2.5 Commitment Signaling: Motivation & Adaptive Value

Like the motivation for shibboleth signaling, the motivation for commitment signaling is the payoff in terms of social benefits. Despite the costs involved in commitment signaling, it compensates the signaler by securing access to the benefits of group membership. For instance, commitment signaling, despite potentially diminishing the signaler's social reputation and prestige in its evaluation among outgroup members and the general public, can enhance and aggrandize the signaler's social reputation and prestige relative to her ingroup by expressing her loyalty and commitment to the group and increasing her trustworthiness as perceived by fellow ingroup members, thus centralizing her place among the fold and securing her status as an ingroup member. Commitment signaling can ensure that one is securely and safely ensconced within the group and that the status of her membership within the ingroup or "affinity group" is protected (Kahan 2017). Commitment signaling then helps secure the signaler's access to the group benefits of group membership, such as access to group resources, shared problem solving and shared labor to common problems, successful coordination and cooperation with interactants, increase in potential cooperators, professional opportunities, and greater selection of and access to mating/romantic partners.

[18] The costly signaling of organic food, however, is perhaps a Zahavian handicap signal, since only those who actually have the spare income can afford to purchase organic foods, or those who actually care about food health and the environment can be bothered to spend the time, effort, and money on purchasing organic food. Yet, as organic foods become cheaper and more readily available at stores, their structure as costly signals may shift from a handicap signal to a public investment in one's social reputation by signaling commitment to a certain movement.

Again, if the foregoing is correct, then the payoff of commitment signaling, despite its potential cost, is that it is socially fitness-enhancing. And again, the social fitness-enhancement here may be thought of as prior to or having structural or causal primacy which grounds, causes, or realizes the other kinds of fitness-enhancement, such as biological as well as personal, psychological, and hedonic. For each of these standards of fitness, the enhancement of social fitness has causal or historical primacy.

And like shibboleth signaling, these benefits are rendered or made possible by the fact that commitment signals modify the receivers who detects such signals, both the behavior of the ingroup and the outgroup. For members of the general public or relevant outgroups, detecting the commitment signal modifies their behavior such that it potentially imposes some cost or sanction on the signaler; thus, modifying R's behavior such that it forms the basis of some potential strategic cost is a necessary condition for belief *b's* being a commitment signal (at least for costs which require an outgroup). This potential to modify the receivers of the relevant outgroup or general public in a way that is costly for signaler S is the basis for modifying behavior of ingroup receivers in a way that is adaptively valuable or fitness-enhancing for S*,* by ensuring receptivity to cooperation with ingroup members.

Unlike shibboleth signaling, where it is the receivers and the signaler who benefit from the signal, in commitment signaling, it is the signaler who is the primary beneficiary of any fitness-enhancing consequences of the signal. This is not to say that commitment signaling does not benefit the group collectively or their individuals. Such doxastic signaling practices can enhance the fitness of the group by contributing to group cohesiveness, loyalty, and trust, improving morale of the group members and assuring them that each member is committed to the group. Costly belief signals like commitment signaling (and ritualistic belief signaling) may also earn their keep insofar as they contribute to the fitness of the group by deterring freeloaders and Machievellian invaders which

would be a detriment to the group, since they would deplete the resources without giving any in return or compensating the group's loss somehow. For those who are not actually committed to the group and willing to incur any costs for the group, beliefs which likely invite costs from the general public, which possibly invite censure and sanctions from the general public and other outgroups are not a stable social strategy and are too costly – or at least are perceived as too costly – by freeloaders who are not actually committed to the group and willing to incur a cost. Also, the commitment signaling cost ensures that these signals are honest, since by holding such beliefs – signaling that one holds such beliefs – one actually incurs some cost on behalf of the group and increases the cost of defecting from the group so that such defection is not a viable option. Furthermore, the investing costs and change in payoff of group defection will increase the perceived, subjective cost of commitment signaling for those who are not actually committed to the group – at least perceived as a higher price to pay than to those who are actually committed to the group. The crucial feature is that commitment signaling is a mechanism for ensuring the signal is honest, since it makes the signaler vulnerable to potential costs just in virtue of having the belief and making it public, which may be required for even partaking of any of the benefits of group membership.

To be clear, there must be some cost paid publicly by signaling the belief in order for the belief to signal commitment since it is via the costliness that the belief signals commitment. In other words, the commitment-signifying aspect of the belief signal depends upon the cost of signaling that belief such that *x* signals commitment *only if x* implements some public cost. If, for instance, I believe that the COVID-19 vaccine is a ruse for Bill Gates implanting the population with microchips, I engender various sorts of social costs. First, I likely lose a significant amount of credibility among those outside my ingroup; second, my rationality and ability to track reality is called into question; third, I likely open myself up to public ridicule and censure; fourth, the pool of potential cooperators diminishes; finally, and following from the first four kinds of cost, I have

constricted myself from migrating to other social networks, making the prospect of group defection too costly to be a viable option. I thus signal my willingness to "ride or die" with my ingroup – my dependency on and loyalty to the ingroup since I have severed my social ties with the outside social world and occluded opportunities within other social networks. (Of course, we may undertake a protracted regimen of reputational rehabilitation which rebuilds these formerly severed social relations).

Thus, we have seen how conceptually, in terms of their cost-benefit structure, shibboleth signals and commitment signals are distinct kinds of social signals, not only in the information they communicate but how they modify the receiver's behavior and who the intended or target receiver is whose behavior is to be modified.

2.3 Doxastic Signals: Are Signals Beliefs?

As many theorists have noted, having to fake a signal, such as a belief, is often more costly than honest signaling (Sterelny 2012), i.e., what one honestly believes. It is cheaper to honest signal that one is an expert at *x* than it is to fake signal one's expertise at *x*. If, for instance, I am a wine connoisseur, since I already possess the requisite knowledge, it is cheaper for me to signal – as a byproduct of my expertise – that I am a wine connoisseur than it is for someone who deceptively signals that she is a wine connoisseur and must therefore spend time, energy, and resources looking up and researching information about wines and the wine industry. This dynamic is what Sterelny (2012: 111) describes as "an automatic cost asymmetry between honest and deceptive signals," continuing with the example of an expert birder compared to a fake birder:

> It is easy and cheap for me to present credibly as an Australian birder, because I am an Australian birder. The resources needed to make my birding credentials credible have been assembled automatically, as a side effect of my birding, and independently of any need to signal that characteristic…suppose someone with no birding skills wanted to credibly impersonate being an Australian birder. That would impose high signal costs. They would

> have to outlay funds on peripherals: optics, field guides. They would have to spend time and energy acquiring at least a passing knowledge of the birds, especially if they could not rely on an utterly ignorant audience. It is possible to fake birding interest and expertise, but it would be expensive. It is cheap for me and expensive for a fraud. If the profit of deception is low, this asymmetry in itself might suffice to keep signals honest.

Comparable to the high costs of sending fake or deceptive signals about one's expertise or interests, signaling fake beliefs is likewise typically involved and expensive, undercutting the optimal payoff matrix. Having to fake a belief signal and engage in constant impression management and self-monitoring is cognitively costly: the subject must devote extra cognitive effort and resources to competently signaling the fake belief, maintaining and managing the presentation of this fake belief signal, as well as preventing her actual belief from leaking. This correspondingly requires constant vigilance, self-monitoring, and self-censorship, curtailing any behavior that might leak her actual belief. Furthermore, there is the risk of accidentally leaking her belief whenever under cognitive load such as performing demanding tasks, running the risk of letting down her cognitive guard. The cost, moreover, of being found out as a fraud presumably risks being subject to the same kinds of costs or punishments as those inflicted on group defectors in order to deter or discourage future fraudsters. Such punishments include group expulsion and social alienation (meaning others are unwilling to coordinate and cooperate with the fraudster), the reappropriation of her resources and other means of social sanctions.

So, it is cheaper and thus a more optimal payoff to signal honest or genuine beliefs – to believe what one signals as her belief – as opposed to the recurrent and risk-prone cost of sending fake belief signals, in which one must constantly look over their proverbial cognitive and behavioral shoulders, monitoring her behavior so that it reflects or is congruous with the fake belief, constantly tracking the social context to make sure she expresses the appropriate belief, etc. It is much cheaper to internalize the belief so that we believe what we signal we believe.

This touches upon a longstanding debate within philosophy of psychology and the philosophy of belief over the thesis of *Doxastic (In)Voluntarism* (Audi 2001; Chignell 2010; Funkhouser 2003), pronounced most forcefully perhaps by Bernard Williams (1973) in his claim that it is a conceptual truth that we cannot believe at will (in contrast to Hume who claimed it was only a contingent restriction on our belief forming practices).

For Hume, our inability to form beliefs at will is a contingent matter of psychology (1993: 37): it is a psychological fact about humans that we are unable to voluntarily form beliefs, being passive agents to their occurrence. This fact, therefore, requires a psychological explanation. For Williams, in contrast, the fact that we cannot form beliefs at will is a necessary fact, based on his conception of belief. It is not a psychological fact but a fact intrinsic to the concept of belief. More strongly, besides just a conceptual fact about what it means to be a belief, some claim it is a metaphysically necessary fact about belief, given the nature of the attitude itself (cf. Papineau 2013). It has nothing to do with our psychology, the limits of our psychology, or any contingent psychological constraints. If we could form putative beliefs at will, then such intentional attitudes would, *ipso facto*, fail to be genuine beliefs.

Bracketing the conceptual and metaphysical questions of whether we necessarily cannot form beliefs at will, given the concept of belief and the nature of belief itself, let us consider the psychological, i.e., empirical, grounds for denying or affirming that we can form beliefs at will. An important aspect of Williams' claim is the qualifier that we cannot form beliefs at will *in full consciousness.* As Funkhouser (2003) notes, even if it is psychological fact that we cannot form beliefs at will that "(a) we know (or suspect) to have acquired at will," (b) in full consciousness as a result of (c) "*direct* (i.e. unmediated) willing" (the conjunction of which Funkhouser calls the *Williams Thesis* and which is considered "necessarily unattainable"), this does not preclude the possibility that we may form belief *b* at will by either (1) forgetting or being unaware we willed *b,* (2) knowing we willed

b, but *b* is not occurently "before [our] consciousness," or (3) we are conscious of belief *b* which we formed via biased belief-forming processes, e.g., psychological mechanisms, motivated reasoning, or discounting of evidence (180-1).[19] Possibilities (1)-(3) turn on the distinction between *synchronic* and *diachronic* belief-formation processes. While we might not be able to form beliefs at will in full consciousness synchronically, it is nevertheless possible that we can come to believe some proposition at will over time via the routes delineated, especially (1) and (3).

If the principal attitudes involved in *wishful thinking, self-deception, delusions* are genuine beliefs, such phenomena mark diachronic belief-forming processes. Presumably such beliefs are acquired as the result of motivated reasoning which modulates how we interact with evidential information (e.g., giving undue weight to hypothesis-confirming evidence while avoiding, discounting, or downplaying counterevidence, or construing non-relevant information as evidence – information that does not in fact furnish the agent with epistemic reasons for believing *that p*[20]) or implemented through other motivational mechanisms. This can be achieved through various biases or motivated strategies of belief formation. This includes the processes involved in *confirmation bias*, for example, in which we seek out or attend to information which confirms our antecedent belief or hypothesis while eschewing or ignoring information that disconfirms it (Nickerson 2016). When we are motivated to believe *that p* because we desire to believe *that p,* we tend to subject favorable arguments to less scrutiny and accept uncritically information which is hypothesis-confirming while tending to vigilantly scrutinize counterarguments and assume a hyper-critical stance to any information that is

[19] Funkhouser delineates these possible conditions as (1) Forgetting (2) Submerging and (3) Biasing. Since I do not have the space to review Funkhouser's full treatment of the Williams Thesis, I outsource that work to his (2003) paper, "Willing Belief and the Norm of Truth."

[20] This treatment of mere information as evidence for a belief prompts the distinction between having evidence for a belief and having *good* evidence. This distinction informs the further distinction which divides *rational* beliefs from belief *simpliciter:* rational beliefs are *responsive* to evidence, having the right kind of evidentiary fit, while beliefs *simpliciter* may be merely sensitive to evidence (Bortolotti 2010), where this involves updating one's belief to some information *i* that does not conform to epistemic norms and/or which involve treating any information as evidence for a belief, irrespective of whether it satisfies the appropriate epistemic norms.

dis-confirmatory or potentially undermining of our desired belief. These distinct processes correspond to the division between *epistemic bubbles* and *echo chambers* (Levy 2021a): in epistemic bubbles, we lack the appropriate evidentiary inputs while in echo chambers we tend to give more weight to fellow ingroup members as informants and evidentiary sources as well as to evidence which supports our prior beliefs and values.

Another motivational mechanism implemented in the willful formation of belief is the diachronic process of biased memory retrieval. Our desires to believe a given hypothesis *h* can bias how we interact with and retrieve memories, recalling more easily and more frequently information which confirms the desired hypothesis. Motivational factors, including conative and emotive states, also influence how the mind structures and organizes information in memory, the associations and links we form among the content of memory, and the compartmentalization of information in memory, which in turn affects how we engage with evidence in working memory and the types of local inferences we make (Bayne & Pacherie 2005). Through biased memory retrieval, we can plausibly willfully form beliefs by attending to certain kind of evidence and drawing certain inferences, thus implementing such formation through biasing processes, or through the condition of forgetting that we willed a certain belief which we now hold in full consciousness.[21]

Assuming that doxastic accounts of self-deception, delusions, and wishful thinking offer the correct theoretical model, we have prima facie grounds for supposing the possibility that some beliefs are formed willfully – albeit unconsciously and overtime. Furthermore, by identifying the underlying psychological mechanisms, we increase the plausibility of our explanation as to how it is such belief-formation processes are possible. Such an account also offers another strategy for

[21] Some may worry that memory is a factive relation and thus cannot do the work we have described here. This worry is moot, however, since nothing in the description assumes these are false or confabulated memories; it's merely that the subject interacts with her memories in a biased – typically motivationally – manner which makes it possible for her to willfully form beliefs along one the possible routes delineated by Funkhouser (2003).

undermining the rationality constraint of belief and throwing greater doubt on the claim that belief is essentially constituted by a norm of truth or essentially aims at truth.

The purpose of the foregoing has been to sketch the conditions under which agents can plausibly form social beliefs, such as beliefs which function as social signals or constitute an agent's self-conception and social identity. That is, it provides a plausible story as to how we come to *internalize* a mental state such that it becomes a belief. Even if social beliefs are not strictly speaking willfully formed, they are similar in important and salient ways by virtue of their motivational component. While willfully formed beliefs are motivated by an agent's will or desire, social beliefs are motivated by the mechanisms of social pressures, social comparison, and social identity, which contributes to our self-conception. Thus, social beliefs, like willful beliefs, are psychologically motivated and plausibly form two different subclasses of motivated beliefs which share great proximity in the doxastic space.

Lastly, I would like to mention a possible route for forming and maintaining motivated beliefs based on a *metacognitive* account of delusions (Currie 2000). Currie's account claims that delusional patients are mistaken in their second-order beliefs about the belief status of their first-order attitude, mistakenly believing that that they believe the delusional content whereas in fact they merely imagine the first-order delusional content. Although I am not endorsing their anti-doxastic conclusions about delusions actually being states of imagining, an interesting consequence of their view is that we may come to form the first-order belief *that p* – which may have started out as some other kind of mental state, e.g., imagining – because of our (mistaken) second-order belief. Thus, a subject's second-order belief can ground and determine her first-order beliefs without such a belief being formed in response to good evidence or obeying the other rationality constraints (Bayne & Pacherie 2005).

Applying such a metacognitive dynamic to social beliefs (or motivated beliefs more generally), this is possibly another route (though not the only) by which we form motivated beliefs, like social beliefs, through second-order beliefs whose content is that we believe that we believe *that p,* thus in fact causing or determining the first-order belief. Regarding social beliefs, when we consider our self-conception as a member of a certain group, we form second-order beliefs about what we believe *qua* member of that group. For instance, when I consider my self-conception as a liberal Democrat, this self-conception, through social comparison with other liberal Democrats and their messaging, can lead me to form the second-order belief that, as a liberal Democrat, I believe that I believe that climate change is anthropogenically caused and an imminent existential threat. I invite the reader to consider applying this model to other real-world examples.

Proposing various models which explain how we form motivated beliefs, such as willful beliefs or social beliefs (assuming they are distinct in kind) lends plausibility to the claim that these attitudes are genuine beliefs and not some other king of mental state proximal to the doxastic genus. This plausibility is garnered by offering potential models that would explain and show how motivated beliefs could be formed given our extant psychological equipment and identifying the psychological mechanisms so implicated.

In sum: one objection to the signaling model of belief is that S is merely expressing or avowing to believe *that p* – an expression which does not correspond to any such belief *that p* for S. Response: in terms of cost-benefit analysis, it is more beneficial or more optimal if the signaler in fact believes the belief that she signals. Sending a deceptive signal, i.e., a fake or imposturous belief, is more costly than sending an honest signal, the content of which the signaler actually believes. This is because sending a fake or deceptive belief signals requires the sustained cognitive task of making sure one doesn't accidentally signal or reveal what she actually believes, which in turn requires keeping track of one's discursive public contexts and knowing which belief to signal in a given

context; it requires constant vigilance to make sure one does not accidentally and unintentionally leak what one truly believes. Lastly, it also runs a greater risk and potentially incurs greater cost if one is revealed to be an imposture, which can potentially induce costs and punishments comparable to those levied against group defectors. Within the social space, *defectors* and *fraudsters* occupy the same vicinity. For these reasons, the optimal strategy is to *internalize* the beliefs of the group – transmitted via group messaging – including those beliefs which perform a signaling function. Again, none of this implies that the strategy is a conscious one or that the signaling agent is even conscious of such motivational pressures; the mechanisms presumably operate at the subpersonal level, such as intrinsic psychological motivations to imitate group models and conform to social norms (Zawidzki 2013) and reasoning guided by motivated biases (Nickerson 2016), which are automatic, unconscious, and outside rational control.

2.4 Does Propositional Content Matter?

What the belief expresses semantically might perhaps be ultimately irrelevant. Not only might it be irrelevant since the truth-value of the proposition is irrelevant, but it might also be irrelevant insofar as the propositional content that does the signaling is often determined by "out-group enmity" (Hornsey & Fielding 2017), such that beliefs which become defining of ingroup *g* are determined arationally, in reaction to what rival out-groups happen to believe. If rival out-group *o* believes that vaccines are a public good and essential for public health, then ingroup *g* will tend to believe that vaccines are a public vice.

Beliefs that become defining of a group and function as shibboleth signals are often not, contrary to common assumptions, determined by any antecedent, conceptually prior or more basic ideological beliefs or values. The propositional content of the shibboleth belief is rather like a placeholder whose principal function is to mark group membership, to facilitate population sorting,

and is not grounded in any prior ideological commitments, except for when individuals have no other information to use, say, when group messaging is not available or accessible, suggesting what they should believe. When lacking other informational cues, group members must rely on antecedent knowledge of the group's ideological commitments and values to determine what to believe with respect to some proposition or issue (Cohen 2003).

This last claim is empirically suggested by studies in sociology and political science which look at how group messaging determines a person's evaluation of specific policies. Instead of being motivated to accept or reject a scientific hypothesis, or endorse a legislative policy, based on shared ideological principles that are "diagnostic" or "definitive" of one's reference group*,* many of our political and scientific beliefs, especially those embroiled in culture wars and public controversies, are formed principally in response to "group influences" or group messaging, communicating the position of the group about a particular issue (Cohen 2003: 808, 819).

In Cohen (2003), for instance, partisan Democrats and Republicans were presented with welfare policies and asked to evaluate them for fairness and other moral virtues. While the welfare policy itself was held constant, what was manipulated is who was purported as proposing such a policy, whether it was Democrats or Republicans, thus manipulating the partisan messaging variable. By now the results should be unsurprising: when participants thought the policy was proposed by their same political, partisan tribe (Democrats, Republicans), they were more likely to positively evaluate the welfare policy as fair, balanced, and justified than when led to believe that the same welfare policy was proposed by their rival political tribes. Cohen also found that group messaging about what to believe (i.e., whether it was proposed by one's own political party) took priority or trumped any ideological inferences or inclinations about what to believe regarding the welfare policy. It was only when there was a lack or absence of social cues or group messaging about what to

believe regarding the presented policy did participants fall back on the ideological commitments they antecedently took to be "diagnostic" of their "reference group" (Cohen 2003: 809).

Such findings suggest that group messaging sometimes functions not to communicate the values or ideological commitments of the group but rather to communicate the position of the group as a value point for others to signal where their tribal allegiances lie. Shibboleth signaling is about signaling that one holds beliefs or values consonant with the group; it's about signaling that one is a member of the group. Furthermore, as this study suggests, group messaging and signaling can as much determine the position or beliefs of the group as antecedent ideological commitments can. Ideological determinants of the group then can plausibly form post-hoc, after the belief has already become a marker of group membership, meaning that some of the ideological commitments of a given group are determined by what the group's position happens to be, which runs counter to how we usually frame the causal order. That is, it is plausible that many ideological commitments and values defining of a group are determined by the content of group-messaging and social signaling, such that the causal chain or direction of fit runs from the content of the belief signal or group message to determining or informing the content of the group's ideological commitments (cf. Achen and Bartels 2016: 223). Thus, the ideologies of the group might be determined post-hoc, or post-hoc become internalized and integrated into the group's identity after such content has functioned to merely communicate social-identity or what ingroup a person is a member of.

§3. Tribal Beliefs – A Character Sketch

For many of our beliefs, there is an undeniable "social life" (Funkhouser 2017: 809), in which beliefs, as cultural items, take on and cultivate robust social roles (Boyd & Richerson 2005; Mercier 2020). The social roles of belief can contribute to the biological and cultural success of individuals who believe *that p* whenever the belief *that p* is socially strategic or valuable to hold.

Signaling certain beliefs has the potential to enhance, rehabilitate, or protect S's social reputation, including S's status as a rational and reliable agent of coordination and cooperation (Mercier & Sperber 2017; Zawidzki 2013; Mercier 2020). The social function and social utility of certain beliefs can make them also personally and biologically valuable to acquire to the extent that they secure access to the benefits of group membership, such as access to mating partners, nutrition, childcare and medical care, shared labor, as well as contributing toward a positive self-conception and self-esteem (Tajfel & Turner 1986). Humans depend on social support as a vital mechanism for fitness, given the extent to which human activities involve and human needs require sophisticated and protracted coordination and cooperation (Sterelny 2012), including "hunting, trade, defense against enemy groups, and probably shared parenting" (Boyer 2018: 43)[22]. For most of us, it is only through social support, collective action, and coordination and cooperation that we have access to many quotidian practical benefits, including financial gains, professional opportunities, and even romantic opportunities. Such access and procurements are possible only through robust support networks and solving problems of coordination and cooperation.

To gain some perspective, consider a similar species of belief that shares proximity in the doxastic space with tribal beliefs, notable for their psychological adaptive value: *positive illusions* and *optimistically biased beliefs* (Taylor & Brown 1988; Bortolotti 2020). Our motivated "tendencies to adopt and maintain positive beliefs about ourselves" and make overly optimistic predictions about our prospects, despite lacking good evidence or even possessing evidence to the contrary, often prove to be psychologically adaptive, e.g., sustaining psychological health, preserving a positive self-conception and self-esteem, which enables us to continue engaging with the world so as to achieve

[22] *See also* Sterelny (2012: 101): "No living humans gather resources needed for a successful life by their own efforts. None do so even with the help only of their immediate family." Thus, I assume in this paper a basic constraint on human life and existence is the need for social support networks to access group benefits. This is the constraint of *social dependency*.

our goals and aims (Bortolotti 2020: 109). The self-protecting or self-enhancing functions of these beliefs allow individuals to continue engaging in goal-oriented behavior, making it both practically possible and more likely that they will attain their goals (or some subset thereof) than if they just became mired in accurate reflections and assessments on their inadequacy, became depressed, or despaired at the prospects of achieving their goals, abandoning such endeavors all together (Katz 1960; Sedikides & Alicke 2012). Such behavior would not only degrade psychological fitness but also other kinds of fitness too, including biological.

Comparable to the adaptive value of positive illusions, many of our beliefs are motivated by social aims, such as securing social benefits and avoiding social costs, and are adaptively valuable in virtue of their performing such social functions. The belief that the COVID-19 vaccine is unsafe or dangerous, for instance, might pay higher social dividends and incur greater social benefits than the cost of not getting vaccinated *if* (1) S believes that she herself has a low probability of contracting the COVID-19 virus (or that there is no such virus) or has a low probability of dying from it (because of her age and good health), and (2) if S's cultural models and leaders, friends, social network, community, neighbors, parents, teachers, influencers, etc. are also skeptical of the COVID-19 virus and vaccines. The obtainment of (1) and (2) modulates, for S, the perceived costs of not getting vaccinated or not wearing a protective mask. It should be noted, furthermore, that this cost-benefit analysis need not be a deliberate or conscious strategy. Nor do social reasons for belief imply impression management, whereby the psychological subject is consciously and deliberately self-monitoring and managing her impression on others (Paulus & Trapnell 2008). Rather, in cases where we are talking about genuine belief, such strategies as conscious impression management and deception are inconsistent with genuine beliefs (though over time S may come to actually believe the content which she deliberately yet disingenuously broadcasts (Hogan 1983; Funkhouser 2003)).

As a category, *tribal belief* describes beliefs that are sensitive to social norms, culturally acquired and transmitted, fixed in virtue of social reasons, and motivated by social goals. Tribal beliefs, for instance, are sensitive to messaging cues about what members of the in-group are supposed to believe. They also communicate important social information about a person's tribal allegiances and about how a person identifies with respect to social categories. Tribal beliefs are also capable of communicating information about the strength or degree of those allegiances, that is, how committed people are to their ingroups, including how central the ingroup is to a person's self-conception and social identity.

The importance of such information is that in novel or complex social environments, receivers of the belief signal can thereby infer or discern what social groups the signaler is a member of and calibrate how committed the signaler is to her respective ingroups. Equipped with such socially strategic information, receivers can determine whether to interact and cooperate with the signaler, preferring to interact with individuals of the same ingroup or tribe, with whom the receivers share greater social similarity or proximity in the social space. Shared group membership is an indicator of important shared social similarities (shared social norms, values, customs and etiquette, narratives, etc.). Interacting with those who are more like us socially – who share the same social norms, customs, and values, etc. – tends to result in interactions with higher net gains or benefits for all parties since they come to the table with similar expectations, similar background assumptions, as well as motivated ingroup biases (Tajfel & Turner 1986) that encourages prosocial behavior that enhances coordination and cooperation. This is further reinforced by social norms and "socially distributed mechanisms" like cultural "pedagogy" and "norm cognition and enforcement" such as sanctioning or punishing those who violate social norms (Zawidzki 2013: 51,61). Jointly, these components encourage the proliferation and selection of prosocial behavior among group members, solving problems of coordination and cooperation, and enabling the group to outcompete or

outperform other, less coherent and less efficient groups within a given social niche (Zawidzki 2013: Ch. 2; Boyd & Richerson 2005: Ch. 6).

Tribal beliefs cry out for explanation beyond the resources of traditional models of belief, demanding we consider the fitness-enhancing virtues of some beliefs not just in terms of navigational or epistemic capacities (except perhaps metaphorically as in successful navigation of our social space) but specifically in terms of social capacities. Furthermore, if beliefs are essentially connected to truth because of selective pressures, as some philosophers contend, then many of our beliefs would seem systematically and persistently fitness-degrading, gaining traction among the populace despite their maladaptive effects or value. It is not just that such beliefs are defective or fail to even count as beliefs according to Veridicalist criteria – since they fail to navigate the world and track reality – but such beliefs often impose a (epistemic) cost or handicap on the cognizer. The *strategic* costliness of beliefs such as those involved in religious rituals (Sosis 2003) or public investment commitment signaling functions (Sterelny 2012; Zawidzki 2013; Mercier 2020) is a feature which distinguishes them from mere unsuccessful sperm cells or defective hearts, blocking potential attempts to analogize the two cases. The analogy, however, presuppose that in different contexts success conditions or criteria are held constant for what appears in fact to be fundamentally different doxastic functions. Hearts and sperm cells are not selected for in virtue of their defectiveness; their defectiveness is merely an incidental side-effect of failing to perform their selected-for function. Hearts which fail to pump blood or sperm cells which fail to fertilize an ovum are not assessed by some other alternative criterion which is grounded in some other proper function they perform – at least it is not clear that there is some other function selected for which hearts and sperm cells are capable of performing and which are adaptively valuable.

Moreover, unlike defective hearts and wayward sperm cells, whose failure to pump blood or fertilize ova do not incur some alternative utility or benefits or satisfy alternative success criteria,

some tribal beliefs depend for their adaptive value, specifically their social utility, upon their epistemic or cognitive costs, that is, upon their epistemic delinquency. These epistemic or cognitive costs consequently incur costs for a person's social reputation among rival outgroups or among the general public, which enables individuals to signal their willingness to both incur a cost on behalf of the group and, more importantly, a willingness to cut ties or "burn bridges" with the social world beyond the ingroup (Mercier 2020; Sterelny 2012). Thus, in one context, where one set of norms and evaluative criteria predominate or are made salient (e.g., epistemic or veridical norms), tribal beliefs may be assessed as defective, yet in another context, such beliefs may be assessed as correct or successful by a different set of norms and evaluative criteria, say, a utility criterion. Beliefs which perform a *commitment signaling* function, for instance, are often epistemically, socially, and even practically costly for the signaler, but it is in virtue of these costs, that the signaling belief ultimately yields a higher payoff in returns of social benefits and utility relative to S's ingroup (and thus in such a context would be judged correct by a utility criterion).

Other tribal beliefs (like shibboleth signals) might be epistemically innocuous with no implication of or dependency on being epistemically costly or defective, being selected for simply in virtue of performing an extra-epistemic function of demarcating group membership. Beliefs within this subclass may even be true and held for good epistemic reasons relative to one group yet held for mere social reasons relative to another group (or for some individuals within the same ingroup). If, however, we extend our model of belief to include proper functions like social signaling, incorporating the concept of commitment signaling which requires some potentially costly display, many beliefs that seem defective, maladaptive, or fitness-degrading are in fact fitness-enhancing and valuable by performing an expressive or communicative function by virtue of their costly signaling capacities (Sterelny 2012: 116; Funkhouser 2017: 817; Sosis 2003).

Tribal beliefs then are adaptively valuable (biologically, culturally, etc.) insofar as they are selected for in order to provide people with a cheap (i.e., low cognitive effort and low resource cost) and reliable heuristic for sorting populations into their respective social groups or tribes – "us" and "them" – whereby social agents sort the population into potential *cooperators* and *defectors*. The upshot is that tribal beliefs help us, among other social situations involving coordination and cooperation, manage prisoner dilemma style problems – both in sustained, iterated interactions or in one-off interactions – in a suitably complex and diverse social environment where we must interact with strangers (Boyd & Richerson 2005; Axelrod 1995). Having a reliable heuristic for discerning with whom to interact contributes both to the success of individual group members (individual fitness) and to the success of the group collectively (group fitness). To the extent that groups with a high degree of cohesiveness and "are better organized and better able to exploit scarce resources" tend to "outcompete other groups" (Zawidzki 2013: 239) due to more effective coordination and cooperation among its members, individual prosocial behaviors and effective signaling systems ground the fitness of the group, whereby selection of the former (individual prosocial dispositions and effective signaling systems) contributes to selection of the latter (group level selection).

3.1 Motivational Schema

What can we glean from the foregoing discussion? First, we can begin to generalize the cost-benefit relations that help determine which beliefs we are motivated and likely to form:

(i) If the social cost, for S, of believing *that p* is greater than the epistemic benefit of believing *that p*, then S is likely to not believe *that p* (or believe *that not-p*)

(ii) If the social benefit, for S, of believing *that p* is greater than the epistemic cost of not believing *that p* (or believing *that not-p)*, then S is likely to believe *that p*

(iii) If the social cost, for S, of not believing *that p* (…) is greater than epistemic benefit of believing *that p*, then S is likely to not believe *that p* (…)

(iv) If the social benefit, for S, of not believing *that p* (…) is greater than the epistemic cost of believing *that p*, then S is likely to not believe *that p* (…)

A few clarificatory works are in order. First, as *likely* signifies, these are probabilistic claims, not logical guarantees. The causal connections will need to be cashed out in terms of likelihoods, tendencies, or dispositions with an "all things considered" clause. Inferring what a subject believes is always a matter of inductive or probabilistic inference vulnerable to exceptions. What a psychological subject believes ultimately is contingent upon a myriad of other factors, including other cognitive, emotive, and conative states, cognitive impairments, individual need for cognition and other psychological idiosyncrasies, etc. that potentially contravene the predicted belief, making it doubtful that any lawlike statements could guarantee their consequent necessarily. Hence the presumption of a *ceteris paribus* clause. So, the conditionals here are intended to capture the empirical generalizations that the antecedents are likely sufficient to bring about consequents.

Second, and building on the first point, the sufficiency condition is best understood as *causal* sufficiency, since the antecedent captures the motivational reasons that are (likely) psychologically causally sufficient for S believing or not believing *that p*. Ultimately, however, I am neutral for present purposes about how best to construe the relation between the antecedents and consequents, that is, how tight or how loose the connection between them ought to be conceived.

Third, by *epistemic benefit* I mean something approximate to Bortolotti's (2020: 14) characterization: a belief is epistemically beneficial "if it makes a positive contribution to the agent's epistemic functionality," where *epistemic functionality* refers to "our capacity to pursue and attain epistemic goals." What *epistemic goals* signifies, however, depends upon our "epistemological commitments": if we are *veritists,* our epistemic goal is to maximize truth by maximizing true beliefs and minimizing false beliefs (ibid.). Yet, if we are *virtue epistemologists,* our epistemic goals presumably include the "promotion of intellectual virtues" (ibid.).

Similarly, *epistemic costs* consist of two different meanings. First, most straightforwardly, epistemic cost signifies the failure of a belief to issue in or constitute knowledge, either because it

does not bear the appropriate relation to evidence and thus is not justified, i.e., rational, or because the belief is false. Second, epistemic costs can also be understood within a virtue epistemology framework where they count as epistemic vices that undermine intellectual virtues. Such vices include a failure to epistemically engage with the world or one's environment, to "acquiring, retaining, and using relevant information" (Bortolotti 2020: 11) in an epistemically virtuous way, or failing to cultivate epistemic virtues such as open-mindedness, curiosity, honesty, wisdom, and understanding (ibid. 14).

The epistemic cost or benefit of a belief can also have important connections to its practical cost or benefit. So, in the case of the belief *that lead is dangerous to ingest,* the truth-value of the belief's propositional content is relevant to the practical value of holding such a belief. The truth of the proposition *S believes that lead is dangerous to ingest* has practical implications for S depending on whether the intentional object of the belief, *that lead is dangerous to ingest,* is true or false. Its truth-value is relevant to the practical costs or benefits of holding such a belief. Epistemic costs in such cases ground the practical costs; epistemic benefits likewise ground the practical benefits.

In contrast, there is no such pairing issue when it comes to the truth-value of some proposition *p* and the social cost or benefits of believing *that p* (except for cases where epistemic cost or benefits enhance the social benefit of believing *that p*). For instance, the falsity of the propositional content, i.e., the epistemic cost of falsely believing *that p* (e.g., *that Obama was born in Kenya*), does not incur a social cost for S (depending on the social context or environment). For the average person, holding the false beliefs *that Obama was born in Kenya* or *that the American Civil War was fought over State's Rights* makes little epistemic difference for the believer yet can and typically does make a substantial difference in terms of social costs or benefits for holding such beliefs. Beliefs after all are not formed in social vacuums. And, as we've seen, believing *that p* even when *p* is false or

even *because p* is false, might be socially beneficial for S in the right kind of social context or environment.

When social and epistemic beliefs pull in opposite directions with respect to some propositional content, and the epistemic costs are sufficiently low, political conformity or some other tribalistic space may be privileged above all other criteria, presumably in virtue of the utility such tribal beliefs incur for S. Even when the epistemic costs are significant, the utility of forming and preserving beliefs that align with the messaging of the group may be paramount to one's cultural and personal success and thus maintained in virtue of their utility. Seemingly this is the dynamic at play when certain groups refuse to social distance, wear masks, or get vaccinated among those on the cultural and political Right where beliefs that vaccines are dangerous or ineffective or that COVID-19 is not real or any worse than the flu underpin such behaviors; it might even be that risk-prone behavior and their underpinning beliefs enhance their social utility by signaling commitment to the group, even at the expense of one's own health and safety, thus communicating that one is a reliable, trustworthy, and loyal group member (cf. Sterelny 2012)

3.2 Social but not Signal

Apart from a *signaling* or *expressive* function (Funkhouser 2017; Boyd & Richerson 2005; Kahan 2015, 2016, 2017; Herek 1986), tribal beliefs also have an existential value insofar as they provided the basis for constructing a self-conception and social identity, defining ourselves in terms of social categories. Through shared cultural narratives and language (Ross 2007), including tribal idioms, clichés, and slang which partitions off the linguistic space of the group from other groups, we participate in a shared social and cultural identity. In turn, this shared social and cultural identity informs in part our self-conception through mechanisms such as social comparison. Tribal language and narrative encode and transmit important social information – tribal propositions -- such as

information about social norms and expectations, rationalizations for the group's position, behavior, and existence, as well as group messaging and talking points for group members to internalize and echo. This information affects how we view ourselves in relation to the world and others – our world orientation – as well as inform the process of constructing a social identity – a social-identity which is psychologically valuable insofar as it contributes positively to our self-conception and self-esteem (Tajfel & Turner 1986). When a group fails to encourage or reinforce a positive self-conception or be the source of positive self-esteem, we tend to seek out another group which can and will fulfill this psychological and existential need (ibid.)

Despite inevitable tensions between some individual beliefs and ingroup messaging, tribal beliefs, perhaps in part because our self-conception is grounded in or entangled with our membership in the tribe (Tajfel & Turner 1986: 277), motivate biased evaluations of the group. Tribal beliefs partly constitute tribal narratives that rationalize the behavior and views of the group. They provide specific content to our social categories that we appropriate into self-constitution (Zawidzki 2013: 60). They also provide just-so talking points for members to rationalize the behavior, values, norms, and views central to the group's identity. Tribal beliefs which perform narrative and justificatory functions may subsequently be taken up as social signals in currency among a cognitive community of group members, whereby they assume a post-justificatory function. Tribal narratives and messaging, partly determine, at the level of the individual, the content of a person's self-conception, influencing how we interpret ourselves holistically as social, moral, and psychological agents. This leads to positing two more plausible social functions for belief that are not exhausted by the signaling category: (i) the justificatory function and (ii) the existential or social-identity function.

Conclusion

In this paper, I have explored the social nature of belief, looking at the social functions which belief potentially performs and the social utility they engender. In particular, I have examined the social signaling function, delineating this function into two kinds: shibboleth and commitment signaling. I have specifically surveyed the way health and climate beliefs (especially those with a conspiratorial bent) signal important and strategic social information. The signaling function crosscuts both individual and group fitness, enhancing individual self-conception and social reputations while also playing a vital role in coalitional psychology. The social signaling function, by offering a sustained systematic model of doxastic plurality wherein beliefs perform a plurality of selected-for proper functions, also offers an alternative framework for thinking about belief that departs from traditional Veridicalist accounts of belief. Lastly, this alternative model attempts to motivate extending the psychological space and positing an additional doxastic kind into our belief taxonomy and psychological ontology: tribal beliefs. Such beliefs are delineated by their sociality, specifically, by their social proper functions which specifically develop in response to social environments subject to the forces and pressures of intergroup dynamics – social spaces divided into ingroups and outgroups.

References:

__Achen, C. H., & Bartels, L. M. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton: Princeton University Press.

__Adler, J. (2006). *Belief's Own Ethics*. Cambridge, MA: MIT Press.

__Audi, R. (2001). "Doxastic Voluntarism and the Ethics of Belief." In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Matthias Steup (ed.), Oxford: Oxford University Press.

__Bayne, T. and Pacherie, E. (2005). "In Defence of the Doxastic Conception of Delusions." *Mind and Language* 20 (2): 163-188.

__Beck, J. 2017. "This Article Won't Change Your Mind: the facts on why facts alone can't change false beliefs." *The Atlantic.* Online. https://www.theatlantic.com/science/archive/2017/03/this-article-wont-change-your-mind/519093/

__Block, Ned (1995). "On a confusion about a function of consciousness." *Brain and Behavioral Sciences* 18 (2): 227-247.

__Boghossian, P. (2008). *Content and Justification.* Oxford: Oxford University Press.

__Bortolotti, Lisa (2020). *The Epistemic Innocence of Irrational Beliefs.* Oxford: Oxford University Press.

__Bortolotti, Lisa (2009). *Delusions and Other Irrational Beliefs.* Oxford: Oxford University Press.

__Boyd, R. and Richerson, P. J. (2005). *The Origin and Evolution of Cultures.* OUP: NY.

__ Boyer, P. (2018). *Minds Make Societies: How Cognition Explains the World Humans Create.* New Haven: Yale University Press.

__Bratman, M. E. (1992). "Practical Reasoning and Acceptance in a Context." *Mind*, *101* (401): 1–15.

__Browning, O. (2021). "Anti-vaxxers storm Covid testing centre during Milton Keynes 'freedom rally.'" *Independent.* https://www.independent.co.uk/tv/editors-picks/anti-vaxxers-storm-covid-centre-vbc0e1c48

__Campbell, J. (2001). "Rationality, meaning, and the analysis of delusion." *Philosophy, Psychiatry, and Psychology* 8 (2-3):89-100.

__Clifford, W. K. (1877). "The Ethics of Belief." In *The Ethics of Belief and Other Essays.* Amherst, New York: Prometheus Books: 70-97.

__Cohen, G.L. (2003). "Party over Policy: The Dominating Impact of Group Influence on Political Beliefs." *Journal of Personality and Social Psychology*, 85 (5): 808-822.

__Cosmides, L. & Tooby, J. (1989). "Evolutionary psychology and the generation of culture, Part II. Case study: A computational theory of social exchange." *Ethology and Sociobiology*, 10: 51-97.

__ Currie, G. (2000). "Imagination, Delusion and Hallucinations." *Mind and Language* 15 (1): 168-183.

__Davidson, D. (1973). "Radical Interpretation." *Dialectica*, 27 (3-4): 313-328.

__Dennett, D. (1987). *The Intentional Stance.* MIT Press.

__Engel, P. (2013). "In Defense of Normativism about the Aim of Belief." In T. Chan (ed.), The Aim of Belief. NY: Oxford University Press. 32-63.

__Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails*. Minneapolis: University of Minnesota Press.

__ Festinger, L. (1954). "A Theory of Social Comparison Processes." *Human Relations*, 7(2): 117–140.

__Frankish, K. (2009) "Adaptive misbelief or judicious pragmatic acceptance?" *Behavioral and Brain Sciences*. Cambridge University Press, 32(6): 520–521.

__Funkhouser, Eric (2020). "A tribal mind: Beliefs that signal group identity or commitment." *Mind and Language*, 1-21.

__Funkhouser, E. (2017). "Beliefs as signals: A new function for belief." *Philosophical Psychology* 30 (6): 809-831.

__Funkhouser, E. (2003). "Willing Belief and the Norm of Truth." *Philosophical Studies* 115 (2): 179-195.

__Gendler, T. S. (2007). "Self-Deception as Pretense." *Philosophical Perspectives* 21 (1): 231-258.

__Glüer, K. and Wikforss, Å. (2013). "Against Belief Normativity." In Timothy Chan (ed.), *The Aim of Belief*. Oxford University Press.

__Goodman, J. and Carmichael, F. (2020). "Coronavirus: False and misleading claims about vaccines debunked." *BBC. https://www.bbc.co.uk/news/53525002*

__Greene, J.D. (2013). *Moral Tribes: Emotion, Reason and the Gap Between Us and Them.* NY: Penguin Press.

__Henrich, J. (2009). "The evolution of costly displays, cooperation and religion: credibility enhancing displays and their implications for cultural evolution." *Evolution and Human Behavior,* 30: 244-260

__Herek, G.M. (1986). "The Instrumentality of Attitudes: Toward a Neofunctional Theory." *Journal of Social Issues*, (42): 99-114.

__Hogan, R. (1983). "A socioanalytic theory of personality." In M.M. Page (ed.), *Nebraska Symposium on Motivation,* (30): 55-89. Lincoln: University of Nebraska Press.

__Hornsey, M.J., Finlayson, M., Chatwood, G., and Begeny, C.T. (2020). "Donald Trump and vaccination: The effect of political identity, conspiracist ideation and presidential tweets on vaccine hesitancy." *Journal of Experimental Psychology,* (88): 1-8.

__Hornsey, M. J., & Fielding, K. S. (2017). "Attitude roots and Jiu Jitsu persuasion: Understanding and overcoming the motivated rejection of science." *American Psychologist, 72* (5), 459-473.

__Hornsey, M., Harris, E., Bain, P., Fielding, K.S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change* 6**:** 622–626.

__ Hume, D. (1993). *An Enquiry Concerning Human Understanding*. In Eric Steinberg (ed.), Indianapolis, IN: Hackett Publishing Company.

__James, W. (1896). "The Will to Believe. "*The New World* 5: 327-347.

__Kahan, D. 2017. "The Expressive Rationality of Inaccurate Perceptions." *Behavioral & Brain Sciences*. Jan (40) 6: 1-5.

__Kahan, D. (2016). "The politically motivated reasoning paradigm, Part 1: What politically motivated reasoning is and how to measure it – Kahan – major reference works – Wiley Online Library. Emerging Trends in Social and Behavioral Sciences: 1-16.

__Lemos, N. (2007). *An Introduction to the Theory of Knowledge*. Cambridge University Press.

__Levy, N. (2021a). "Echoes of covid misinformation." Philosophical Psychology. https://doi.org/10.1080/09515089.2021.2009452

__Levy, N. (2021b). "Virtue signaling is virtuous." *Synthese* 198 (10): 9545-9562.

__Maynard-Smith, J. & Harper, D. (2003). *Animal Signals*. Oxford: Oxford University Press.

__McHugh, C. and Whiting, D. (2014). "The Normativity of Belief." *Analysis,* 74 (4): 698-713.

__Mercier, H. 2020. *Not Born Yesterday: The science of who we trust and what we believe*. New Jersey: Princeton University Press.

__Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. MIT Press.

__ Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. MIT Press.

__Nickerson, R. S. (1998). "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology, 2*(2): 175-220.

__Nyhan, B., Reifler, J. (2010). "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32**:** 303–330.

__Owens, David John (2003). "Does belief have an aim?" *Philosophical Studies* 115 (3): 283-305.

__Papineau, D. (2013). "There Are No Norms of Belief." In Timothy Chan (ed.), *The Aim of Belief*. Oxford University Press.

__Paulus, D. and Trapnell, P.D. (2008). "Self-presentation of personality: An agency-communion framework." In O.P. John, R.W. Robins, & L.A. Pervin (eds.), *Handbook of Personality Psychology* New York: Guilford: 492-517.

__Pivetti, M., Giannino M., Bonomo, M., & Hakoköngäs, E. (2021). "Conspiracy Beliefs and Acceptance of COVID-Vaccine: An Exploratory Study in Italy." *Social Sciences*, 10 (108): 1-17.

__Price, H. H. *Belief: The Gifford Lectures Delivered at the University of Aberdeen in 1960*. NY, Humanities P.

__Ramsey, F. P. (1990). *F. P. Ramsey: Philosophical Papers*. Cambridge: Cambridge University Press.

__Romer, D., & Jamieson, K. H. (2020). "Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S." *Social science & medicine (1982)*, *263*, 113356. https://doi.org/10.1016/j.socscimed.2020.113356

__Ross, D. (2007). "*H. Sapiens* as ecologically special: what does language contribute?" *Language Sciences*, 29 (5): 710-731.

__Shah, Nishi (2002). "Clearing Space for Doxastic Voluntarism." *The Monist* 85 (3): 436-445.

__Shah, N. and Velleman, J.D. (2005). "Doxastic Deliberation." *Philosophical Review* 114: 497-534.

__Shepherd, H., MacKendrick, N., and Mora, C. (2020). "Pandemic Politics: Political Worldviews and COVID-19 Beliefs and Practices in an Unsettled Time." SOCIUS (6): 1-18.

__Sosis, R. (2003). "Why aren't We All Hutterites? Costly Signaling Theory and Religious Behavior." *Human Nature*, 14: 91-127.

__Sperber, D. & Mercier, H. (eds.) (2017). *The Enigma of Reason*. Cambridge, MA, USA: Harvard University Press.

__Stanovich, K. (2011). *Rationality and the Reflective Mind.* Oxford: Oxford University Press.

__Steglich-Petersen, A. (2017). "Fictional Persuasion and the Nature of Belief." In Ema Sullivan-Bissett, Helen Bradley & Paul Noordhof (eds.), Art and Belief. Oxford: Oxford University Press: 174-19.

__Steglich-Petersen, A. (2013). "Doxastic norms and the aim of belief." *Teorema: International Journal of Philosophy* 32 (3):59-74.

__Steglich-Petersen, A. (2009). "Weighing the aim of belief." *Philosophical Studies* 145 (3): 395-405.

__Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique.* Cambridge, MA. MIT Press.

__Sunstein, C. R. & Vermeule, A. (2009). "Conspiracy theories: Causes and cures." *Journal of Political Philosophy* 17 (2):202-227.

__Taylor, S. and Asmundson, G. (2021). "Negative attitudes about facemasks during the COVID-19 pandemic: The dual importance of perceived ineffectiveness and psychological reactance." *PloS one*, *16*(2) e0246317. https://doi.org/10.1371/journal.pone.0246317

__Tajfel, H., & Turner, J. C. (1986). "The Social Identity Theory of Intergroup Behavior." In J. T. Jost & J. Sidanius (Eds.), *Key readings in social psychology. Political psychology: Key readings*. NY: Psychology Press: 276-293.

__Taylor, S. E., & Brown, J. D. (1988). "Illusion and well-being: A social psychological perspective on mental health." *Psychological Bulletin, 103* (2): 193-210.

__Tosi, J. and Warmke, B. (2020). "Moral grandstanding as a threat to free expression." *Social Philosophy and Policy* 37 (2): 170-189.

__Tosi, J. and Warmke, B. (2016). "Moral Grandstanding." *Philosophy and Public Affairs* 44 (3): 197-217.

__Veblen, T. (1899). *The Theory of the Leisure Class*. New York: NY, Macmillan.

__Velleman, D. (2000). "On the aim of belief." In The Possibility of Practical Reason. Oxford: OUP: 244-81.

__Walsh, B. "The U.S. on Coronavirus Masks." *Axios*. Jun 24, 2020. https://www.axios.com/political-divide-coronavirus-masks-1053d5bd-deb3-4cf4-9570-0ba492134f3e.html

__Williams, B. A. O. (1973). *Problems of the Self*. Cambridge University Press.

__Williams, D. (2021). "Signalling, commitment, and strategic absurdities." *Mind and Language*: 1-19.

__Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

__Zalabardo, José L. (2010). "Why believe the truth? Shah and Velleman on the aim of belief." *Philosophical Explorations* 13 (1):1 - 21.

__Zangwill, N. (2005). "The normativity of the mental." *Philosophical Explorations* 8: 1-19

__Zawidzki, T.W. (2013). *Mindshaping: A New Framework for Understanding Human Social Cognition*. Bradford.