Appears in: Pettit, Philip ‘A Pragmatic Genealogy of Rule-Following’ in Joshua Gert, ed., Neopragmatism: Interventions in First-order Philosophy, Oxford University Press, 2023, 140-69 DOI: 10.1093/oso/9780192894809.003.0006

<recto>

Style Definition: CHBM:BIB

Style Definition: CN: Font: Bold

PART III

CP3

LANGUAGE, TRUTH, AND LOGIC

Abstract

How do creatures like us intentionally track certain properties when we use words to predicate them, yet have no means of defining those terms? This is the rule-following problem posed by Wittgenstein and Kripke. The answer defended is that we do so as a byproduct of practices that are well-documented as common across our species: sensitization, joint action, and teaching and learning. We can be sensitized to instances of a property or class, as even a simple animal can be sensitized under conditioning. Being committed to acting jointly with one another, we can become aware of such a class as an abstract entity. And being creatures who teach and learn from one another, within and across generations, we can recognize that if we diverge in assignments to a class, predications of a property, then at least one of us is not operating properly.

Keywords

rule-following, Wittgenstein, Kripke, sensitization, joint action, triangulation, apprentice learning, realism

<recto>

6

C6 # A Pragmatic Genealogy of Rule-~~following~~Following

Philip Pettit

C6S1 ## 1. Introduction

C6P1 How to make sense of our ability to follow rules, especially those basic rules for doing things that we may not be able to define? How, for example, to make sense of the rules we presumably follow in judging properly that something is a tool or a game, or is smooth or crimson, or regular or irregular in shape, where we have no other words in which to analyze those properties? Or if it seems that we might be able to define those rules and employ the definitions as formulae to guide us, how to make sense of the basic rules we presumably follow in making judgments about the properties—on pain of regress, there must be some—that we cannot define in that way? This I take to be the main rule-following problem highlighted in Wittgenstein's (1958) classic discussion, in the well-known commentary provided by Saul Kripke (1982), and in various other sources.[1]

C6P2 I approach the problem here in a novel fashion, asking about how creatures otherwise like us in make-up, might develop the ability to follow basic rules—and so rules in general—beginning from a point where that capacity hasn't yet materialized. I argue that some features

[1] For an overview of other approaches to the problem, and attempted resolutions, see (Miller 2018: ~~Ch's~~ chs. 5–6). I concentrate generally on properties, although the rule-following problem arises for other entities too such as the plus-function on which Kripke (1982) focuses.

that they have in common with us, going on accepted psychological theories, would make certain practices more or less inevitable and that those practices would give rise to a skill that answers to our notion of rule-following. Those practices and that skill would emerge with robust probability, so the proposal goes: the likelihood of their emergence would be independent of fortuitous events.[2]

C6P3 The thought experiment I propose is an example of the sort of counterfactual genealogy—by a recent account a 'pragmatic genealogy' (Queloz 2021)—that Bernard Williams (2002) hails as a novel approach to philosophical problems. It was used explicitly by Edward Craig (1990) in his account of knowledge, as it was used by Williams in his explanation of the importance of truth and truthfulness, and as I myself used it in a recent reconstruction of the nature of ethics (Pettit 2018). But it was also employed, without being named, in Wilfred Sellars's (1997) explanation of psychology, originally published in 1956, in Herbert Hart's (2012) account of law in 1961, and in David Lewis's (1969) theory of convention.[3]

C6P4 Lewis offers a particularly clear example of a counterfactual or pragmatic genealogy. He begins with creatures otherwise like us but lacking conventions or the idea of conventions;

[2] The genealogy provided builds on my earlier work but reworks and develops it in novel ways. Much of the earlier work is in ~~Part~~ part I of (Pettit 2002). I do use the word 'genealogy' in that work to describe my approach but do not spell it out in the way I do here. This essay might be seen as a response to (Miller (2018: ~~Ch~~ ch. 6.2), who suggests that the genealogy I offer may be circular, presupposing rule-following in the course of explaining it.

[3] Huw Price (1988) also gestures at something close to a genealogy in this sense but unlike the examples given, his book is often cast, fairly or unfairly, as debunking in character: it represents truth as something less important or real than it is commonly taken to be. For a general account of the genealogy approach, see (Queloz (2021).

argues that they would face various coordination predicaments, as in deciding what side of the road to take in meeting one another; shows that they would individually adjust to such problems on the basis of precedence or salience; and maintains that such adjustments would aggregate into recognized social regularities, serving a purpose that conventions typically serve. His suggestion, then, is that the regularities that would appear amongst our counterfactual counterparts can serve as models for conventions, providing a good sense of what they involve and what pragmatic purpose they serve.

C6P5 I think that a genealogical account of broadly the same sort can help to illuminate even something as cognitively fundamental to our species as rule-following. I take two well-documented, evolutionarily explicable features of human beings as given: the disposition to act jointly for shared, otherwise unattainable goals, and a form of teaching the young that does not just rely on their tendency to copy their elders. And I try to show that with features like these, our counterparts in the thought experiment—for short, humanoids—would be robustly likely to follow rules, including the basic rules where analyses run out.

C6P6 My hope is that the humanoid practice of rule-following that would appear under those imagined conditions can serve as a model of rule-following amongst our own kind, illuminating its nature and demystifying its existence. Insofar as our account of rule-following traces it to the pressures and opportunities opened up by independently sourced practices, it has a distinctively pragmatist character, although it is realist rather than skeptical about rules and the following of rules.

C6P7 The chapter is in ~~four~~ six sections. In ~~the first~~section 2, I offer an account of rule-following, explaining why it has seemed so problematic to many. In the following three (sections 3–5), I look respectively at three developments we can expect among the humanoids—sensitization to patterns, identification of patterns, and triangulation on patterns—arguing that together they would lead the humanoids into following rules. I argue

in a brief conclusion (section 6) that rule-following among humanoids offers a model of what rule-following amongst human beings involves as well. On this model, there really are rules, even basic rules, that we humans can be said to follow, but the rules present to us as important and discernible realities only within the perspective of practices that have an independent appeal.

~~1~~2. Rule-~~following~~ Following and the ~~problem~~ Problem ~~it~~ It ~~raises~~Raises

C6S2

2.1 Following ~~rules~~Rules, ~~basic~~ Basic and ~~otherwise~~Otherwise

C6S3

C6P8

A distinctive feature of our species is that we ask one another questions, we even ask ourselves questions, and try when the evidence is available to generate answers. Among the questions raised, we ask about whether properties that we cannot analyze or define—for short, basic properties—are realized in this case or not. And without having personal access to definitions, we try to provide answers. Is that a *tool* in your hands? Is that a *game* that they are playing? Is this number the *sum* of those? Is that coat *red* in color? Is that a *regular* shape or not? That we lack personal access to definitions in such cases does not mean that the properties are indefinable; they may be basic-for-us without being basic-for-experts. But some properties—some suitably determinate properties—must be basic for us, on pain of regress, as indeed some must be basic for experts. And we may assume that the cases given illustrate that category.[4]

[4] It may be, for all we assume, that what is basic-for-us or basic-for-experts are inter-defined packages of properties, not properties in isolation: this, in the way a line is defined, roughly, as the

C6P9 When we try to answer a question of this kind, the manifest assumption is that we understand the property involved, being aware of the sorts of conditions determining whether it is present, and that the aim of the exercise is to check on whether they are realized according to the evidence at our disposal. We think of that exercise as one in which we may fail but, at the same time, as one in which a greater effort on our part—paying closer attention to the property and to the evidence at hand—can reduce the likelihood of failure.

C6P10 That there are stable conditions linked with the property, determining whether it is present or not, means that there is a regularity in that linkage. That regularity will serve intuitively as a rule governing judgments about its presence, insofar as we have the ability to consciously try to conform to it—equivalently, to try to track the property—expecting that the attempt may raise our chances of success ~~success~~, but without ensuring~~, success~~ it. We cannot try to conform to it by consulting a formula that lays out the realization conditions of the property, since such a formula would require the sort of analysis or definition that we are taking to be absent. We can only try to conform to it by seeking to track the property directly, looking for evidence of its realization in this or that situation.

2.2 Modes of ~~rule~~Rule-~~following~~Following

C6S4

C6P11 When a regularity serves as a rule in this sense, then we control for conforming to it in a conscious and intentional manner; we manifestly try to ensure conformity. That the control is conscious follows from the part that the property plays as an object we ~~identify as~~take to be a target to track; that it is intentional is implied by the need for effort in this tracking. The fact that control is necessary to make conformity to the rule likely means that besides being conscious and intentional, rule-following has the further feature of being defeasible, indeed

shortest distance between two points and a point as the intersection of two lines. We ignore that possibility here for reasons of convenience.

defeasible in a way that must be salient to the agent. Despite our best efforts, we may fail to conform: the control may be unsuccessful.

C6P12 The notion of control invoked here is quite straightforward, being relevant in the case of mechanisms as well as agents. The air-conditioning system controls for the ambient temperature in a room, keeping it within certain limits. It ensures that the temperature will remain within the pre-set limits, robustly across a range of variations in circumstances—say, an influx of hot or cold air—although success is not guaranteed; after all, a mechanical glitch is always going to remain a possibility. In following a rule, we control in a similarly robust, if defeasible, fashion for conformity with the rule, adjusting the judgments we make on the questions before us as circumstance and evidence require. And we do that consciously, of course, and by acting out of an intention to get the correct answer.

C6P13 Or at least we do ~~that~~ this consciously and intentionally~~,~~ when rule-following has an active character that requires conscious effort and an explicit intention to conform. But consistently with being a form of control, rule-following may have a virtual or standby character instead. In that form it would not involve consciousness or intention but, as we shall see, it would still represent a form of agential control.

C6P14 Think of how the cowboy in the classic western controls for the direction the cattle will take when he lets them follow their noses and rides herd from the rear. Riding herd may just involve strolling along singing a song, as in the kitsch image, provided the cowboy is poised to notice any animal going off track, thereby raising a red flag, and provided he is ready to intervene in such a case and put it back on the desired route. The cowboy exercises virtual or standby control over the cattle and, even if he is never required to intervene, he will still expect to be remunerated for the job he ~~did~~has done.

C6P15 We may follow rules, including basic rules, in the same standby sense. Suppose we let habits dictate the judgments we make, and the words we utter, in response to various

questions about the instantiation of this or that property; suppose, in other words, that we let habits take over as the cowboy lets the cattle follow their noses. We will still be in control of the judgments made if the fact that the habits prompt an implausible judgment—a judgment that conflicts with standing assumptions—is likely to raise a red flag, and if that in turn is likely to prompt us to intervene as the errant animal would prompt the intervention of the cowboy. Habit might lead us to mistake a pepper for a tomato but biting into it would certainly raise a red flag. And such a flag would prompt us to resort to conscious, intentional effort in determining what it is we put in our mouth.

C6P16 This possibility is worth noting, because otherwise it might seem that our earlier description of what rule-following involves is excessively reflective and intellectual. It is likely that on most of the occasions when we follow a basic rule, we do so without thinking, as we naturally say: we let habit take over. Even when we do this, however, we will remain in control to the extent that we are disposed to notice a red flag—'Is this really a tomato?'—and to let that force us to think reflectively about the case. We will enjoy the virtual or standby control involved in riding herd on our habitual responses.

C6P17 Although the possibility and prevalence of standby control is worth noting, however, the focus of our discussion will be on active rule-following only. By definition, standby rule-following is possible only if active rule-following is possible, for it can occupy the stage only when the active mode is in the wings. In any case, it is the active mode that makes the problem of rule-following salient.

C6S5 2.3 The ~~rule~~Rule ~~-following~~ Following ~~problem~~Problem

C6P18 The problem is this. How do we manage to directly track a property that is basic for us in the sense explained: a property or indeed any similar entity, such as the plus-function implied in the property of being a sum? Rule-following in familiar cases involves explicit or implicit

definitions or formulae and, while it too may raise some problems, we shall concentrate here, following Wittgenstein, on the basic case where definitions are lacking.[5] In the definitional case, the resources that enable us to control appropriately for conformity are provided by the formula available. The problem in the case of following basic rules, specifically the active version of this case, is to explain what the resources are that enable us to track a property directly rather than definitionally.

C6P19 The idea that certain rules are basic connects with the idea of basic acts. That an act is basic means that while we perform it intentionally, we do not perform it by performing any other acts intentionally (Hornsby 1980). I may fasten my shoes intentionally by tying my shoelaces intentionally. But if I am proficient in doing so, I will tie my shoelaces intentionally without relying on doing anything else intentionally. The child may have to learn to move ~~its~~ their fingers, now in this way, now in that, to tie ~~its~~ their shoelaces, so that the act of tying is not basic for ~~it~~them. But when the child becomes proficient, ~~it~~ they will tie ~~them~~ the laces intentionally without any awareness of what ~~it~~ they do~~es~~ with ~~its~~ their fingers, and so without intentionally moving those fingers in any independently characterized way. ~~It~~ They will tie ~~its~~ their shoelaces intentionally; and ~~it~~ they will tie them without relying on intentionally taking a distinct step as a means to that end.

C6P20 Actively following basic rules is not only an intentional act but a basic intentional act. Following a rule that is defined for us in other terms means intentionally conforming to the rule in suitable situations by means of intentionally applying the relevant definition or formula. Following a basic rule means intentionally conforming to it in response to appropriate evidence but not by means of doing anything else intentionally—anything more basic—such as applying a definition.

[5] Paul Boghossian (2012) focuses on a more general problem, which we shall ignore here. For a response, see (Pettit (2024: ~~Ch~~ ch. 2).

C6P21 With plausible examples of basic rule-following, such as those given above, it may be hard to imagine that any creature, human or otherwise, could try to track a property without having a word to refer to that property or, equivalently, a word to express the concept under which it is viewed. Trying to track the property may consist as a matter of practice in trying to use a relevant word only in the presence of the appropriate property. But even if the effort to track a property is essentially connected with having a word for it—or at least being able to use words to identify it—this should not be taken to indicate that the problem primarily concerns how we learn to use that word. The main problem is how we can identify a basic property, or any such basic entity, as something to guide us in judging that it is present in certain cases and absent in others.

C6S6 2.4 The ~~problem~~ Problem in Wittgenstein

C6P22 The rule-following issue, as we have described it here, is at the center of Wittgenstein's discussion, and indeed Kripke's commentary. Kripke (1982: 24) puts the problem nicely when he formulates the requirement for the direct tracking of a property or other basic entity. The requirement is normative, and manifestly normative, in character: it requires that the property involved 'should *tell* me what I ought to do in each new instance'. If it didn't speak to me in some such sense, he suggests, then to judge that the property is present in ~~a~~ this or that case would be to 'make an unjustified leap in the dark' (Kripke 1982: 10).

C6P23 Wittgenstein's (1958: §175) discussion highlights the challenge of explaining basic rule-following in much the same way, arguing that the target involved—the property tracked—should *guide* me, as he puts it. It may seem that, in following a rule, I was just moved to go one way rather than another. But that seems wrong, he writes: 'I feel as if there must have been something else'. '"For surely," I tell myself, "I was being *guided*"'. The problem,

then, is to say how a basic property could *guide* me across cases: how, in Kripke's phrase, it could *tell* me which cases are instances of the property, which not.

C6P24 It would be unilluminating, according to Wittgenstein, to say in the sort of example he has in mind that I just grasp the nature of the property, the sense of the concept or word that ascribes it (see too Kripke 1982: 54). ~~"~~"~~It's~~ It's as if we could grasp the whole use of a word in a flash~~",~~," so I may think (Wittgenstein 1958: §197). But how is that grasp, that intuition, to guide me? An 'inner voice' tells me, someone may suggest. But 'how do I know *how* I am to obey it? And how do I know that it ~~doesn't~~ doesn't mislead me? For if it can guide me right, it can also guide me wrong' (Wittgenstein 1958: §213).

C6P25 Putting intuition aside, Wittgenstein asks whether I might grasp a basic property by surveying a set or series of instances and then extrapolating from those to other instances. Might I cotton on to the property by looking at instances, for example, of 'the same colours, the same lengths, the same shapes' and thereby learn to 'continue' the 'pattern uniformly' (Wittgenstein 1958: §208)? No, he claims. For how am I to know how 'to continue a pattern' (Wittgenstein 1958: §211)? Might I find reasons to go this way rather than that? No, for 'my reasons will soon give out' and then I can only 'act, without reasons'. The problem is that there is nothing about a finite series of any items that gives me reason to think of extrapolating to further items as following a rule. 'Whatever I do is, on some interpretation, in accord with the rule' (Wittgenstein 1958: §198).

C6P26 It is plausible with any series of items, of course—say, examples of addition which illustrate numbers as the sum of others—that we develop a disposition to continue in one way rather than others; that is part of what happens when we learn to add. So perhaps rule-following just consists in forming such a disposition and then acting as it prompts me in extrapolating to further instances? Perhaps to follow the rule for detecting sums, for example, is just 'to be disposed, when asked for any sum "x+y" to give the sum of x and y as the

answer' (Kripke 1982: 23). In a plausible interpretation of Wittgenstein, Kripke (1982: 24) argues that this won't work either, principally for the reason that to be subject to a disposition in proceeding is not to be guided towards what one ought to do or is justified in doing; it conflicts with the assumption in such a case that 'whatever in fact I (am disposed to) do, there is a unique thing that I should do'.

C6P27 On the face of it, Wittgenstein holds a realist view of rule-following, never doubting that we do follow rules even in the basic cases on which he focuses. But he defends that view in a somewhat aphoristic and opaque manner. Thus, he argues that when you follow a sign-post, in his own analogue, you will not only 'have been trained to react to this sign in a particular way',; you will go 'by a sign-post only in so far as there exists a regular use of sign-posts, a custom' (Wittgenstein 1958,: §198). He suggests that we will each have learned those customs—those uses or institutions, as he also says—insofar as we teach them to one another. And he thinks that such teaching will proceed 'by means of examples and by practice', 'by expressions of agreement, rejection, expectation, encouragement', and by the sort of 'gesture that means "go on like this"' (Wittgenstein 1958: §208).

C6P28 In Kripke's (1982: Ch ch. 3) interpretation, however, Wittgenstein assumes a very different profile, conceding in an anti-realist or skeptical spirit that no explanation of rule-following can save the phenomenon. I can be said to give the right answers in different cases, so the idea goes, but that is just to say that I give the sorts of answers most others in the community would give. And if I am described in such a case as following a rule, that only has the force of an honorific: it may imply approval for my conformity with established habits—I properly belong to the community—but it does not imply that in any literal sense I am actually following a rule: I am controlling my responses with a view to conforming to the rule.

~~2~~3. Being ~~sensitized~~ Sensitized to ~~patterns~~Patterns

C6S7
C6P29
In this section we begin to go through three practices and capacities that human beings generally display and that, by a range of accepted accounts, have been features of the human make-up well back into the history of the species. These capacities will belong to humanoids as well and the question is whether their exercise would engage them in following basic rules, by our account of what that involves. If it would, then the capacity of humanoids to follow rules would be explicable in naturalistically unproblematic terms. And if it could be explained in that way amongst the humanoids, it may lend itself to a similar explanation in our own kind.

C6P30
The capacities we explore enable human beings, and would enable the humanoids, first, to be sensitized to basic patterns; second, to identify those patterns as such; and, third, to triangulate on the patterns, making them objects of purportedly common attention. We discuss sensitization in this section, identification in section ~~three~~4, and triangulation in section ~~four~~5. Except when context implies otherwise, the patterns we have in mind throughout the discussion are all basic patterns that we cannot analyze in other terms.

3.1 Patterns ~~unlocked~~ Unlocked by a ~~key~~Key

C6S8
C6P31
The notion of a pattern invoked here is best introduced by contrast with a random set. A set of items will be patterned as distinct from random insofar as it is possible to present them more compactly than just by listing the members (Chaitin 1975, 1988). Thus, given one or another proper subset of the members, it should be possible to determine other members in the patterned set without having to list them one by one; in any case, listing members would be impossible with an unbounded set (Dennett 1991; Jackson, Pettit, and Smith 1999).

C6P32 That which makes it possible to grasp the extension or membership of a patterned set, without having to list all the members, may be an independently identifiable key to the pattern, in which case the pattern will be analyzable. Take the Fibonacci series of numbers, 0, 1, 1, 2, 3, 5, 8, 13…~~.~~ This is a patterned rather than ~~a~~ random set, as is probably intuitively obvious. But we can confirm that appearance insofar as there is an explicit key to unlock the pattern, so to speak. The key is that the series begins with 0 and 1 and that later members are each the sum of the two preceding numbers: the value of x_n, a member of the series at the *n*th position, is given by the formula, $x_n = x_{n-1} + x_{n-2}$.

C6P33 This way of unlocking or reading a pattern is special because the possibility of applying the key depends on the independent possibility of seeing a pattern in applications of the key: seeing them as applications of the same key. In the example, this would require an ability to see a pattern in the notion of summing, in the notion of two, and in the notion of preceding. And, of course, there may be no key available for doing that: no key to unlock the key itself. If patterns are to be detected, therefore, including a pattern as simple as that in the Fibonacci series, then there must be some patterns that can be unlocked without reliance on an explicit key; there must be some basic patterns.[6]

C6S9 3.2 Patterns ~~unlocked~~ Unlocked without a ~~key~~Key

C6P34 Happily, this is not a problem, since there is a host of patterns that natural creatures engage with, and effectively unlock, without applying such a key. The dog that learns to expect an outing on hearing the word ‘Walk’ engages with a pattern in that manner. And so does the

[6] The idea is reminiscent of the lesson taught by ~~<CE: Reference Lewis Carroll's (1895) has not been provided in the Bibliography. Please check.>~~Lewis Carroll's (1895) classic paper on Achilles and the tortoise: viz. that if someone is to be moved by any explicit deductive argument, they must subscribe to a rule of inference that does not appear as a premise in the argument.

pigeon that is conditioned to peck for food at doors that have a triangular shape. In such cases, as in a variety of cases that involve human beings and humanoids, the key to the pattern, if we can speak of a key, is presumably implemented in the brain of the creature involved as distinct from being applied in the manner of the key to the Fibonacci series. It works to unlock the pattern insofar as it means that, having been exposed to just a few instances of the pattern, the subject becomes sensitized to an open range of other instances, including instances that may vary in all sorts of other ways.

C6P35 Thus, when the dog or the pigeon is sensitized to a pattern in this engaged sense, instances of that pattern—instances of a 'walk' call in the dog's case, a triangular door in the pigeon's—will present to it as belonging to a single similarity class, despite varying in other ways. That the animal is sensitized to that class or property means that instances will prompt the same response robustly over variations in other features: variations of the voice and accent in which the call is made, for example, variations in the color and shape that a triangular door may assume.

3.3 Back to ~~rule~~Rule-~~following~~Following

C6S10

C6P36 Any agential creature that engages in the designated sense with a given pattern will form beliefs, on a functionalist account of belief, that that pattern—in our standard case, that property—is present in this or that situation and will be led on that basis to act accordingly for the satisfaction of its desire (Stalnaker 1984).[7] Thus, in recognizing triangular doors, the pigeon will form beliefs, now in this case, now in that, that a door is triangular—or that it has

[7] For simplicity, I ignore various complexities in the functionalist account of belief such as its connection with desire and the fact that beliefs and desires come in degrees.

some equivalent property—, and will be prompted to act as such a door would make sense for it to act, given its other beliefs and desires; it will peck at the door and receive its reward.

C6P37 This should not be in any way surprising. After all, we might construct a simple robot to emulate the performance of the pigeon, identifying triangular doors and doing something that corresponds to the pigeon's pecking. That robot will also merit the ascription of beliefs, on the functionalist conception of belief that I favor, since it will operate as an agent that aims at the achievement of a goal—pecking on regular doors—and adjusts its behavior for achieving that goal in light of how it represents the situation: it pecks on a door just when it believes that it is triangular in shape.

C6P38 As this is true of the simple pigeon and robot, so something similar will be true of the humanoids—as it is true of humans—across a much wider domain. They will be engaged in the relevant sense by patterns on any front where they learn to form beliefs and select means for satisfying their desires. Thus, for example, they will recognize the foods that nourish and the foods that don't; the animals on which they can prey, and the animals that prey on them; the areas that provide a safe refuge and those that don't; the materials that can be used to make tools and those that can't. In each category they will respond to the similarity that binds things together, the pattern that they display, and form corresponding beliefs that the associated property is present here, absent there, and so on.

C6P39 Might this capacity enable the humanoids to track the properties that they thereby engage with in a rule-following manner? Might they be able to control in a conscious and intentional way for making judgments, given the evidence available in any situation, on whether the relevant property is present or not? Might they be able, for example, to control for making a judgment—and forming a judgment-based belief—that this activity is indeed a game, this artifact a tool?

C6P40 No, for reasons similar to those emphasized by Kripke's Wittgenstein. He argues that while the humanoids might become disposed to track certain properties, such a disposition, however reliable it may be by our lights, would not enable them to identify something that might guide them on how they ought to proceed, telling them what to do in each new case; it would not play the required normative role. Their sensitization might lead the humanoids in the right direction, by our lights, in ascribing a certain property in this instance or that~~:~~; it might constitute a suitable extrapolative disposition.~~:~~ B~~But b~~ut it would not give them the ability to identify that property as such—that is, in abstraction from its instances—and to try consciously and intentionally to track it in judgment.

C6P41 That ability would require the humanoids to view the property tracked as an object of attention and awareness distinct from its instances. And sensitization to patterns, on its own, will not do the job. It will make it possible for them to form beliefs and desires about the concrete objects available for perceptual attention; if something is a game, for example, they may form the belief that it is a game or form the desire to take part. But this sensitization to particulars will not deliver sensitization to properties. It will not enable them, abstracting from particulars, to pay attention to the properties as such—say, the property of being a game or a tool—and to form a conscious controlling intention to look out for instances of that property and to make judgments about its presence or absence only when that is appropriate.

C6P42 Putting this point in another way, sensitization ensures that a property will be represented *in* any humanoid who is suitably sensitized but not ensure that it is represented *for* them: that is, as a representation they can consult (Cummins 1989). A representation in an agent may figure in explaining their performance, as when it plays a causal role in triggering certain beliefs and desires and generating actions. But the representation in the agent may play do this without playing the distinct role of enabling the agent to pay attention to the property itself and to form beliefs or desires about that abstract entity.

~~3~~4. Identifying ~~p~~Patterns

C6S11
C6P43 The capacity to be sensitized to patterns that the humanoids enjoy is replicated in a range of creatures, human and non-human. But while such a species-general capacity would not involve the humanoids in following rules, in particular basic rules, perhaps some species-specific practices and abilities would have this effect. We begin to explore that idea in this section, focusing on the capacity for joint action that human beings, and perhaps only human beings—the issue is debated in the empirical literature—display.

C6P44 The problem that sensitization to patterns leaves in place is that it may materialize in the humanoids, as in many other creatures, without giving them the ability to view any patterns or properties as objects of awareness, forming beliefs and other attitudes about them. And that means, of course, that they cannot aspire to track basic patterns, following them as rules. We shall now see that the pressure to act jointly explains why the humanoids will come to treat certain patterns as objects of awareness, thereby overcoming at least this particular obstacle to rule-following.

4.1 Joint ~~activities~~Activities

C6S12

C6P45 Human beings, and hence the humanoids of our narrative, are creatures who spontaneously act together for various common goals, going beyond the sort of action that mere sensitization would support. They have a distinctive capacity and inclination to combine their efforts to advance any goal where it is manifest, first, that they each desire its realization; and, second, that they can only or best achieve this in tandem, with each playing their part in a salient plan. Thus, if they are on the beach and they observe that a swimmer is in difficulty, they will be likely to get together to save the swimmer when it is manifest that this is a goal they share, that there is a salient plan under which they can realize it together, and that

anyone who begins to enact the plan will be joined by others. They may save the swimmer under such conditions, for example, by getting together to form a chain of people into the water; this may be the salient thing to do, perhaps because someone suggests it.[8]

C6P46 Michael Tomasello (2016) argues that this predisposition towards jointly intentional action is one of the most distinctive features of human beings, and is indeed exclusive to human beings. We may go along with him, if only for reasons of convenience, in thinking that it is exclusive to human and humanoid subjects, but our argument at this point does not depend on that extra claim. The main point is that human and humanoid agents may be taken to be capable of joint activity.

C6P47 Tomasello relies on two sources of evidence to support the claim that joint action is a characteristic of human beings. The first source of evidence is that in a crucial period of human evolution, between about 400K 400,000 and 150,000K years ago, the environment was such that our human forebears would have been forced to forage and hunt together—this, or die alone—and that that this would have created a selectional pressure in favor of a natural disposition to act jointly. They would have had to be able to distinguish edible from poisonous plants, and to collaborate in picking the edible and avoiding the poisonous. And equally they would have had to be able to recognize potential prey and potential predators and to combine in hunting the animals of the one sort and in defending against animals of the other.

C6P48 The second source of evidence on which Tomasello draws is that the disposition to act jointly with others is displayed by children between the ages of 1 and 3, although it is

[8] This is a very stark statement of what is involved in joint action, broadly in line with (Bratman (2014); see (Pettit (2017, 2023: Ch ch. 4). For other approaches to the analysis of this notion, any one of which would work for our purposes here, see (Tuomela (2007); Searle (2010(; Gilbert (2015).

generally not displayed by other primates. ‘These young children coordinate on a joint goal’, Tomasello (2014: 41) says, ~~:~~

> commit themselves to that joint goal until all get their reward, expect others to be similarly committed to the joint goal, divide the common spoils of a collaboration equally, take leave when breaking a commitment, understand their own and the partner’s role in the joint activity, and even help the partner in her role when necessary~~.~~.

C6P49

C6P50

As examples of such collaboration among very young children—a form of cooperation in which they hold one another to their expectations and protest at non-compliance—he mentions ‘giving and taking objects, rolling a ball back and forth, building a block tower together, putting away toys together, and “reading” books together’ (Tomasello 2014: 44).[9]

C6S13

4.2 Why ~~patterns~~ Patterns ~~will~~ Will ~~become~~ Become ~~salient~~Salient

C6P51

In order for any agents to practice jointly intentional activity of this kind, they must not only be sensitized to the pattern in this or that particular, be it an activity or an object~~. T~~, they must also be able to direct their attention to the pattern or class itself. They will have to do this when they form a belief that a partner is seeking a joint action in a certain class: say, that of playing some sort of game. And they will have to do it when they seek with others to find an object in a certain class: say, a plant of such and such a kind. They must be attuned to the property that unites instances of that activity and instances of that object. And, more than

[9] See too ~~(~~Tomasello (2008, 2009, 2016~~)~~.

that, they must assume that their partners in the enterprise are also attuned to the property and that the partners assume the same thing about them.[10]

C6P52 If agents did not have this extra capacity, then they could hardly plan to pursue a certain class of animals with others, or search out a certain class of plants. Indeed they could not plan to engage with others in any class of activity, even something as simple as playing a game together. Thus, without that capacity, to return to Tomasello's case, no child could expect collaboration on the part of another in rolling a ball or building a tower or reading a book. And no child would have grounds to remonstrate with another, as is apparently common among children, about their not enacting the pattern properly: not adding to the tower, not rolling back the ball, not joining in looking at a book, and so on.

C6P53 These observations show that in order to pursue joint activities as we human beings do, the humanoids will have to be able to classify the items they seek in gathering or hunting, as well as the sorts of activity involved in that exercise, seeing them as belonging to a class with which they are familiar. And, furthermore, the observations show that the humanoids will have to ascribe the same classificatory practice to others—the same practice of assigning particulars to one or another familiar class—and, more generally, must take it to be a practice that others ascribe to them.

C6P54 Sensitization would enable the humanoids to have beliefs about the particular objects they confront, in the way it would enable the pigeon to believe in one or another case that that

[10] This consideration may also support the stronger claim that the assumption has to be manifest or a matter that is public between them. Such manifestness may be interpreted for current purposes as involving common awareness in the sense of David Lewis (1969): that is, a hierarchy of assumptions involving the assumption by each, not only that they all assume a commonality, but that they all assume that they all assume the commonality, and so on. On the case for rival interpretations, see (Lederman (2018).

object is a triangular door. And it would enable them individually to pursue specific goals, with a goal of a general form of behavior—say, the search for food—crystallizing only under a suitable stimulus into the goal of gathering those plants or pursuing those prey. But the capacity for joint action would require the extra ability to have beliefs about classes of objects or activities: in effect, about properties. The humanoids would have to be able to form beliefs that others want to take part in a certain kind of activity like hunting or gathering, that they want to catch this or that sort of animal or find this or that variety of plant. They would have to be acquainted with the kinds or sorts or varieties involved in such cases.

C6P55 The shift required by joint action might be cast as one of intentional or semantical ascent. Previously, the humanoids, like other animals, would have been able to form beliefs that predicate certain properties of particular objects or assign the objects to certain classes. But now they must be able to form beliefs that take those properties or classes as items of which they can predicate other properties, as in thinking that this is the sort of thing others propose that they seek together, or this is the sort of activity that they intend to undertake jointly. Where particulars were previously the only items that served as objects of attention, now properties or classes can also enjoy that status.

4.3 Salient ~~patterns~~ Patterns and ~~signs~~Signs

C6S14

C6P56 It may be that, under the pressure of acting together, the humanoids will find certain patterns salient without having signs or words for them. The pattern associated with a certain game may be salient to them, for example, insofar as they can see any instance, or perhaps the simulacrum of an instance, as exemplifying the class of game in question (Goodman 1969). But whether not that is the case, their capacity for joint activity would certainly be much improved if they had signs, gestural or sounded, to denote the objects on which they act or

the pursuits they conduct. Such signs would enhance their capacity to initiate joint practices, and might even be essential for their success and spread.

C6P57 But can we assume that the humanoids will be able to support their joint activities with the development of suitable signs? Two pieces of evidence from non-human animals suggest that we can. The first is evidence of the use of standardized calls or signs to register this or that scenario and the other is evidence of a capacity to use a sign—in this case, a simple gesture—intentionally for a communicative effect.

C6P58 Taking up the first, we know that various animals utter cries that serve as signs, prompting in others the sorts of activity that they would perform if they had witnessed the situation eliciting the cry in the first place. The clearest example is from the vervet monkeys in Kenya studied by Dorothy Cheney and Robert Seyfarth (1990). These animals regularly warn one another of dangers by making appropriate calls: one call for an approaching leopard, another for a hovering eagle, another for a snake in the grass. Those hearing such a call generally come to believe that the relevant danger is present, taking the action that is appropriate, depending on whether they are on the ground or in a tree, for example.

C6P59 The calls emitted by the vervet monkeys may be more or less automatic, not intentional, but to turn to the second piece of evidence, there are also examples among non-human animals of intentional, communicative signing, this time by means of gestures. The outstanding example here is the way chimpanzees communicate in situations of targeted help, as they are called.

C6P60 In one oft cited study, for example, a first chimp needs a stick to reach food outside its enclosure and notices a stick that would do the job lying in an adjoining enclosure occupied by a second chimp (Yamamoto, Humle, and Tanaka 2009, 2012). The first chimp draws the attention of the second to the stick, and to its desire for that stick, by reaching towards it as if it were within its grasp, which it clearly is not. In taking that action the first chimp reveals,

not just a desire to get the stick, but a desire to get it by means of having the second chimp recognize its desire—see the significance of the futile reaching—and to respond appropriately. The first chimp acts in a broadly communicative manner and generally succeeds in conveying the message and winning the cooperation of the second (Grice 1957; Neale 1992; Moore 2016).

C6P61 Given such capacities among non-human animals, and given the utility that signs would have for the humanoids in conducting joint action, it is plausible to think that they will come to intentionally use different sounds as names, as we might say, for the objects on which they do or might act together and for the activities they might jointly perform. And with such a naming capacity at their disposal, they will surely extend the names they use to just about any pattern that becomes salient to them, whether in the world around them, or in their own individual or collective efforts.

C6P62 The development envisaged here among the humanoids can be highlighted by returning to the contrast with our simple, sensitized pigeon. Even the pigeon might be trained to make a sound, say 'troor', on forming the perceptual belief that a door is triangular, where the belief is a disposition to act on its desire as it is appropriate for it to act in the presence of such a door. But while the pigeon might be trained in this way, there would be no reason to think that it uses the sound 'troor' intentionally to mark the presence of the property assigned in its belief. That property does not exist for it as an object of attention and so the sign cannot function in its mind—though it might function in ours—as giving information about the property: viz.~~,~~ that it is instantiated here or there.

C6P63 Things are going to be very different for the humanoids, given the use of signs that we are positing among them. Suppose that they come to use 'tigroo' of animals in a certain class. Since they will be attuned to that class as such, and the property common to members of the class, the sound 'tigroo' will presumably function for them as a marker of the property; it will

not require any great cognitive insight for them to link the sound with the property. And that means in turn that an utterance like ‘tigroo’ amongst them will naturally be taken by them and by others—assuming no incompetence or insincerity—as an expression of the belief that there is a tigroo present. That utterance on the part of an individual humanoid will publicly assign the animal perceived to the kind to which their belief assigns it in their own mind.

C6P64 Again, suppose that the humanoids have a sound like ‘ganting’ that they use to identify instances of a certain sort of activity. As they learn to use ‘ganting’ of a kind of activity presented here or there, they will take the utterance of that sound to correspond to a belief they form, assigning to the activity on display the sort of property assigned in the belief. And likewise in a suitable context—say, in the absence of any evidence of ganting—they will naturally take and use the call to propose a joint ganting venture.

C6P65 In the foregoing discussion, we have assumed that joint activity will require certain classes or patterns to become salient objects of attention and that the humanoids will naturally develop signs to facilitate such activity. For all we need assume, however, it may be that joint activities and signing practices would evolve in tandem, with each being enabled or at least facilitated by the other. For all we need assume, indeed, it may even be that signing is a pre-requisite for making classes or patterns into objects of attention; this is close to a view maintained by Thomas Hobbes (Pettit 2008: ~~Ch~~ ch. 2).

4.4 Back to ~~rule~~Rule-~~following~~Following

C6S15

C6P66 We saw in the last section that just being sensitized to a basic pattern or property, being able to detect it instance- by- instance, will not enable the humanoids to have that pattern as an object of awareness, forming beliefs or other attitudes about it. That means that it cannot explain why or how they might come to track such a property consciously and intentionally. The discussion in this section shows why and how patterns might become objects of

awareness for the humanoids, thereby getting over that particular obstacle. The humanoids will be able to identify the pattern that binds any items together, and they will be able to form beliefs and desires about such a kind. Where the pigeon of the last section can only form the belief about a particular object that it, as we would say, is triangular, the humanoids will be able to form this or that belief about the class of triangles or about triangularity itself.

C6P67 Will the capacity to identify and name patterns, as distinct from being merely sensitized to those patterns, enable the humanoids to follow basic rules? No, it will not.

C6P68 The appearance of signs might seem to make it possible for any one of the humanoids to use a relevant name—say, ‘tigroo’ for an animal or ‘ganting’ for an activity—with the intention of using it just when the name is appropriate: to assert the presence of a tigroo only when such an animal is present or to use ‘ganting’ only of an activity where the name applies. But there are two problems that stand in the way of that possibility.

C6P69 The first problem is that the humanoids may form the intention, now in this case, now in that, to use ‘tigroo’ in response to the presence of such an animal—say, to use it to communicate with others—and yet have no general intention, however tacit, to use the word only when there is a tigroo present: only when the sign is appropriate. But even if we put aside that difficulty, there is a second problem to notice. This is that ~~even~~ if we ascribe such a general intention about sign-usage to the humanoids, there will be no suitable criterion of appropriate usage available to them, and so no basis for taking them to be following a rule.

C6P70 While each of the humanoids in joint action will identify one or another pattern that they expect others to identify too, they will always identify it as a class or kind or property corresponding to their own sensitization. Suppose, then, that two humanoids diverge from one another in a given case, with one using ‘tigroo’ to signal the presence of that sort of animal, while the other refuses to do so and, more generally, refuses to go along with the first by acting as the presence of a tigroo would make it appropriate for them to act. There is no

reason in the story so far why either of them might balk at the divergence, as they would presumably do ~~so~~ if they could be cast as intentionally and consciously seeking to follow a common rule. For all that we have assumed, they may simply turn away from such a conflict in their signaling and in their beliefs; they may just give up on the joint activity that convergence would likely have triggered.

C6P71 The fact that the humanoids generally succeed in marshalling joint activities, and in converging on the use of the corresponding sounds, means that ~~that~~ they are more or less commonly disposed to perceive and respond to a common world. But consistently with that being the case, they may each act as appropriate to a given activity, and use corresponding signs appropriately,~~,~~ only to the extent that their own sensitization to the patterns in question prompts those responses. Thus, they may be surprised or perplexed by the divergence envisaged. But there is no reason to think that they will not just walk away from the episode rather than reacting as would be appropriate if they were each seeking to follow a shared rule.

C6P72 If there is room for rule-following among the humanoids, by all that the account so far has suggested, it can only involve an attempt on the part of each to follow a solipsistic rule, as we might call it. This would be the rule each might follow of tracking a pattern—say, the tigroo pattern—across different instances by using the tigroo sign only when their sensitization supports that response: only when they are prompted independently to believe that an animal is a tigroo.

C6P73 Might this make sense? Might an individual humanoid intentionally and consciously seek to conform in their usage of a term like ‘tigroo’ to a pattern in the world: the pattern displayed by all and only tigroos? Might they do that, in particular, if their basis for using the term in a given case is provided by their own disposition, grounded in sensitization, to believe that this or that is or is not a tigroo?

C6P74 Not if it is the case, as we have assumed, that the idea of intentionally conforming to a pattern—that is, trying to conform to it—implies that manifestly the effort may not be successful: manifestly, the subject may fail to get the rule right. For on the picture presented, the disposition that prompts a humanoid to say ‘tigroo’ of a certain animal is just going to be the disposition that determines that that animal is indeed a tigroo. The diagnostic of success will itself be a guarantor of success, ruling out the possibility that they might get things wrong. They might claim to be following a rule, of course, but they could hardly claim to follow an elusive rule: a rule that they might miss or mistake.[11]

C6S16 ~~4~~5. Triangulating on ~~patterns~~Patterns

C6P75 Sensitization to a pattern or property, as we saw in ~~the~~ section ~~before last~~3, can enable the humanoids to form beliefs about concrete items to the effect that they instantiate the property or not. And, as we saw in ~~the last~~ section 4, the awareness of a pattern that is required in creatures capable of joint action, can enable the humanoids to go one better and form beliefs and other attitudes about the property relevant in such a case. But as we have just noted, even with this awareness of patterns, the humanoids may be unable to track a basic pattern or property in the manifestly defeasible or fallible way that rule-following requires.

[11] It is conceivable in principle, but hardly robustly likely, that an individual humanoid might recognize over time that their responses vary, might identify obstructing factors that occasionally affect them, and might try to track the property—still, of course, a potentially idiosyncratic property—that shows up only in the absence of obstruction~~;~~ (o~~n~~ n the idea of obstruction, see the discussion of restriction and distortion in the next section). For defenses of rule-following that is private in this sense, see {Blackburn (1984) and {Azzouni (2017).

C6P76 Apart from the general practice of joint action, however, there is a distinct practice of teaching and learning—a distinctive species of joint action—that is also characteristic of human beings and, by all accounts, exclusive to them. And this, as we shall see, can help us out of the defeasibility problem just raised; it can explain how humanoids might get to be able to intentionally and consciously track patterns that remain elusive: patterns that, as they recognize, they may miss or mistake in a given case.

C6S17 5.1 Teaching and ~~learning~~Learning

C6P77 In its developed form, teaching and learning involve jointly intentional action in which one party plays the role of teacher, the other or others the role of learner, and the shared aim is to achieve the transfer of practical skills or of received pieces of knowledge (Sterelny 2012; Lalande 2017). It is a kind of apprentice learning in which the man or woman who occupies the position of master gets the learner up to speed, as they work with one another on the job. Each makes an appropriate effort, whether in teaching or in learning, with a view to achieving the transfer of expertise that is sought by each. This might be the ability to prepare food, build a home, or fashion a hunting tool; a form of know-how about the dangers and opportunities of their environment; or the skill of recognizing edible plants, sowing and harvesting a crop, raising or hunting various animals.

C6P78 The evidence is that human beings have long interacted across generations, and indeed within generations too, to communicate how to perform the infinite variety of tasks on which human culture depends. Unlike other animals, or at least most other animals, they do not just rely on the young to copy what their elders do. They do not merely hope that the members of each generation will pick up skills in foraging for plants, hunting prey, or keeping clear of predators. Those in each generation teach those in the next generation how to do those things, eliciting the required pattern-reading dispositions in them.

C6P79 In illustrating this practice, Kim Sterelny (2012: 37–38) maintains that ‘a full apprentice model of expertise transmission’ came on stream amongst our ancestors a few hundred thousand years ago, applying in activities like tool-making, child-minding, and foraging. In this version, the practice presumably represented a distinctive species of joint action. But Sterelny cites reasons for thinking that an earlier version was already present among early hominins—*homo erectus* rather than *homo sapiens*—a couple of million years ago. The young ‘learned by doing, in environments that advantageously shaped individual trial-and-error learning’ and that were ‘structured advantageously by adults through the exercise of the adults’ own expertise’; in these environments, for example, ‘tools, partially completed tools, and raw materials were readily available as objects of play, experiment, and exploration’.

C6P80 Assuming that humankind ~~have~~ has long been involved in the practice of teaching and learning, we can also ascribe that practice to the humanoids of our narrative. Developed as a species of joint action, so we shall now see, the practice will enable them to consciously control for tracking patterns in such a way that they may manifestly fail in certain cases. It will enable them to bring the activity into line with our conception of rule-following.

C6S18 5.2 Defeasible ~~tracking~~Tracking

C6P81 For all that joint action in general guarantees, as we saw, the most salient response to divergence in the use of a sign like ‘tigroo’ may be for the parties to assume that they are tracking different patterns, albeit patterns that coincide in enough cases to make joint action possible. Will the presence of teaching and learning among the humanoids make a difference to the response they are likely to make? Suppose that one does not take something to instantiate a pattern that the other does take to be an instance, or takes something to be an

instance that the other does not view in that way. How would we expect them to respond to that divergence?

C6P82 The teaching-and-learning assumption means that they are unlikely to respond by concluding that they must not be targeting the same property. Or at least that will be so insofar as the case falls within the domain, as it surely will, where in principle teaching and learning is possible. Assuming that there is a common pattern targeted on both sides—a pattern-for-us, as they might cast it—and authorizing one another as generally capable of tracking that pattern, they will balk at the divergence. One of them, so it will seem to each, must fail in the tracking enterprise; they must miss an instance of the pattern that is there, or mistake another property for such an instance.

C6P83 But why would either party fail? Why, in particular, would either fail if they are both competent participants in the practice, trained up to a passing degree in sensitivity to the pattern? The only possible answer for them to endorse is that the sensitization of one or the other is affected by a restriction in the evidence presented—the prompt for triggering their sensitivity—or by a distortion in their perception of that evidence: something that perturbs the triggering of that sensitivity. In the tigroo example, they must assume that one or the other cannot see the animal properly or that their perception is not working properly.[12]

C6P84 What in theory might establish that such a difference or distortion—such an evidential hindrance—is present on one side of a dispute rather than on the other? The factor would have to differentiate the parties, for sure, and it ought to be the sort of thing that could conceivably have a restricting or distorting effect. It might be independently obvious that one or the other party is subject to such a hindering effect: that they are too far away, for example, to be sure of what they see or hear or smell. But again, it might not. So, which

[12] Both might be subject to a hindrance of some kind, while only one of them is lucky enough to get the pattern right. I put aside that possibility for ease of presentation.

factor in that case will deserve to be indicted as the hindrance? The answer presumably is: that which would best explain the divergence, consistently with the assumption of a common pattern and shared sensitization.

C6P85 But how might the humanoids identify such a culpable factor? The most salient method would be to see which of the competing judgments a majority of others would endorse and to defer to their view. If only one of the parties is out of step with most others, after all, that suggests that it is they who are evidentially hindered. The majority will necessarily be reliable if the disputed instance of the property at issue is in a domain where more or less arbitrary convention rules: say, in determining whether a telephone booth, as in some usages, counts as an instance of a box. And it will be reliable on empirical grounds in other cases: say, in determining the exact color of an object, where it is more likely that a single individual is color blind or impaired in some other way than that they alone see the color properly.

C6P86 Whatever the reason why the majority view should be taken as correct in this way, it is important to note that when an individual defers to the majority, in the scenario imagined, they will not do so just for the sake of social ease: they will not ~~coordination~~ coordinate just for coordination's sake. Rather, they will defer to others for the sake of triangulating reliably on a property they seek in common to track. Each individual will be guided by their own sensitization to the property they identify, but they will rely on that guidance only under the proviso that they are not subject to evidential restriction or distortion. If they defer to the majority, then, that will be for the sake of triangulating on the property they target—a pattern-for-us—not for the sake of coordinating, despite divergence, on how they publicly respond: say, on what word they use for that feature (Davidson 2001).

C6P87 Thus, if one party to a dispute finds that they are in the minority, then other things being equal, they will self-correct and restore convergence with others. It may take some time for

the parties involved to negotiate with one another and to reach such a resolution, but we can leave out such details here. We need only register that they will recognize the case for triangulating intersubjectively on one another to determine what is objectively so and that they will routinely do this to resolve their differences. Even the color-blind subject can be expected to treat the color that is there objectively to be one that they identify only very vaguely, due to their impaired vision: due to an impairment of which they will have become aware in the course of intersubjective triangulation.

C6P88 This story does not require, of course, that negotiation will never fail and that every dispute will be resolved. One response to failure might be to ignore and insulate the problematic case, with each agreeing that there is no saying who is tracking the pattern, and who is not. Another might be for each to assume that they are tracking the pattern, others not, and that there is an unrecognized restriction or distortion that is putting others astray. And a third might be to deem cases where there is continual divergence as unimportant, taking the property tracked to be one that is not defined for that range: on this account, the property tracked would be cast as a vague or indeterminate target in the manner of baldness.

C6P89 The use of something like the majoritarian method would presumably lead the humanoids to a position where, like folk epistemologists, they can track the properties of evidential restriction or distortion across different cases, and introduce words to name them. And that would facilitate triangulation as a means of distinguishing what is objectively so, by their intersubjective lights, from what merely seems to be so. It would mean that they could identify when they or others were likely to be going wrong, without explicit reference to majority judgment. They would do this on the basis of detecting the presence on one or another side of a recognized source of restriction or distortion.

C6P90 The upshot of these considerations is that if the humanoids practice teaching and learning of the kind that has long distinguished human beings, then their tracking of basic

properties is likely to be defeasible, and defeasible in a way that will be manifest to them. They will target properties that are revealed, not necessarily via their subjective sensitization, but via a corrected counterpart: their sensitization in the absence of factors that deserve to be identified as evidential restrictions and distortions.

C6S19 ## 5.3 Back to ~~rule~~Rule-~~following~~Following

C6P91 On the picture developed so far, the humanoids will each be sensitized like any animals to certain basic patterns: in our paradigm case, properties; they will be able to attend to those properties, making them into candidates about which to form beliefs and other attitudes; and specifically, they will be able to attend to them as properties that they may occasionally miss or mistake.

C6P92 To the extent that this picture fits, the humanoids will count as following basic rules. Each of them will be able to control consciously and intentionally for identifying instances of this or that basic property, making an effort to get things right but without a guarantee of success. They will realize that the patterns they track are those that show up for them only in the absence of the hindrances that are revealed in triangulation with others. And they will see that while their efforts can help to promote the chance of success, they will not make success inevitable.

C6P93 The rule-following that we ascribe to the humanoids does not presuppose any great intellectual sophistication. The presentation given of the achievement is liable to be misleading in that regard, since it may suggest that the humanoids will reach the goal described in a series of insights and inferences. But the achievement is really the more or less inevitable precipitate of a network of assumptions that are encoded in the practice we have described.

C6P94 Jointly acting as teachers and learners, the humanoids must assume that there are basic patterns in the shared world they inhabit—patterns-for-us—and that they in general are sensitized to such patterns; they could not reject those assumptions and continue with the practice. But they may make those two assumptions only in the sense of being disposed, perhaps unwittingly, to act in relevant contexts as their truth would require. The two assumptions, operating at this level, will force the humanoids to conclude in the case of divergence that something is amiss with the sensitization of those on one or another side: they are affected, as suggested, by a restriction of the evidence that triggers their sensitivity or a distortion of the effect of the evidence on their sensitivity. And the presumptive way of identifying and locating such a hindrance will be to check for how other individuals will respond in the relevant case.

C6P95 Let working assumptions that register points like these be encoded in the practice and responses of the humanoids, and rule-following will materialize amongst them without the need for reflective thought. It will prompt them without hesitation to treat their dispositions, on the one side, as lenses in which basic patterns reveal themselves instance by instance; but, on the other, as lenses that are likely to require correction in light of the hindrances that may arise. In actively making a judgment on some basic case, then, they will naturally see themselves as trying to attend to the pattern or property targeted, conscious that effort is needed to avoid the danger of going astray. In other words, they will control consciously, intentionally and defeasibly for tracking that property. They will fit the bill for following basic rules.

C6P96 How might a rule show up in the experience of the humanoids? Looking at an example~~s~~ of a tool or a game or a sum, to revert to earlier cases, how might they be able to see an abstract pattern there that they might try to track over further instances? Presumably the extrapolative disposition resulting from their sensitization, modulated by the disposition to

interrogate that disposition for the presence of a hindering factor, would let that pattern become salient. It will not be a pattern with which they can be acquainted in the way a defining formula would allow but a pattern discerned proleptically in their anticipation of where those dispositions would lead.

C6P97 For a model of what is likely to happen, think of how a group of friends might be said to know a route across a complex mountain park from points *A* to *B*. They might not be able to draw that route or describe it or give instructions to others about how to follow it. Yet, they can be said to know the route in virtue of knowing that they each have a generally reliable disposition to move correctly from the starting point to a particular landmark; that at each landmark short of the terminus they will have a similar disposition to move on correctly to another landmark; and that any failure on the side of one is likely to be corrected in negotiation about the discrepancy with others. They will know the route proleptically, by grace of the interacting dispositions on which they rely.

C6P98 As that group of friends have this dispositional, intersubjectively dependent grasp of the route from *A* to *B*, so we may think of the humanoids as having such a grasp of the pattern associated with a basic rule. And as the friends can set out consciously and intentionally to follow the route across the mountains, knowing that despite their best efforts, they may fail, so something similar is going to be true of the humanoids as they control for following a basic rule.

C6S20 6. Conclusion

C6P99 The genealogy provided makes sense of how basic rule-following might emerge among humanoids and what it would constitute for them. But the practice described in the genealogy may also be the practice that allows us human beings to follow basic rules, as we clearly do,

in our own thought and talk. If rule-following does not consist in the sort of practice described, so we might urge, what does it involve? If not this, what?

C6P100 The model of rule-following proposed suggests that basic properties—and other basic items like the addition function—become salient for us and present as features of the world only in virtue of practices that we undertake for pragmatic reasons associated with joint action and with teaching and learning. They are not data revealed in the pure light of theoretical reason, but patterns summoned to view under the pressure of practical concerns. Does this pragmatist aspect of the model count against it? Surely not, for it allows us still to embrace an important form of realism—pragmatic realism, if you will—about the basic properties and about the properties they serve in turn to analyze.

C6P101 The first point to make in support of this realism is that the model is consistent with holding that there is a fact of the matter about whether this or that basic property is instantiated in one or another case. It may be that we get to be able to access basic properties and take them as guides only to the extent that we are sensitized by nature to certain similarity classes and only to the extent that we operate with practices that allow in principle for the reconciliation of differences. But this still allows us to think that the properties that become visible by grace of those dispositions and practices are discovered by us, not invented or created.

C6P102 This way of thinking about those properties is supported by the model insofar as it implies that when we negotiate about differences we do not seek to coordinate for coordination's sake; we do not look for convergence, at whatever cost to our sense of being subject to pressures from without. We seek the sort of convergence that our practices make important, to be sure, but this form of convergence is built around the assumption that there is something in principle available to all that may not inevitably be accessed in practice. We display that assumption insofar as we authorize one another as potential sources of correction

and, seeking to identify and neutralize evidential hindrances, look for what is accessible on all sides.[13]

C6P103 It may be said, however, that the model offends against realism on a second count. The charge is that while the basic patterns or properties acknowledged are sustained by how the world proves to be—they are not our invention—still they can only be patterns of species-specific interest, even perhaps of just culture-specific interest. They cannot constitute properties that have a wide cosmological role in making sense of the world as a whole (Wright 1992).

C6P104 But that is not correct. The pattern that makes something tasty to human beings may be highly idiosyncratic and play no significant role in explaining anything other than our disposition to eat it. But a pattern like that of solidity may also be tracked in the manner of a basic property and enjoy much greater explanatory significance. It is capable of detection by more than one sensory modality, capable of detection by many other creatures as well, and capable of affecting how even non-sentient objects relate to one another. Thus, it can play a much wider role in explaining how the world operates than the tastiness of various foods to human beings.

C6P105 Finally, the pragmatic character of the line defended here may seem to offend against realism in a third respect. On a realist image of the world, there is only a contingent connection between what there is in our view and what there is in fact, so that human ignorance and error remain a permanent possibility. But, it may be said, our model of rule-following suggests that this is not so: that a majority of individuals cannot be wrong about

[13] The anti-realism of Kripke's (1982) Wittgenstein takes coordination to serve a coordinating purpose only. A realist account of rule-following, similar to that developed here, is present in the essays collected in (Pettit 2002: ~~Pt~~ part 1). For pragmatic accounts that apparently look for a middle way between anti-realism and realism in those senses, see (Price (1988) and (Gert (2012).

what holds and does not hold at the basic level. Conceiving of the world as revealed in our practices, so the idea would go, we fail in a characteristically pragmatist fashion to distinguish how the world appears within those practices and how it is in itself.

C6P106 It is certainly true according to the model developed here that if someone operates in a normally competent manner and is free from the hindering effects of evidential restriction and distortion, then they are bound to be accurate in the identification of a basic property: their sensitization to the property will dispose them to ascribe it just when it is present. This means that basic properties must satisfy a biconditional like this: (*x*)(*x* is an instance of a basic property *F* if and only if it is disposed to present as an instance of *F* to competent observers in ideal, unhindered conditions).

C6P107 An example of such a biconditional might be: (*x*)(*x* is red if and only it is disposed to look red to competent observers in ideal, unhindered conditions). Basic properties will satisfy this condition in virtue of the fact, not that observers think of *F*-ness as a disposition to evoke that effect—they will think of it as a categorical property, in the way they think of a color like red—but because mastering the concept of *F* requires being disposed to ascribe *F*-ness to anything under such conditions (Jackson and Pettit 2002). Reflection on how we use terms for basic properties—presumptively, a term like 'red'—ought to make that clear.

C6P108 This observation does not reflect a failure endemic among pragmatists to distinguish the world in itself from the world as it appears in human practices. It is supported, rather, by a highly plausible assumption: that the conditions that explain why our basic words have certain denotata—why 'red' ascribes red—must reflect the conditions that we have to meet if we are to master those words. I would scarcely count as understanding 'red', giving it an appropriate referent, if I were not disposed to use it of things that looked red, at least when I had no reason to think that I was subject to some hindering factor.

C6P109 But not only is the claim implied in our model independently plausible~~.~~ ~~I~~, it also does not seriously compromise realism. The truth of the biconditionals to which we are committed under the model of basic rule-following does not give us a title to claim any individual or collective infallibility. Any one human being, and any group or generation of human beings, may ~~fall~~ fail to recognize some of the hindrances that affect judgment—new hindrances are always likely to show up—and may fail to see that there is a hindrance present in a particular case. And so any individual or group, even a whole generation, may miss or mistake one or another basic property, being subject to a hindrance that only becomes obvious later. It has recently been suggested, for example, that continuing human evolution has increased sensitivity to violet, and that previous generations simply missed this color; that would explain why violet only began to appear in paintings from the 1860~~'~~s on (Tager, Kirchner, and Fedorovskaya 2021).

C6P110 The upshot is that qualms about betraying realist instincts need not inhibit us from endorsing the pragmatically oriented model of rule-following suggested by our genealogy. The model is decidedly different in that respect from the account of rule-following that Kripke ascribes to Wittgenstein. Wittgenstein's own remarks leave rule-following somewhat obscure, but it may be worth mentioning in conclusion that the model developed from the humanoid genealogy might be taken to make sense of those remarks.

C6P111 Wittgenstein (1958: §201) insists that 'there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call "~~—~~"obeying the rule"~~—~~" and "~~—~~"going against it"~~—~~" in actual cases'.[14] Our model plausibly explains how that can be the case, with assumptions built into practices driving judgments that follow appropriate rules. Wittgenstein compares such non-interpretational rule-following to following a sign-post, as we saw,

[14] For a fine account of Wittgenstein's rejection of the role of interpretation in basic rule-following, see (Miller (2015)~~.~~); ~~See~~ see also (Swindlehurst (2020).

arguing that an individual will only be able to do this ~~this~~ insofar as they 'have been trained to react to this sign in a particular way' and 'there exists a regular use of sign-posts, a custom' (Wittgenstein 1958: §198). Those remarks can be read in various ways, but they certainly make good sense on the model of rule-following recommended here.[15]

C6S21

References

Azzouni, J. (2017). *The Rule ~~following~~ Following Paradox and ~~its~~ Its Implications for Metaphysics*. New York, Springer.

Blackburn, S. (1984). "The Individual Strikes Back." *Synthese* 58: 281–301.

Boghossian, P. (2012). "What ~~is~~ Is Inference?" *Philosophical Studies* 169: 1–18.

Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford, Oxford University Press.

Carroll, L. (1895). "What the Tortoise ~~said~~ Said to Achilles." *Mind* 4: 278–80.

Chaitin, G. J. (1975). "Randomness and Mathematical Proof." *Scientific American* 232, May: 47–52.

Chaitin, G. J. (1988). "Randomness in Arithmetic." *Scientific American* 259, July: 80–5.

Cheney, D. L. and R. M. Seyfarth (1990). *How Monkeys See the World: Inside the Mind of Another Species*. Chicago, Chicago University Press.

Craig, E. (1990). *Knowledge and the State of Nature*. Oxford, Oxford University Press.

Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, MA ~~Mass.~~, MIT Press.

Davidson, D. (2001). *Subjective, Intersubjective, Objective*. Oxford, Oxford University Press.

[15] My thanks for helpful comments on an earlier draft by Alex Miller and for detailed and insightful comments by Joshua Gert on ~~a~~ the penultimate version of ~~the~~ this chapter.

Dennett, D. (1991). "Real Patterns." *Journal of Philosophy* 88: 27–51.

Gert, J. (2012). *Normative Bedrock: Response-dependence, Rationality, and Reasons*. Oxford, Oxford University Press.

Gilbert, M. (2015). *Joint Commitment: How ~~we~~ We Make the Social World*. Oxford, Oxford University Press.

Goodman, N. (1969). *Languages of Art*. ~~London~~Oxford, Oxford University Press.

Grice, H. P. (1957). "Meaning." *Philosophical Review* 66: 377–88.

Hart, H. L. A. (2012). *The Concept of Law*, 3rd ed~~ition~~n. Oxford, Oxford University Press.

Hornsby, J. (1980). *Actions*. London, Routledge.

Jackson, F. and P. Pettit (2002). "Response-~~dependence~~ Dependence without Tears." *Philosophical Issues* 12: 97–117.

Jackson, F., P. Pettit, and M. Smith (1999). "Ethical Particularism and Patterns." *Particularism*. London, B. Hooker and M. Little: 79–99; reprinted in F. Jackson, P. Pettit, and M. Smith~~,~~ (2004~~,~~). *Mind, Morality and Explanation*, Oxford, Oxford University Press.

Kripke, S. A. (1982). *Wittgenstein on Rules and Private Language*. Oxford, Blackwell.

Lalande, K. L. (2017). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton, NJ~~Princeton~~, Princeton University Press.

Lederman, H. (2018). "Common Knowledge." *Handbook of Social Intentionality~~.~~*, ed. M. Jankovic and K. Ludwig. London, Routledge.

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

Miller, A. (2015). "~~"~~"Blind Rule-Folllowing and the "~~'~~'Antinomy of Pure Reason~~".".~~'." *Philosophical Quarterly* 65: 396–416.

Miller, A. (2018). *The Philosophy of Language*. London, Routledge.

Moore, R. (2016). "Meaning and ~~ostension~~ Ostension in ~~great~~ Great ~~ape~~ Ape ~~gestural~~ Gestural ~~communication~~Communication." *Animal Cognition* 19: 223–31.

Neale, S. (1992). "Paul Grice and the Philosophy of Language." *Linguistics and Philosophy* 15: 509–59.

Pettit, P. (2002). *Rules, Reasons, and Norms: Selected Essays*. Oxford, Oxford University Press.

Pettit, P. (2008). *Made with Words: Hobbes on Language, Mind and Politics*. Princeton, NJ~~Princeton~~, Princeton University Press.

Pettit, P. (2017). "Corporate Agency—The Lesson of the Discursive Dilemma." *Routledge Companion to Collective Intentionality*~~.~~, ed. M. Jankovic and K. Ludwig. London, Routledge.

Pettit, P. (2018). *The Birth of Ethics: Reconstructing the Role and Nature of Morality*. Oxford, Oxford University Press.

Pettit, P. (2023). *The State*. Princeton, NJ, Princeton University Press.

Pettit, P. (2024). *When Minds Speak: A Genealogy of Human Thought and Capacity*. Oxford, Oxford University Press.

Price, H. (1988). *Facts and the Function of Truth*. Oxford, Blackwell.

Queloz, M. (2021). *The Practical Origins of Ideas: ~~Genelogy~~Genealogy as Conceptual Reverse-Engineering*. Oxford, Oxford University Press.

Searle, J. (2010). *Making the Social World: The Structure of Human Civilization*. Oxford, Oxford University Press.

Sellars, W. (1997). *Empiricism and the Philosophy of Mind*. Cambridge, MA~~Mass.~~, Harvard University Press.

Stalnaker, R. C. (1984). *Inquiry*. Cambridge, MA~~Mass.~~, MIT Press.

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

Swindlehurst, Z. M. (2020). "Blind ~~rule~~Rule ~~following~~ Following and the ~~regress~~ Regress of ~~motivations~~Motivations." *Inquiry* 64.

Tager, A., E. Kirchner, and E. Fedorovskaya (2021). "Computational Evidence of First Extensive Usage of Violet in the 1860's." *Color Research and Application* 46: 961–77.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA~~Mass~~, MIT Press.

Tomasello, M. (2009). *Why We Cooperate*. Cambridge, MA~~Mass~~, MIT Press.

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

Tomasello, M. (2016). *A Natural History of Human Morality*. Cambridge, MA~~Mass~~, Harvard University Press.

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

Williams, B. (2002). *Truth and Truthfulness*. Princeton, NJ~~Princeton~~, Princeton University Press.

Wittgenstein, L. (1958). *Philosophical Investigations*. Oxford, Blackwell.

Wright, C. (1992). *Truth and Objectivity*. Cambridge, MA~~Mass~~, Harvard University Press.

Yamamoto, S., T. Humle, and M. Tanaka (2009). "Chimpanzees Help Each Other upon Request." *PLoS OnNE* 4: e7416.

Yamamoto, S., T. Humle, and M. Tanaka (2012). "Chimpanzees' Flexible Targeted Helping ~~Basef~~ Based on an Understanding of Conspecifics' Goals." *Proceedings of the National Academy of Sciences* 109: 3588–92.

Commented [Ma1]: AU: pls supply page range