Abstract: Here I will develop a naturalistic account of epistemic reflection and its significance for epistemology. I will first argue that thought, as opposed to mere information processing, requires a capacity for cognitive self-regulation. After discussing the basic capacities necessary for cognitive self-regulation of any kind, I will consider qualitatively different kinds of thought that can emerge when the basic capacities enable the creature to interiorize a form of social cooperation. First, I will discuss second-personal cooperation and the kind of thought that emerges from its interiorization. Then, I will discuss third-personal cooperation and the kind of thought that emerges from its interiorization. We will see that epistemic reflection is the interiorized version of interpersonal argumentation, which is the epistemic component of third-personal cooperation. In developing this account, I will draw heavily on the work of Michael Tomasello and other cognitive scientists advocating the "social intentionality hypothesis". However, I will show how work done in the defeasible reasoning tradition can provide us with a deeper explanation of some claims made by advocates of the social intentionality hypothesis. Additionally, we will see that work done on social intentionality can help us better understand the significance of knowledge and justification as understood by the defeasible reasoning tradition. We will see that the social intentionality hypothesis and the defeasible reasoning tradition are mutually illuminating. By drawing equally on both, I will provide a novel account of the foundations of knowledge.

Word Count: 8,900.

Keywords: Reflection; Social Intentionality; Defeasible Reasoning; Foundationalism; Naturalism

Reflective Naturalism

"As I understand it, being justified is a status one has in relation to a person or group of people. If that's not what epistemologists have in mind when they speak of epistemic justification, I have no idea what they mean."

-Sydney Morgenbesser

In this paper, I will develop a naturalistic account of the epistemic significance of higher-order reflection. In one sense of "naturalist", the naturalist is committed to reducing normative vocabulary to non-normative vocabulary, finding a supervenience base for it, or something of the sort. This is not what I will be doing. In another sense of "naturalist", naturalists think philosophy should be informed by, and itself capable of contributing to, the worldview developed in the sciences. This is the kind of naturalism I espouse, and I hope to show that this project will vindicate the philosophical significance of epistemic reflection.

I will proceed by showing that work on social intentionality[1] in cognitive science and work on defeasible reasoning[2] are mutually illuminating and mutually supporting. The work in cognitive science helps us understand why justification and knowledge, as understood by proponents of the defeasible reasoning tradition, matter for creatures like us. The work in defeasible reasoning helps us understand how the game of giving and asking for reasons does what the cognitive scientists say it does. I will develop an account of the foundations of knowledge that draws equally from both traditions to this end.

---

[1] The leading proponent of this movement is Michael Tomasello (1999; 2014; 2021). Related ideas are pursued in Karmiloff-Smith (1992), Dunbar (1996), Frawley (1997), Mercier & Sperber (2011; 2017), Godfey-Smith & Yegnashankaran (2011), Scott-Philips (2015).

[2] Lehrer (1965; 1970), Lehrer & Paxson (1968), Hilpinen (1971), Swain (1974), Barker (1976), Klein (1971; 1976; 1980), Pollock (1986), Moser (1989), Schroeder (2015), de Almeida & Fett (2016), Paulson (2023a, b; forthcoming) a, b). See Shope (1983) for a useful overview of the history of defeasible reasoning in epistemology.

I begin by motivating the claim that thought differs from mere information-processing because it involves cognitive self-regulation on the part of the subject.[3] I will consider three kinds of self-regulation and three corresponding kinds of thought associated with them.[4] After describing the most basic kind, I generate the second kind by describing a way that the capacities at play in the previous kind of self-regulation can take a social turn by interiorizing a form of social cooperation: second-personal cooperation (more on "interiorization" in §2). I then generate a third kind that involves the capacities at play in the previous one taking another turn by interiorizing a more complicated form of social cooperation: third-personal cooperation. The epistemic component of this form of social cooperation is the game of giving and asking for reasons. The defeasible reasoning tradition gives us the normative kinematics of this game. A capacity for higher-order reflection will be shown to be necessary for the third kind of thought. Reflection's philosophical significance consists, then, not only in the benefits of the highly adaptive form of social cooperation it makes possible. It also makes our thought different in kind from other forms of thought.[5]

In section (I) I begin my taxonomy of kinds of thought or, as Tomasello (2014) puts it, "intentionality". Each kind of intentionality emerges from the previous kind, both evolutionarily and conceptually. In section (I) I do some stage setting and discuss the most basic kind, "Individual Intentionality". In section (II), I discuss the emergence of "joint intentionality" when self-regulation takes a social turn by interiorizing a form of social cooperation. In section (III), I discuss the emergence of "collective intentionality" when self-regulation takes another turn by interiorizing a more complicated form of social cooperation. Throughout section (III) we will encounter claims made in the cognitive sciences that seem right but cry out for further explanation. In section (IV) I

---

[3] Here I will be developing ideas found in Hurley (2003) and Camp (2009).
[4] Here I follow the taxonomy found in Tomasello (2014).
[5] This aspect of its significance is often overlooked by reflection's detractors. See for example Kornblith (2011).

lay out the core claims of the defeasible reasoning tradition. In section (V) I draw on the defeasible

reasoning tradition to further explain the claims from section (III). Part of the explanation consists

in my account of the foundations of knowledge that ties the work on social intentionality together

with the work on defeasible reasoning.

(I) Background on Cognition and Individual Intentionality

Thought isn't something that simply occurs within us: it involves cognitive activity on our part.

This is what differentiates genuine thought from varieties of cognition that are less robust, such as

stimulus-response learning and the "innate releasing mechanisms" of early ethology.[6]Unlike thought,

these are sub-agential processes. They occur within the agent, but the agent takes no active role in

them.[7]

A common narrative in the history of cognitive science is that the discipline was born with the

overthrow of behaviorism. This happened when mentalistic vocabulary was no longer used as

shorthand for behavioral dispositions but rather to refer to theoretical entities (often agent-level

cognitive states) capable of explaining those behavioral dispositions.[8]Part of the reason the

overthrow was needed in the first place was the difficulty of accounting for flexible and creative

problem solving, sometimes on the first try, with operant and classical conditioning alone.[9] It

became increasingly apparent that we can't explain the cognitive feats of humans and many non-

human animals without positing an agent-level ability to monitor the processing, storage and

---

[6] Cf. Lorenz (1965), Tinbergen (1951).
[7] By "agent", I don't necessarily mean "moral agent". I consider three kinds of agency below, only the last of which is moral agency.
[8] See Bruner (1972), Fodor (1975), Baars (1986), Miller (2003). Greenwood (1999) offers criticism but agrees with something close to this.
[9] Camp (2009) makes a similar point.

deployment of information.[10] I think this reflects the fact that these creatures think and the capacity for thought requires a cognitively active agent.[11] If you take issue with my use of "thought" for whatever reason, I'll give you the term. You will presumably acknowledge more and less robust kinds of cognition. Read my remarks about thought as being about an explanatorily significant kind of cognition, more robust than other kinds due to the activity of the subject.

Going forward, I will flesh this picture out by first identifying the basic capacities that make a subject cognitively active. This will enable us to distinguish different forms of thought later by identifying qualitative changes in these capacities.

The first form of thought I will consider is "Individual Intentionality".[12] Individual intentionality involves a capacity for causal inference. A capacity for causal inference is required because robust cognition requires "some kind of mental representation that goes 'beyond the information given' to direct perception" (Tomasello & Call 1997: 8). Similarly, Elisabeth Camp argues persuasively that the capacity for stimulus-independent representation is an important aspect of cognitive activity (Camp 2009: §3). To enjoy individual intentionality, the subject must be able to represent situations both as goal states and as already-obtaining states and flexibly adapt her behavior to discrepancies between the two. Activity is a kind of control and, as Susan Hurley puts it, "Control just is the maintenance of a target value by *endogenous* adjustments for *exogenous* disturbances" (Hurley 2003: 235, her italics). This involves a model of the causal structure of the relevant features of the environment and the outcomes that would result from various interventions in it.

---

[10] In the case of non-human animals, pioneering work includes that of Griffin (1978), Premack & Woodruff (1978), and Cheney & Seyfarth (1980).

[11] Cf. Hurley (2003), Camp (2009), Tomasello (2014; 2021).

[12] I get the term from Tomasello (2014: Chapter 1).

Plausibly, this involves a capacity to simulate the outcome of possible courses of action offline.[13] The subject considers a possible course of action directed at the goal state, "presses play" and imagines the outcome of that action as determined by her causal model of the situation. As Barsalou puts it, "the ease of simulating a scenario underlies the acceptability of a causal explanation" (1999: 606). This enables the subject to anticipate the results of courses of action she has never performed before, which makes individual intentionality more robust and potentially more useful than mere operant conditioning. For example, in an experiment performed by Manrique et al. (2010), chimpanzees were presented with a novel food extraction problem. Its solution required a tool of certain dimensions in a different room out of sight. Chimpanzees could sometimes solve the problem on their first trial, demonstrating their possession of a causal model they retained even after they went to the other room and their proximal stimulation changed. As Tomasello explains, "They then simulated the use of at least some of the tools and the likely outcome in each case" (2014: 16).

Another important difference between individual intentionality and mere associative learning is the capacity for self-monitoring that it involves. The ability to simulate the outcomes of possible courses of action involves the subject being sensitive to its own simulations as well as the external features of the environment that it is tracking in the simulation. The subject needn't conceptualize the simulation as a simulation but must be implicitly sensitive to the fact that it is one. This involves a capacity for uncertainty monitoring: the subject must be able to determine how well equipped they are to simulate the situation accurately offline. Along with causal inference, this capacity comprises the most basic form of executive functioning, and it is what makes subjects with robust individual intentionality capable of "taking charge" in a way that creatures with only stimulus-response connections and fixed action patterns are not. So long as there is a capacity to reconsider one's own

---

[13] See Barsalou (1999; 2005; 2008).

mental states adaptively and implicitly, the creature is engaged in cognitive self-monitoring. There is empirical evidence that dolphins, pigeons, rats, monkeys and apes enjoy this capacity (Cf. Smith, Couchman & Beran 2014).[14]

Many creatures with individual cognition have a primitive theory of mind as well (cf. Premack & Woodruff 1978; Tomasello & Call 2008). That is, they encode information about the intentional states of others and factor that into their decision-making. Creatures with individual cognition are sometimes social animals, although they understand their relation to their conspecifics in a "Machiavellian" way (cf. Humphrey 1976). That is, conspecifics are much like the other features of their environment with which they must cope, except they have intentional states.[15]

(II)   Joint Intentionality

Here is a more robust form of social thought, one I will call "joint intentionality".[16] Instead of understanding one another as agents going our separate (though sometimes mutually beneficial) ways, we might understand ourselves as jointly pursuing a common goal and each performing complementary roles to that end.

This requires a capacity for socially recursive inference. That is, in order for us to form a joint intention to pursue a common goal, we have to each understand that the other understands that we understand that… this is what we are doing. The recursive reasoning will only ever be taken a few layers deep in occurrent mental episodes, but the creature implicitly understands that the pattern of inference can go on in the same way indefinitely (Tomasello 2014: 38).[17]

---

[14] See also Hampton (2001), Washburn, Smith & Shields (2006), Kornell, Son & Terrance (2007), Call (2010).
[15] See Hare et al. (2000) for more on this.
[16] Again, in line with the terminology of Tomasello (2014: Chapter 3), here borrowing heavily from Tuomela (2007).
[17] Cf. Clark (1996), Tomasello (2008).

Socially recursive inference makes "common ground"[18] between the parties to the joint activity possible. The common ground will be formed by *ad hoc* gestural communication and pantomiming (Tomasello 2014: 59ff.) since discursive reasoning and communication only emerges alongside the next kind of intentionality to be considered.

This kind of common ground makes possible another key feature of joint intentionality: social self-monitoring. Since we each have to know a certain amount about each other's mental states to form a joint intention in the first place, when we have a joint intention, we will also be able to understand what we are doing from the perspective of our counterpart. If you and I are foraging together, I know enough about your mental states to assess my own performance from your perspective. So, even though it might be tempting for me to free ride and let you do all the hard work, I am also capable of understanding the situation from your perspective. Joint intentionality does not include the ability to make this understanding explicit. That would require discursive capacities foreign to joint intentionality. Creatures with joint intentionality can only understand this implicitly via simulated assessment from the perspective of their collaborator.

To sum up, second-personal activity involves each participant playing a distinct role in a shared activity. When this kind of activity is interiorized, the subject incorporates both roles into her own thinking by toggling between viewing the activity from her own perspective and the perspective of the agent fulfilling the complementary role. This involves evaluating her own performance from the perspective of the other agent through simulated self-assessment. In this way, she views herself second-personally, evaluates herself from the second-person perspective and regulates her behavior in light of this evaluation. This is what "interiorization" involves.

---

[18] This is somewhat like Stalnaker's (1978) account of common ground, save that it is "quasi-propositional". For more on this, see Tomasello (2008; 2014: Chapter 3).

(III) Collective Intentionality

At this stage we have seen a purely first-personal form of thought and a second-personal form of thought emerge from it. This latter form of thought is distinctive in that it involves interiorizing the perspective of a collaborator into one's own thinking and using that perspective to assess prospective courses of action and regulate them in a way tailored to cooperative activity. The final turn is the emergence of third-personal thought or "collective intentionality (cf. Tomasello 2014: chapter 4). Here we build on the capacities at play in second-personal "joint intentionality" but with a twist: the subject interiorizes the perspective of the "generalized" other rather than just the perspective of this or that collaborator. This enables one to view oneself third-personally and evaluate oneself from the third-person perspective.

As in the case of joint intentionality, decision-making is collaborative. However, now the coordination is discursive. At this stage, coordination goes beyond pantomiming and involves giving and asking for reasons.[19] Instead of gesturing at the antelope tracks and hoping you will recognize what I think we should do, I can now explicitly state the intention I recommend we form together and explicitly support my recommendation with reasons you could either accept or challenge. Indeed, the primary biological function and philosophical significance of reasoning is not the regulation of individual thought, but rather interpersonal argumentation. The role reasoning plays in regulating the private thought of individuals is a derivative phenomenon that results from the interiorization of interpersonal argumentation. As Mercier & Sperber (2011: 60) put it,

---

[19] Brandom (1994) also gives an account of human thought in terms of scorekeeping in the game of giving and asking for reasons. The account on offer here owes much to his work, particularly his demonstration of the significance of toggling perspectives for doxastic scorekeeping, as well as the role discourse plays in rendering our practices explicit.

Reasoning contributes to the effectiveness and reliability of communication by allowing communicators to argue for their claim and by allowing addressees to assess these arguments. It thus increases both in quantity and in epistemic quality the information humans are able to share. Claiming as we do that this role of reasoning in social interaction is its main function fits well with much current work stressing the role of sociality in the unique cognitive capacities of humans.[20]

By "reliability" they mean "truth conduciveness". So, one of their claims is,

(1) The game of giving and asking for reasons is truth conducive.

I will refer to the activity of engaging in interpersonal argument as "the game of giving and asking for reasons" following Sellars (1956/1997). I do this because, as I will argue later, we can best understand interpersonal argumentation epistemologically by drawing on the defeasible reasoning tradition and defeasible reasoning, I hope to show, is helpfully thought of as a kind of game (cf. Lehrer 2017).

(1) sounds plausible, but empirical researchers haven't, to my knowledge, offered a deeper explanation of it. This is fair enough since it is a plausible claim, and their primary task is to explain other things in terms of it. Nonetheless, it would be good to have a deeper explanation of how the game of giving and asking for reasons is conducive to true belief. In fact, the connection between the game of giving and asking for reasons and the truth often takes a backseat to its function in facilitating social coordination in the empirical work. Indeed, when Tomasello talks about good reasons he sometimes puts "good" in scare-quotes.[21] The thought seems to be that the reasons can

---

[20] They then go on to cite the following, Byrne & Whiten (1988), Dunbar (1996) Dunbar & Shultz (2003) Hrdy (2009) Humphrey (1976), Tomasello et al. (2005), Whiten & Byrne (1997).
[21] For example, see Tomasello (2021: 20).

only be "good" in a sociologically relativized way constructed by the attitudes of the participants in the game. While Tomasello is right that the game of giving and asking for reasons is a social phenomenon, we should be hesitant to understand its norms in a way that severs the connection between the beliefs we offer reasons for and the truth of those beliefs. When this connection is severed, it is hard to see how the "reasons" put forth within the game are *reasons* to believe anything at all.

Going forward, I won't commit myself to any claims about what grounds the norms of the game or what the ultimate source of their normativity is.[22] Rather, I will show that we can draw on the defeasible reasoning tradition to articulate the normative kinematics of the game and doing so will help us flesh out the picture Tomasello and others give us in a way that helps us answer some of the outstanding questions their work leaves us.

One of Mercier & Sperber's main reasons for claiming (1) is that it helps explain how reasoning can be adaptively beneficial for humans even though we are prone to fallacious reasoning in non-argumentative situations. People display poor logical reasoning skills in contexts detached from interpersonal argument (Evans 2002: 981) but do much better when not so detached (Resnick et al 1993; Pennington & Hastie 1993: 155; Blum-Kulka et al 2002; Thompson et al 2005; Hagler & Brem 2008).[23]

The primacy of social reasoning does not mean that solitary reasoning is dispensable nor is it to deny that it also serves the function of improving our cognition. Rather, the point is that the quality of individual reasoning is improved as the reasoner's inner monologue starts to sound more

[22] Tomasello draws heavily on Searle (1995; 2001) in his account of where the norms come from. I remain agnostic.
[23] Many of the studies cited here test the ability to recognize formal fallacies. Similar findings pertain to informal fallacies. See Hahn & Oaksford (2007), Neuman (2003), Neuman et al. (2006), Weinstock et al. (2004), Rips (2002).

like a dialogue (cf. Mercier & Sperber 2011: 73[24]). For instance, Descartes' *Meditations* are the monologue of a solitary meditator, but that meditator has a knack for anticipating points at which others will resist and uses those points of resistance as a way of improving the arguments under consideration.[25] We might say that the monologue reads like a dialogue for this reason. My claim here is that individual thought serves its epistemic-regulative function by internalizing interpersonal argumentative dialogue, the purpose of which is social coordination on truth, and this is why individual thought that better mimics interpersonal dialogue better serves its regulative function (cf. Godfrey-Smith & Yegnashankaran 2011: 80). It may also explain why certain congenital disorders that disrupt inner speech (e.g., autism and Williams Syndrome) also disrupt cognitive self-regulation (cf. Frawley 1997).[26]

This is related to another point Mercier & Sperber make in the above quote: reasoning contributes to the effectiveness of communication. "Effectiveness" can mean many things, but they seem to be using it synonymously with "persuasiveness" here. This is particularly important because social cooperation obviously requires a fair amount of doxastic coordination. However, it would be good to have a deeper understanding of how interpersonal reasoning contributes to coordination and, most importantly, how to square this with the truth-directedness of doxastic states. So, I will now mark another important claim to be revisited and further explained in section (V),

    (2) The game of giving and asking for reasons is coordination conducive.

---

[24] Baier (1981: 182) makes a similar claim about the methodology of the *Meditations.* There she also sympathetically discusses a view of human thought much like the one I propose here and its relation to Descartes' views.

[25] Consider, for example, the transition from the sensory fallibility argument (7:18) to the dream argument (7:19).

[26] The above gives us reason to think that the epistemological significance of inner speech goes deeper than is typically thought. For alternative accounts of its significance, see Jackendoff (1996), Bermudez (2003) and Munroe (2021).

It is important to bear in mind here that it is not just our social relations with one another that change as we shift from joint to collective intentionality. As with the shift from individual to joint intentionality, a new form of self-regulation emerges. In joint intentionality, we saw a new form of self-regulation that involves self-monitoring and self-assessment from the perspective of this or that potential collaborator. In collective intentionality, a new form of self-regulation emerges when the standards of interpersonal argumentation are interiorized. That is, the same norms of argumentation that we can use to coordinate with others can also be used to regulate our own thought by simulating self-assessment from the standpoint of an imagined audience. However, it needn't be from the standpoint of anyone in particular. Collective intentionality was selected for in virtue of its ability to foster large-scale collaboration (Tomasello 2014: 80ff.). To this end, it enables us to coordinate with others generally, even those with whom we have yet to jointly attend to local goings on. As Tomasello puts it,

> The Capstone of all this- recognized by modern thinkers who take a sociocultural view of human thinking- is the internalization of these various interpersonal processes of making things explicit into individual rational thinking or reasoning. Making things explicit to facilitate the comprehension of a recipient leads the communicator to simulate, before actually producing an utterance, how his planned communicative act might be comprehended-perhaps in a kind of inner dialogue. Making things explicit to persuade someone in an argument leads the disputant to simulate ahead of time how a potential opponent might counter his argument. (Tomasello 2014: 112).

The philosophically significant claim here is that individual thought results from the interiorization of paradigmatically interpersonal activities, particularly those involving persuasion. In other words,

(3) For those enjoying collective intentionality, private individual thought is the interiorized form of the game of giving and asking for reasons.

This claim will be important later. However, that is not because I will draw on work in epistemology to give a deeper explanation of it but rather because I will draw upon it to give a deeper explanation of the significance of work done in epistemology.

With the internalization of interpersonal argumentation, we see the emergence of a new, essentially reflective, form of self-governance. We interiorize the standpoint of the critic (not this or that critic, but rather the moves available to any critic whomever they might happen to be) and use that interiorized standpoint to regulate our own cognition.

One might worry that although this kind of self-conscious and deliberate rehearsal before an imagined audience is something we do sometimes, it isn't something we do all the time. This is true, but it is consistent with my proposal. Since self-regulation is a necessary condition for robust cognition, the kind of self-regulation we enjoy determines the kind of cognition we have, and the epistemic standards appropriate to assessing intentional states of that kind. Even those states the subject doesn't bother to regulate can be epistemically assessed by reference to the standard. Since self-regulation here takes the form of interiorized argumentation, the standards of argumentation are the standards appropriate to its assessment. I will have more to say about these standards in the next section. To foreshadow a bit, the normative kinematics of the game of giving and asking for reasons have a default and challenge structure.[27] Since the subject is default entitled to rely on the most often-used belief-forming processes, we shouldn't expect the need for constant simulated self-

---

[27] Cf. Brandom (1994).

assessments. Similarly, we shouldn't expect most claims made in interpersonal communication to be challenged.

So far, we have been considering the evolutionary function of the game of giving and asking for reasons and the way higher-order epistemic reflection emerges from its interiorization. The evolutionary function is agonistic collaboration meant to coordinate on the truth. None of this is to deny that people sometimes play the game of giving and asking for reasons in bad faith or that the outcome of argumentative exchanges can be influenced by social power relations in untoward ways. Rather, it is to understand the practice in its ideal form, which will then allow us to see these things as corruptions of it. It is only by understanding the practice in its healthy form that we can later recognize pathology in it.

To better understand the ideal form of the practice, we need to consider what makes it capable of serving the purpose of truth-conducive generalized belief coordination. I will now draw on the defeasible reasoning tradition in epistemology to make sense of this.

(IV)   Defeasible Reasoning and Collective Intentionality

(1)-(3) require further explanation and defense. To provide that explanation and defense, we need to turn to the defeasible reasoning tradition in epistemology. After a brief explanation of the core tenets of that tradition, I will draw on its resources to explain and defend (1)-(3).

I will note before getting started that this section might at first appear to be a detour, but the appearance is misleading. The defeasible reasoning literature is often thought of as a self-contained game meant to address the Gettier problem. Although many people make use of defeasible reasoning in this way, this was not the original intention behind it. The tradition was inaugurated by Keith Lehrer in his (1965) solution to the Gettier problem. However, in more recent work, he has made clear his position that knowledge is a standing in the "justification game" (Lehrer 2017: 174),

which is to be understood in terms of the interpersonal practice of justifying ourselves to one another. Defeasibility constraints are just his way of spelling out the normative kinematics of that game. We found compelling reasons to think that our kind of thought is the interiorized version of the game of giving an asking for reasons. A natural next step is to try to better understand our kind of thought by considering the normative kinematics of the game we interiorize.

According to the defeasible reasoning tradition, a necessary condition for a subject to justifiably believe p is that they believe p on the basis of a *prima facie* good reason. *Prima facie* reasons are foundational reasons. This doesn't imply infallibility or even incorrigibility. It is consistent with moderate foundationalism. According to moderate foundationalists, *prima facie* reasons don't depend positively on other reasons for support. They do, however, depend negatively on the absence of countervailing reasons.[28]

Let us work through the basic idea with an example. Perceptual experience as of a red wall is, quite plausibly, a *prima facie* reason to believe that the wall is red. It is an *ultima facie* reason to so believe just in case it is undefeated. Testimony that the lighting in the room is non-standard is a defeater. Defeaters are *prima facie* reasons to give up a belief.[29] That is the intuitive idea behind the defeasible reasoning tradition. I won't try to give a further analysis of reasons here. We need only concern ourselves with how reasons are assessed in this paper. The things I will say about how they are assessed will need to be accommodated by the correct theory of what they are.

---

[28] The first person to put it this way was Robert Audi (1993: Chapter 3), although the idea is latent in earlier work on defeasible reasoning. Cf. Lehrer (1965; 1970), Lehrer & Paxson (1968), Hilpinen (1971), Swain (1974), Barker (1976), Klein (1971; 1976; 1980), Pollock (1986), Moser (1989), de Almeida & Fett (2016), Schroeder (2015), Paulson (2023a, b; forthcoming). See Shope (1983) for a useful overview of the history of defeasible reasoning in epistemology.
[29] Cf. Ballantyne (2015).

Beliefs can have a defeater but nonetheless be justified. This sometimes happens when the defeater is itself defeated. Suppose you are told that the person who told you the lighting is non-standard is a pathological liar. In such a case, the defeater would be defeated, and your *prima facie* reason would be restored. Interestingly, not all defeater-defeaters have this restorative effect. Some give you new reasons rather than restoring your (actual) justifying reasons.[30] Now consider a case in which you learn that the lighting on the wall really is non-standard, but the wall was also painted red so it would still look the same even if the lighting were to change. Here we have a non-restoring defeater. The justification conferred by your perceptual experience is now replaced, not restored, by the testimony about the painting of the wall. Following de Almeida & Fett (2016), I call defeater-defeaters that do not restore the subject's justifying reasons "non-restoring defeaters"[31] and those that do restore their justifying reasons "restoring defeaters".

To determine whether a subject's belief is justified, we first check to see that she has *prima facie* reason for that belief. If so, we check for defeaters in her possessed evidence. Non-restoring defeaters in her possessed evidence defeat her justification, so, following Steup (1996), I call them "justificational defeaters". But a belief can be justified, and true, without being knowledge. This is what happens in Gettier cases. In these cases, the subject's justification is defeated by a fact of which she is unaware. I am going to call these "factual defeaters".[32]

The key insight of the defeasible reasoning tradition is that justification and knowledge are a matter of the standing of the subject's reasons.[33] To determine whether the subject has the former, we must look at those reasons from the subject's perspective. To determine the latter, we look at

---

[30] Cf. Klein (1980), Pollock (1986), de Almeida & Fett (2016).
[31] Similar terminology can be found in Klein (1980).
[32] Some call them "knowledge defeaters" (Audi 1993) or "propositional defeaters" (Bergmann 2006).
[33] Some working in the tradition, such as Peter Klein (2017), say it is a matter of the quality of her evidence. I talk of reasons to better connect the theory of justification with the philosophy of mind developed earlier in the paper.

them *sub specie aeternitatis*. Either way, the procedure for epistemic assessment is the same, it is just performed relative to a different body of potential defeaters.[34]

We should clear up a common misunderstanding before proceeding. According to the defeasible reasoning tradition, you can only know that p if your justifying reason is ultimately undefeated by the facts. However, if you have a false belief, your justifying reason will be defeated by an ultimately stronger reason supporting its true negation. So, people sometimes think that the defeasible reasoning tradition entails infallibilism.[35] This is not the case. The defeasible reasoning tradition allows that one can know on the basis of reasons that are susceptible to defeat. Perceptual reasons, for example, are susceptible to defeat yet they can also provide knowledge. They can't provide knowledge when their potential to be defeated is actualized, but one can still have fallible knowledge on their basis because the potential is nonetheless there. Just as a bottle of acetone can be flammable but not on fire, a subject's beliefs can be defeasible but not defeated.

I propose we understand the game of giving and asking for reasons in terms of the way that reasons put forward within that game are epistemically assessed. We just saw that the defeasible reasoning tradition gives us an account of this at a structural level. That is, according to the defeasible reasoning tradition epistemic statuses are a function of rational relations between the subject's justifying reason(s) and other potentially countervailing reasons.

We can think of it like a card game. When it is time for me to put my cards on the table, we assess my hand by comparing it to the other hands on the table. For me to have any shot at winning, my hand must at least be a provisional winner (i.e., must involve a *prima facie* reason). A provisionally winning hand is an ultimately winning hand just in case either a) no other hand on the table even

---

[34] I develop this claim at greater length in my (2023b). There I argue that this helps explain the philosophical significance of knowledge for reasons that are not generally recognized in the literature.
[35] I mostly encounter this misunderstanding in conversation rather than in print.

provisionally overrules it (i.e., no defeaters whatsoever) or b) somebody else's hand provisionally overrules it, but only provisionally. In the case of b), your hand is provisionally overruled, but the hand that overrules the hand that overrules your hand does so in such a way that your hand goes back to the top of the leaderboard, rather than a different hand speaking in favor of the same propositional content. This is the upshot of the restoring/non-restoring point made earlier about defeaters.

(V)   Explaining (1)-(3)

Here are (1)-(3) again, but this time with a bit more flesh on (1) and (2) meant to foreshadow the explanation promised earlier,

1a) The game of giving and asking for reasons is truth-conducive because it is potentially self-correcting.

2a) The game of giving and asking for reasons is a coordination-conducive procedure: we lay down our cards and defer to the winning hand.

3) For those enjoying collective intentionality, private individual thought is the interiorized form of the game of giving and asking for reasons.

The explanation and defense of (1a) has several components. To make it easier to follow, I will begin with a summary of the components individually and how they relate to one another. I will also dedicate a different subsection to each component to help orient the reader.

§5.1: The ideal endpoint of the game of giving and asking for reasons (GoGAR) is maximally truth conducive.

§5.2: GoGAR requires the potential for self-correction for methodological reasons.

§5.3: For the same reasons, GoGAR has the potential to move participants infinitesimally close to the ideal endpoint. The potential is fully realized in ideal inquiry.

19

§5.4: Our actual participation in GoGAR is in many respects non-ideal, but we approximate the ideal enough for cumulative cultural evolution to take place. This shows that the potential truth-conduciveness of GoGAR is largely realized.

§5.5: Here I explain and defend (2a).

§5.6: Here I explain and defend (3).

§5.1:

The game of giving and asking for reasons works with one party giving a *prima facie* reason to believe something and someone (usually someone else) assessing the reason by checking for defeaters. If there is none, then the *prima facie* reason is an all-things considered reason. If there is a defeater, then we need to determine if that defeater is itself defeated. If it is not defeated, then the subject's *prima facie* reason is not an *ultima facie* reason. If it is defeated, then the *prima facie* reason is an *ultima facie* reason only if the defeater-defeater is restoring. This tells us what it is for a reason to be an *ultima facie* reason relative to a body of evidence. The next thing to determine is which body of evidence constitutes the ideal endpoint of GoGAR.

I propose that the ideal end point of GoGAR is a body of total evidence that is coextensive with the facts.[36] In other words, the ideal end point of GoGAR is a stage where the distinction between knowledge and justification collapses. At this stage, it is impossible to have a justified false belief or a Gettiered belief. If you have a false belief, then it is defeated by some fact. For this reason, GoGAR is maximally truth conducive at its ideal end point. That is, if you have an undefeated reason to believe p at the end point of GoGAR, then p is true.

Of course, no finite subject finds herself at the ideal end point. What now needs to be shown is that finite subjects playing GoGAR can advance themselves closer to the ideal end point

---

[36] Swain (1974) calls this claim (or an analogue about knowledge) "an epistemologist's pipe dream". Once I show that GoGAR has the potential to move us toward the ideal endpoint and this potential is partially realized in the actual world, his claim will have been undermined.

by playing the game. Since the ideal end point is maximally truth conducive, GoGAR becomes increasingly truth conducive as we approach it. The ideal is regulative since finite subjects can never reach it. We can nonetheless improve ourselves epistemically as we approach it.

§5.2

To understand how GoGAR can advance participants toward the ideal end point, we must understand its self-correcting potential. Although actual inquiry is not fully ideal, we will best understand the self-correcting potential of GoGAR by considering the case of ideal inquiry in which this potential is fully realized. We will then be in a position to determine the extent to which that potential is actually realized.

To understand the self-correcting potential of GoGAR, we need to understand why there is a need for self-correction in the first place. The need for self-correction arises because it is methodologically necessary that we take certain things for granted in inquiry, but to do so without dogma we must take those things for granted only provisionally. It is methodologically necessary to assume that tokens of certain belief-types are for the most part true because we would otherwise face a regress. These tokens are our foundational beliefs.

Since the assumption that they are mostly true is methodologically necessary, we inquire into it through conducting inquiry that provisionally assumes it. That is, we assume that our foundational beliefs are mostly true but expect (in both the normative and descriptive senses of the term) that further inquiry within the scope of that assumption will lend further confirmation. The assumption is only provisional because further inquiry could do the opposite. Recall that, according to the defeasible reasoning tradition, foundational beliefs don't depend positively on further reasons for support but depend negatively on the absence of defeaters. That is, a token perceptual belief does not depend for its justification on positive reasons to think perception is reliable in this case, but

21

rather depends negatively on the absence of defeaters giving one reason to think it is unreliable in this case. I propose that the same applies to the background assumptions of our epistemic practices. The background assumption that perceptual beliefs in general are mostly true requires no positive support but nonetheless depends on the absence of defeat.

Furthermore, part of the inquiry conducted within the scope of these background assumptions faces the burden of explaining why each kind of foundational belief is mostly true. For example, we must assume our perceptual states are mostly accurate to do perceptual psychology. Perceptual psychology is an empirical science that requires testing hypotheses by observation. However, we expect our findings in perceptual psychology to further vindicate the background assumption by accounting for the mechanisms that enable us to discriminate objects perceptually.[37] Complacent inquiry that doesn't probe into its own background assumptions in this way is conceivable, but not without epistemic cost. The cost is that the foundations become a fixed tradition rather than a provisional starting point.

It is always possible that inquiry within the scope of these background assumptions will fail to provide the sought-after vindication. In such a case, the strength of defeat for the background assumption is proportional to the maturity of the discipline investigating it. So, although the background assumptions don't require positive support (only the absence of defeat) for their epistemic potency, at a certain point in inquiry the absence of positive support will itself be a defeater. This is because, as the discipline investigating the assumption becomes increasingly mature, it becomes increasingly reasonable to think that if the assumption were true, the discipline would

---

[37] This is exactly what we see, for example, in when Marr (1982) explains the computational processes that enable us to discern luminance contours.

have been able to explain why by now. This fact makes the provisional reliance on the background assumptions less dogmatic than it might at first appear.

The picture I recommend is one according to which it is methodologically necessary that we assign some non-inferentially formed beliefs foundational status. Which ones receive this status is determined by our intellectual inheritance. Our cultural initiation involves the inheritance not only of beliefs[38] but also a practice of argumentatively defending them (cf. Tomasello 2017: 111). Part of that practice will involve treating tokens of some belief types as the unmoved movers of the practice. This involves the assumption that those beliefs are for the most part true. This assumption is, of methodological necessity, brute and uncritical at first (cf. Alfano & Levy 2019). However, it is not mere dogma because the inquiry conducted within its scope has the potential to put that very assumption in jeopardy.

§5.3

Let us now consider ideal inquiry for agents with collective intentionality. We have seen the need for background assumptions that tokens of certain belief-types are mostly true. Let us consider a case in which these assumptions are true for the agents in question. Furthermore, let us imagine a community such that nobody in it suffers from memory loss. This is obviously unrealistic, but it is nonetheless a useful simplification that we will abandon later.

We have a community of agents with collective intentionality conducting inquiry by playing GoGAR with one another. They are engaged in a form of agonistic collaboration: they are jointly pursuing the truth by exchanging information and occasionally debating with one another about which claims are true. Any time one of them forms a belief, she informs the others. They check for

---

[38] See Clark (1996) for more on this.

defeaters and, if there are none or the defeaters are defeated, they adopt that belief. Since their foundational beliefs are, we are assuming, mostly true, the web of beliefs for any given one of them is increasing in number as time goes on. Defeasible reasoning is not only an organon, but also a canon. So, sometimes they acquire defeaters for beliefs that were previously formed and their web of beliefs momentarily contracts. However, expansion exceeds contraction since the foundational beliefs are mostly true, the defeasible inferences drawn from them are as well and some of the beliefs that are revised when a defeater is discovered are re-instated later when a defeater-defeater is discovered.

As the web of beliefs expands and the false ones are gradually uncovered and discarded by the critical resources of defeasible reasoning, the web gradually approaches the ideal end point of GoGAR: the web that includes belief in all and only facts.

§5.4

We now see that GoGAR is *potentially* self-correcting in a way that makes it increasingly truth conducive as inquiry proceeds. We now need to ask whether it is *actually* truth conducive for us. There are clearly many respects in which actual inquiry does not resemble ideal inquiry. We suffer memory loss. We don't share all the information we acquire. Ideal inquiry takes the form of agonistic collaboration, but non-epistemic motives curtail earnest collaboration frequently. In ideal inquiry, the background assumption that foundational beliefs are mostly true is itself true. The extent to which actual human inquiry approximates ideal inquiry in this respect is debatable.

I won't try to determine the exact effect any of these factors has individually on the relation between actual and ideal inquiry. Rather, I will argue that despite these impediments, human inquiry approximates ideal inquiry enough to serve the needs of cumulative cultural evolution. Cumulative cultural evolution, of the sort exhibited by behaviorally modern *homo sapiens* at any rate, requires

incremental epistemic improvement. This requires that we gradually approach the ideal end point of GoGAR, at least in some domains.

Tomasello (1999) persuasively argues that cumulative cultural evolution in general is made possible by inheriting a goal-directed practice from one's ancestors.[39] Without this inheritance, everyone would need to reinvent the wheel. Once one has inherited the goal-directed practice, one can refine its methods so that they are more conducive to the goal. This makes intergenerational incremental improvement possible.

The game of giving and asking for reasons facilitates the epistemic component of cumulative cultural evolution. As Kim Sterelny puts it, "human cognitive competence is a collective achievement and a collective legacy; at any one moment of time, we depend on each other, and over time, we stand on the shoulders not of a few giants but a myriad of ordinary agents who have made and passed on intact the informational resources on which human lives depend." (Sterelny 2012: xii). We inherit what Quine called the "lore of our fathers" (1951): essentially long shelf-life information.[40] Each generation inherits the information and misinformation of the previous one and is able to expand upon and occasionally overhaul a subset of it because they do not need to re-discover the insights of their forebears.

Some amount of cumulative cultural evolution is possible for those with joint intentionality, since socially recursive inference is sufficient for figuring out what conspecifics are trying to do and how they are trying to do it. However, cumulative cultural evolution is significantly advanced when the game of giving and asking for reasons is introduced. When our cultural inheritance is partitioned

---

[39] It is an empirical question the extent to which this inheritance is the product of unconscious imitation (Henrich & Boyd 2002; Muthkurishna & Henrich 2016) vs. socially recursive inference (Scott-Philips 2015). See Sterelny (2017) for an overview of the debate.
[40] See Laland & Hoppitt (2003); Danchin & Luc-Alain (2004) and Sterelny (2012: 29).

into discrete, debatable statements, we are in a better position to determine which parts of it require revision. This, of course, requires higher-order reflection.

As I said above, inquiry requires that we provisionally assume that our foundational beliefs are for the most part true. The assumption is methodologically necessary but not dogmatic because inquiry conducted within its scope could potentially provide countervailing evidence. Furthermore, the more mature such inquiry gets, the narrower the gap between lack of positive evidential support for the assumption, on the one hand, and defeat on the other. Since you inherit not only beliefs but also a practice of assigning certain beliefs foundational status from previous generations, it follows that you outsource some of the necessary inquiry to them. That is, the previous generation (often) regards the same belief types as foundational as yours does. Part of your intellectual inheritance is the progress they've made on repaying the debt incurred by the background assumptions you both share. You inherit their lore and part of that lore is the progress they've made toward explaining how vision, for instance, is generally reliable.

Our relation to the previous generation and their lore is not entirely deferential. At a certain point our relation to them is a diachronic version of the game we play synchronically with one another. Sometimes the attempt to expand upon the discoveries of previous generations results in the acquisition of recalcitrant evidence: that is, the accumulation of defeaters. In this case we adjudicate the conflict with them just as we would one with a contemporary, except for the fact that they aren't around to advocate for themselves. So, just as in the case of private thought, we play both the part of the participant and the critic. However, instead of assessing our own thought from the perspective of the generalized other, we instead simulate the previous generation's attempt to do just this as we play the critical other. In either case, we go through the procedure laid down by defeasible reasoning: we check for *prima facie* reasons, then for defeaters, then check to see if there

are defeater-defeaters and, if so, whether they are restoring or not. Because the game of giving and asking for reasons can adjudicate diachronic and synchronic conflict in the same way, it enables us to incrementally accrue cognitive capital across generations the same way we accrue it within a generation.

§5.5

It is now time to explain and defend,

(2a) The game of giving and asking for reasons is a coordination-conducive procedure: we lay down our cards and defer to the winning hand.

If someone offers an argument that p, there is rational pressure on their audience to either accept the conclusion or find fault with the argument.[41] You could find fault with the argument in a few ways. You could deny that it even provides *prima facie* reason to believe p. You could also try to show that even though it provides a *prima facie* reason to believe p, it doesn't provide an undefeated reason to believe p. To do this you would have to find a defeater for the reason offered. If you were unable to find one, then you would face rational pressure to accept p. If you found one and the speaker were unable to find a restoring defeater-defeater, there would be rational pressure on her to retract her assertion that p and revise her belief that p.[42] In any case, there is rational pressure on one party to yield to the other. GoGAR is conducive to belief coordination because of this. Each party is responsive to rational pressure because rational pressure is a standing within GoGAR. Furthermore, their own individual thought is regulated by internalizing the standards of GoGAR and self-monitoring for discrepancies between the standards and their actual beliefs. This is what higher-order epistemic reflection involves.

---

[41] For more on rational pressure, see Goldberg (2020).
[42] Cf. MacFarlane (2014: 109).

§5.6

Finally, that gets us to,

(3) For those enjoying collective intentionality, private individual thought is the interiorized form of the game of giving and asking for reasons.

The game is interiorized when a single person plays both the role of gives reasons and asks for them. In the interpersonal form of the game, at any given point one person will be giving reasons why we should believe what she recommends, and the other person will be playing a critical role by asking for reasons and assessing the quality of the reasons she gives. In the interiorized form of the game, a single person plays both roles. Tomasello illustrates the point as follows,

Thinking would seem to be a solitary activity. And so it is for other animal species. But for humans, thinking is like a jazz musician improvising a novel riff in the privacy of his own room. It is a solitary activity all right, but on an instrument made by others for that general purpose, after years of playing with and learning from other practitioners, in a musical genre with a rich history of legendary riffs, for an imagined audience of jazz afficionados. (Tomasello 2014: 1).

In the solitary case, the subject only has her own total evidence to work with, both in making arguments and in assessing them. This is what explains the significance of justification as understood by the defeasible reasoning tradition. According to the defeasible reasoning tradition, justification is the standing one has when one's *prima facie* reasons are undefeated by one's total evidence. Individual thought is itself the solitaire version of the game of giving and asking for reasons, so justification is determined by the subject's standing in a version of the game where the deck is restricted to her internally accessible evidence.

(VI)   Conclusion

We now have a naturalistic account of epistemic reflection. To achieve this, it was necessary to first consider what differentiates thought from mere information processing. The difference is cognitive self-regulation. It was then necessary to distinguish qualitatively different kinds of cognitive self-regulation that emerge when the most basic regulatory capacities take a social turn by interiorizing increasingly complicated forms of social cooperation. Our own form of cognitive self-regulation is what emerges from the interiorization of the game of giving and asking for reasons. This account synthesizes the first-personal approach to epistemology common among internalists and the third-personal approach preferred by externalists. To give an account of the normative kinematics of GoGAR, it was necessary to consider the defeater-checking procedure a participant would employ as she makes moves within the game. This involves considering the game from the first-person perspective of a participant. However, it was also necessary to consider the truth-conduciveness and self-correcting potential of GoGAR itself from a third-person perspective not unlike the perspective an engineer might take toward her invention. In doing so, we secure the truth-connection that people often worry goes missing when we take the first-person perspective.

Works Cited:

Audi, Robert. (1993) *The Structure of Justification*. Cambridge: Cambridge University Press.

Baars, B. J. (1986). *The Cognitive Revolution in Psychology*. New York: Guilford Press.

Baier, A. (1981) "Cartesian Persons" *Philosophia* 10 (3-4):169-188.

Ballantyne, N. (2015). "The Significance of Unpossessed Evidence". *Philosophical Quarterly* 65 (260): 315-335.

Barker, John. (1976). "What You Don't Know Won't Hurt You?". *American Philosophical Quarterly* 13:303-8.

Barsalou, L.W. (1999). "Perceptual Symbol Systems". *Behavioral and Brain Sciences*, 22, 588-609.

Barsalou, L.W. (2005). "Continuity of the Conceptual System Across Species". *Trends in Cognitive Sciences*, 9, 309-11.

Barsalou, L.W. (2008). "Grounded Cognition". *Annual Review of Psychology*, 59, 617-45.

Bergmann, M. (2006). *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford U.K., Oxford University Press.

Bermudez, J. (2003). *Thinking Without Words*. Oxford University Press: Oxford, UK.

Blum-Kulka, S., Blondheim, M. & Hacohen, G. (2002). "Traditions of Dispute: From Negotiations of Talmudic texts to the Arena of Political Discourse in the Media". *Journal of Pragmatics*. 34(10 – 11):1569– 94.

Brandom, R. (1994). *Making it Explicit*. Harvard University Press.

Bruner, J. (1972). "The Nature and Uses of Immaturity". *American Psychologist*, 27, 687-708.

Byrne, R. W. & Whiten, A., eds. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford University Press.

Call, J. & Tomasello, M. (2007). *The Gestural Communication of Apes and Monkeys*. Lawrence Erlbaum: Mahwah, NJ.

Call, J. (2010). Do apes know that they can be wrong? *Animal Cognition*, 13, 689 –700.

Camp, E. (2009). "Putting Thoughts to Work: Concepts, Systematicity, and Stimulus-Independence," Philosophy and Phenomenological Research 78:2, 275-311.

Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.

Chang, H. (2007). "Scientific Progress: Beyond Foundationalism & Coherentism". *Royal Institute of Philosophy Supplement* 61:1-20.

Cheney, D.L. & Seyfarth, R.M. (1980). "Vocal Recognition in Free Ranging Vervet Monkeys". *Animal Behavior*, 28, 362-7.

Chomsky, N. (1959). "Review of B. F. Skinner, *Verbal Behavior*". *Language, 35*, 26–58.

Clark, H. (1996). *Uses of Language*. Cambridge University Press: Cambridge, MA.

Danchin, E., & Luc-Alain, G. (2004). "Public Information: From nosy neighbors to cultural evolution". *Science* 305 (July 23).

de Almeida, Claudio. & Fett, J.R. (2016). "Defeasibility and Gettierization: A Reminder", Australasian Journal of Philosophy", 94:1, 152-169, DOI:10.1080/00048402.2015.1009127

Descartes, R. (2009). *Selected Philosophical Writings*. Trans. Cottingham, J., Stoothoff, R. Murdoch, D. Cambridge, MA: Cambridge University Press.

Dunbar, R. I. M. (1996) "The Social Brain Hypothesis". *Evolutionary Anthropology* 6:178– 90.

Dunbar, R. I. M. & Shultz, S. (2003). "Evolution of the Social Brain". *Science*. 302:1160 – 61.

Evans, J. St. B. T. (2002). "Logic and Human Reasoning: An Assessment of the Deduction Paradigm". *Psychological Bulletin* 128(6):978 – 96.

Fodor, J. (1975). "The Language of Thought". Cambridge, MA: Harvard University Press.

Frawley, W. (1997). *Vygotsky & Cognitive Science: Language and the Unification of the Social and Computational Mind.* Cambridge, MA: Harvard University Press.

Godfrey-Smth, P. & Yegnashanakaran, K. (2011). "Reasoning as Deliberative in Function but Dialogic in Structure and Origin". *Behavioral and Brain Sciences* 34(2), 80.

Goldberg, S. *Conversational Pressure.* Oxford: Oxford University Press.

Greenwood, J. (1999). "Understanding the 'Cognitive Revolution' in Psychology". *Journal of the History of the Behavioral Sciences*, 35(1): 1-22.

Griffin, D.R. (1978). "Prospects for a Cognitive Ethology". *Behavioral and Brain Sciences,* 4, 527-38.

Hagler, D. A. & Brem, S. K. (2008). "Reaching Agreement: The Structure & Pragmatics of Critical Care Nurses' Informal Argument. *Contemporary Educational Psychology* 33(3):403 – 24.

Hahn, U. & Oaksford, M. (2007). "The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies". *Psychological Review* 114:704 –32.

Hampton, R. R. (2001). "Rhesus monkeys know when they remember". *Proceedings of the National Academy of Sciences of the United States of America*, 98, 5359 –5362.

Hare, B., Call, J. & Tomasello, M. (2000). "Chimpanzees Know What Conspecifics Do and Do Not See". *Animal Behaviour,* 68, 571-81.

Henrich, J. & Boyd, R. (2002). "On Modeling Cognition and Culture: Why Cultural Evolution Does Not Require Replication of Representations". *Journal of Cognition and Culture,* 2:87-112.

Hilpinen, Risto. (1971). "Knowledge & Justification" *Ajatus.* 33:7-39.

Hrdy, S. B. (2009). *Mothers and Others.* Belknap Press.

Humphrey, N. K. (1976) "The social function of Intellect". In ed. P. P. G. Bateson & R. A. Hinde, *Growing points in ethology*, pp. 303 – 17. Cambridge, MA: Cambridge University Press.

Hurley, S. (2003). "Animal Action in the Space of Reasons". *Mind & Language,* 18(3):231-56.

Jackendoff, R. S. (1996). "How Language Helps us Think". *Pragmatics and Cognition*, 4(1), 1–34.

Khalifa, K., & Goldberg, S. C. (2022). "Socio-functional Foundations in Science: The Case of Measurement". *Philosophical Issues*, 32, 382–397

Klein, Peter. (1971) "A Proposed Definition of Propositional Knowledge". *Journal of Philosophy* 68: 471-82.

Klein, Peter. (1976) "Knowledge, Causality, Indefeasibility" *Journal of Philosophy* 76:792-812.

Klein, Peter. (1980) "Misleading Evidence and the Restoration of Justification" *Philosophical Studies* 37:81-9.

Klein, Peter. (2017) "The Nature of Knowledge" in *Explaining Knowledge: New Essays on the Gettier Problem* eds. Borges, R. de Almeida, C. and Klein, P. New York: Oxford University Press. pp. 35-57.

Kornblith, H. (2011). *On Reflection.* Oxford University Press: Oxford, UK.

Kornell, N., Son, L., & Terrace, H. (2007). "Transfer of Metacognitive Skills and Hint Seeking in Monkeys". *Psychological Science*, 18, 64 –71.

Laland, K. & Hoppitt, W. (2003). "Do Animals Have Culture?". *Evolutionary Anthropology* 12:150-9.

Lassonen-Aarnio, M. (2014)." Higher-Order Evidence and the Limits of Defeat", *Philosophy and Phenomenological Research*. 88.2: 314-345.

Lehrer, K. (1965) "Knowledge, Truth and Evidence" *Analysis* 25: 168-75.

Lehrer, K. (1970) "The Fourth Condition for Knowledge: A Defense" *Review of Metaphysics* 24:122-128.

Lehrer, K. (2017). "Defeasible Reasoning and Representation: The Lesson of Gettier". in *Explaining Knowledge: New Essays on the Gettier Problem* eds. Borges, R. de Almeida, C. and Klein, P. New York: Oxford University Press.

Lehrer, K. & Paxson, Thomas. (1968) "Knowledge: Undefeated Justified True Belief". *Journal of Philosophy* 66:225-37.

Lewis, D. (1969). *Convention*. Harvard University Press: Cambridge, MA.

Lorenz, K. (1965). *Evolution and the Modification of Behavior*. University of Chicago Press: Chicago, IL.

MacFarlane, J. (2014). *Assessment-Sensitivity: Relative Truth and its Appliccations*. Oxford University Press, Oxford, U.K.

Marr, D. (1982). *Vision*. San Francisco, W.H. Freeman & Company.

Melis, A., Call, J. & Tomasello, M. (2006). "Chimapnzees Conceal Visual and Auditory Information from Others". *Journal of Comparative Psychology*. 120, 154-62.

Mercier, H. & Sperber, D. (2011). "Why Do Humans Reason? Arguments for an Argumentative Theory". *Behavioral and Brain Sciences* 34(2), 57-74.

Moretti, Luca. & Piazza, Tommasso. (2018) "Defeaters in Epistemology: Introduction to the Special Issue", *Synthese* 195:2845-54.

Moser, Paul K. (1989) *Knowledge & Evidence*. Cambridge: Cambridge University Press.

Munroe, W. (2021). "Why are You Talking to Yourself? The Epistemic Role of Inner Speech in Reasoning". *Noûs* (56)4: 841-66.

Muthkurishna, M, & Henrich, J. (2016). *Innovation in the Collective Brain*. *Philosophical Transactions of the Royal Society Series B* 317(1690).

Neuman, Y. (2003.) "Go Ahead, prove that God Does not Exist! On High school Students' Ability to Deal with Fallacious arguments". *Learning and Instruction* 13(4):367 – 80.

Neuman, Y., Weinstock, M. P. & Glasner, A. (2006) "The Effect of Contextual Factors on the Judgment of Informal Reasoning Fallacies. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology* 59:411– 25.

Paulson, S. (2023a). "First-Class & Coach-Class Knowledge". *Episteme* 20(3): 736-56.

Paulson, S. (2023b). "Good Reasons are Apparent to the Knowing Subject". *Synthese* 202(1): 1-18.

Paulson, S. (forthcoming). "Luck and Reasons". *Episteme.*

Penn, D.C., Holyoak K.J. & Povinelli, D.J. (2008). "Darwin's Mistake: Explaining Discontinuity Between Human and Non-Human Minds". *Behavioral and Brain Sciences*, 31, 109-78.

Pennington, N. & Hastie, R. (1993). "Reasoning in Explanation-Based Decision Making". *Cognition,* 49(1 – 2):123.

Pollock, John. (1986) *Contemporary Theories of Knowledge.* 1ˢᵗ Edition. Rowman & Littlefield: Totowa, NJ.

Premack, D. & Woodruff, A.J. (1978). "'Does the Chimpanzee Have a Theory of Mind' Revisited" In Byrne, R.W. and Whiten, A. (eds.) *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans.* Oxford University Press: Oxford, UK. pp. 160-79.

Quine, W. (1951). "Two Dogmas of Empiricism". *Philosophical Review* 60 (1):20–43.

Resnick, L. B., Salmon, M., Zeitz, C. M., Wathen, S. H. & Holowchak, M. (1993). "Reasoning in Conversation". *Cognition and Instruction* 11(3–4):347–64.

Rips, L. J. (2002) "Circular Reasoning". *Cognitive Science* 26(6):767–95.

Schelling, T.C. (1960). *The Strategy of Conflict.* Harvard University Press: Cambridge, MA.

Schroeder, Mark. (2015) "Knowledge is Belief for Sufficient (Objective and Subjective) Reason." In *Oxford Studies in Epistemology,* volume 5, 226-252.

Scott-Phillips, T. (2015). *Speaking Our Minds.* London: Palgrave-Macmillan.

Searle, J. (1995). *The Construction of Social Reality.* New York: Free Press.

Searle, J. (2001). *Rationality in Action.* Cambridge, MA: M.I.T. Press.

Sellars, W. (1956/1997). *Empiricism and the Philosophy of Mind.* Cambridge: Harvard University Press.

Shope, Robert. (1983) *The Analysis of Knowing: A Decade of Research.* Princeton University Press.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Sociality.* Cambridge University Press: Cambridge, MA.

Smith, J.D., Couchman, J.J., & Beran, M.J. (2014). "Animal Metacognition: A Tale of Two Comparative Psychologies". 128(2), 115-31.

Stalnaker, R. (1978). Assertion. *Syntax and Semantics, 9,* 315-332.

Sterelny, K. (2012). *The Evolved Apprentice.* Oxford, U.K.: Oxford University Press.

Sterelny, K. (2017). "Cultural Evolution in California and Paris". *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 62: 42–50.

Steup, Matthias. (1996) *An Introduction to Contemporary Epistemology.* Prentice Hall: Saddle River, NJ.

Stiner, M.C., Barkai, R., & Gopher, A. (2009). "Cooperative Hunting and Meat Sharing 400-200 kya at Qesem Cave, Israel". *Proceedings of the National Academy of Sciences of the United States of America*, 106(32), pp. 13207-13212.

Swain, Marshall. (1974) "Epistemic Defeasibility" *American Philosophical Quarterly* 11:15-25.

Thompson, V. A., Evans, J. St. B. T. & Handley, S. J. (2005b) "Persuading and Dissuading by Conditional Argument". *Journal of Memory and Language,* 53(2):238 – 57.

Tinbergen, N. (1951). *The Study of Instinct.* Oxford University Press: Oxford, UK.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of Human Communication.* Cambridge, MA, M.I.T Press.

Tomasello, M. (2014). *A Natural History of Human Thinking.* Harvard University Press: Cambridge, MA.

Tomasello, M. (2021). *Becoming Human: A Theory of Ontogeny.* Harvard University Press: Cambridge, MA.

Tomasello, M. & Call, J. (1997). *Primate Cognition.* Oxford University Press: Oxford, UK.

Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005). "Understanding and Sharing Intentions: The Origins of Cultural Cognition". *Behavioral and Brain Sciences* 28(5):675 – 91.

Tuomela, R. (2007). *The Philosophy of Sociality*: *The Shared Point of View.* Oxford University Press: Oxford, UK.

Washburn, D. A., Smith, J. D., & Shields, W. E. (2006). "Rhesus monkeys (Macaca mulatta) Immediately Generalize the Uncertain Response". *Journal of Experimental Psychology: Animal Behavior Processes*, 32, 185– 189.

Weinstock, M., Neuman, Y. & Tabak, I. (2004). "Missing the Point or Missing the Norms? Epistemological Norms as Predictors of Students' Ability to Identify Fallacious Arguments". *Contemporary Educational Psychology* 29(1):77–94.

Whiten, A. & Byrne, R. W., eds. (1997) *Machiavellian Intelligence II: Extensions and Evaluations.* Cambridge University Press.