dialectica

International Journal of Philosophy

Contents

dialectica

dialectica

December 2021

Contents

Lying, Tell-Tale Signs, and Intending to Deceive

Vladimir Krstić

Arguably, the existence of bald-faced (i.e., knowingly undisguised) lies entails that not all lies are intended to deceive. Two kinds of bald-faced lies exist in the literature: those based on some common knowledge that implies that you are lying and those that involve tell-tale signs (e.g., blushing) that show that you are lying. I designed the tell-tale sign bald-faced lies to avoid objections raised against the common knowledge bald-faced lies but I now see that they are more problematic than what I initially thought. Therefore, I will discuss these lies in more detail, refine the existing cases, and resolve some anticipated objections. I conclude that tell-tale sign bald-faced lies are genuine lies not intended to deceive.

Consider this case (derived from Krstić 2020, 758–759):

> PINARTIO. A vicious murderer, Tony, is hiding from the police in Pinocchio's house. In search of Tony, the police knock on Pinocchio's door asking whether Tony is hiding in his house. Pinocchio wants to give Tony away but he is afraid that, if he gives any indication of this to Tony, Tony will hurt him. Luckily, Pinocchio knows both that the police know that his nose starts to grow at the very instant he forms the intention to lie and that they know that he knows that they know how his nose behaves, but that Tony does not know anything about this. Therefore, he asserts "Tony definitely isn't in my house" to the police. Pinocchio does this not because he intends to deceive the police in any sense (he doesn't want them to think that he is protecting a murderer), but rather because he intends to let them know that Tony *is* in his house by having them recognize the full content of his intention.

Pinocchio intends to cause the police to realize (i) that Tony is in Pinocchio's house, (ii) that Pinocchio is lying by saying that Tony is not in Pinocchio's house, and (iii) that he is lying because he intends to cause them to deductively infer the relevant true proposition from his assertion and the behaviour of his nose. Since he intends to cause them to learn the whole truth—i.e., where Tony is, that he (Pinocchio) is lying, and why—Pinocchio does not count as intending to deceive the addressee and, because the nose is an indicator of lying rather than uttering something he believes to be false, Pinocchio seems to be genuinely lying.

PINARTIO appears to be an excellent counterexample to the view according to which intending to deceive is necessary for lying. Mahon (2015) names this view *deceptionist*. Some proponents of deceptionism are Davidson (1998), Williams (2002), Derrida (2002), Faulkner (2007, 2013), Lackey (2013, 2019), Keiser (2016), Meibauer (2014a, 2014b, 2016), Maitra (2018), and Harris (forthcoming). PINARTIO is a counterexample to this view because Pinocchio does not intend to cause the police believe what he asserts (or to make them more confident in this proposition) and he does not intend to cause them to believe that he believes what he asserts (or to make them more confident in this proposition), which are the standard ways of deceiving discussed in the literature on the nature of lies. In fact, Pinocchio does not intend to mislead the police with respect to anything or to conceal any information.[1] He is trying to help them.

PINARTIO suggests that the *non-deceptionist* analysis, according to which asserting what you believe (or judge) is false is (necessary and) sufficient for lying, is correct.[2] Some proponents of the non-deceptionist view are Aquinas (*Summa Theologica*, article 1), Johnson (1755a), Carson (2006, 2010), Sørensen (2007, 2010, 2022), Fallis (2009, 2012, 2013, 2015), Saul (2012), Stokke (2013, 2016, 2017, 2018), Rutschmann and Wiegmann (2017), Krstić (2018, 2019, 2020), Marsili (2021), Sneddon (2021), and Michaelson and Stokke (2021). And while Pinocchio's lie is a so-called bald-faced lie, i.e., knowingly undisguised lie, it is different from all other cases of bald-faced lies in a very important

---

1 According to Lackey (2013, 246), liars merely need to intend to *be deceptive* towards their hearer in stating that *p*, where this may involve concealing information from the hearer regarding whether *p*. My idea is that Pinocchio is not being deceptive because he intentionally *reveals* (rather than conceals) the whole truth. Lackey could reply that it is not the saying that does the truth-revealing but rather the nose growing and that the statement is thus deceptive. This reply fails because, even if the statement was deceptive, it was not *intended* to be deceptive: the statement *plus* the nose are supposed to reveal the whole truth.

2 However, please see note 3.

way. Standardly, bald-faced lies involve situations in which the liar believes that it is common knowledge that what the liar says is false. The *common knowledge* bald-faced liar does not intend to deceive their addressee because they think that the addressee already knows the truth. Say that a gambler asserts to his wife that he was not gambling when she caught him with the betting tickets from that afternoon's races (Arico and Fallis 2013); it seems sensible to think that he did not try to make his wife believe him, since he should think that she already knows everything.

However, the deceptionists promptly responded to the argument from the existence of these bald-faced lies by saying that the proposed examples (1) are either lies intended to deceive in some of the senses I mentioned above or (2) are not genuine lies in the sense in which they do not involve genuine assertions—and you need to assert in order to lie.[3] These replies do make the common knowledge bald-faced lies much less effective. A desperate gambler may hope that his lie could cause his wife to become slightly less confident in her true belief and he could add that the ticket belongs to a friend. Alternatively, it might be that, since he does not intend his wife to believe him, the gambler does not assert what he says but rather merely makes it look like he does; he could be playing a kind of a (language) game, he could be doing something similar to acting or even to being verbally aggressive (e.g., Keiser 2016; Maitra 2018; Harris forthcoming; Meibauer 2014a; against, e.g., Marques 2020; Viebahn 2019a; Marsili 2021).

These two general objections are well-known and, in one form or another, they have been put forward in detail by many contemporary philosophers (e.g., Faulkner 2007, 2013; Kenyon 2010; Lackey 2013, 2019; Meibauer 2014a, 2014b, 2016; Leland 2015; Dynel 2011, 2015; Hawley 2018; Keiser 2016; Maitra 2018; Harris forthcoming). I do not intend to discuss their application to common knowledge bald-faced lies for two reasons. The first is that some replies already exist (e.g., Fallis 2015; Stokke 2017, 2018; Marques 2020; Viebahn 2019a; Marsili 2021). The second, and more important, is that I designed cases like PINARTIO to avoid both deceptionist objections. PINARTIO involves what I named a "tell-tale kind" bald-faced lie (Krstić 2020), in which the addressee not only does not know the truth (no common knowledge) but rather *learns* the truth by observing the behaviour of the lie-disclosing sign and, vitally,

---

3 Not all scholars believe that the only way to lie is by asserting what you say—some think that one may lie by implicating false information (e.g., Meibauer 2014b; Reins and Wiegmann 2021; Wiegmann, Willemsen and Eibauer 2022), or by adding false presuppositions (Viebahn 2019b) or by making false promises (Marsili 2016, 2021)—but this is the predominant view.

Pinocchio *intends* this to happen. He wants the police to deductively infer the truth from his false assertion and the behaviour of his nose. Therefore, the lie is not intended to deceive the addressee in any sense. Since it is clear that the speaker does not intend to deceive, PINARTIO seems to be an importantly different counterexample to the deceptionist analysis of lying; it gives us a new perspective on the issue and thus it opens the door for a new and promising debate. But we have yet to see a reply to my argument and Pinocchio-like cases.

This paper is designed to fill this void and bring the debate regarding lies and intending to deceive closer to a fruitful end. In section 1, I argue that cases such as PINARTIO are not uncontroversial but that more convincing cases can be developed from them. In section 2, I discuss two objections to my argument. In section 2.1, I reject the popular objection according to which bald-faced lies are not genuine lies on the count of them not involving genuine assertions. In section 2.2, I argue against the interpretation that tell-tale liars (indirectly) assert not the literal meaning of the descriptive sentence they utter but rather the proposition they want the hearer to infer from their behaviour. In section 3, I conclude my argument.

1 Finding the Right Case

My original cases do avoid some standard objections but I see now that they also generate new problems. In this section, I discuss these problems and offer a case that avoids them. In the next section, I discuss two further objections that arise.

In my analysis of PINARTIO, I write:

> The non-[deceptionist] definition counts this [Pinocchio saying "Tony definitely isn't in my house"] as lying because Pinocchio asserts what he believes is false, I count this as lying because the nose grows, and the [deceptionist] definition does not count this as lying because Pinocchio does not intend to deceive his addressee notwithstanding the fact that *the nose indicates that Pinocchio is lying*—this is why (PINARTIO) is a counterexample to the [deceptionist] definition. [Krstić (2019), 653; italics added]

In short, I argue that PINARTIO is a counterexample to the deceptionist analysis because (i) Pinocchio does not intend to deceive his addresses and (ii)

the nose indicates that he is lying. PINARTIO cannot be dismissed on the count of it being a fairy-tale case. True, Pinocchio is a fictional character and the situation I put him in is uncommon. However, this is not relevant to the question under discussion. My main idea is to present a situation in which a speaker *believes* (correctly or incorrectly, it does not matter) that their lie will undoubtedly be disclosed to the addressee and they use this circumstance to communicate the truth. In real life, police may convince their suspect that their lie detector makes no mistakes and there could be a real-life person who mistakenly believes that they always blush when they lie or that their interlocutor is some kind of a holy person who can read their thoughts or a person skilled enough in detecting lie-betraying cues. In *Meet the Parents* (Universal Pictures, 2000), Jack Byrnes convinces his future son-in-law, Greg Focker, that he will unmistakably detect when Greg lies just by feeling Greg's pulse. I merely flesh this situation out using the character of Pinocchio as a communicative device.

Moreover, the issue of whether the nose will grow or not is irrelevant: the deceptionist definition assumes that the *intention* to deceive is necessary for lying. If the nose does not grow, Pinocchio will still lie; he will just fail to communicate the truth. The examples are effective as long as (i) *the liar believes* that something will show that he is lying and (ii) the liar and their audience are in a standard context (i.e., unless some further conditions apply) in which it is common ground that one asserts what one says. Thus, even if Pinocchio misleads the police by asserting what he says, this would be against his intention, which still sits uneasily with the deceptionist analysis. Consider the following case.

ARTOCCHIO. A vicious murderer, Tony, is hiding in Artie's house. The police come to question Artie about Tony's whereabouts but Artie is too afraid to tell them that Tony is right here—Tony might hear him. Luckily enough, Artie believes that he blushes only and always when he lies and that Tony does not know about this. Artie decides to use this to let the police know the truth without thereby alarming Tony. Therefore, he asserts "Tony is *not* in my house" to them excepting that he will immediately start blushing, that this will be a clear sign to the police that he is lying, and that they will realise from this that Tony *is* in Artie's house. Of course, Artie blushes not because blushing is somehow connected to whether he believes

what he asserts but rather due to the highly stressful situation he is in.[4]

In Artocchio, Artie believes that there is a perfectly reliable giveaway of lying and he intends to use it to cause the police to infer the truth from his lie. Artie's blushing is not a real tell-tale sign of lying but it worked as one nonetheless. Furthermore, Artie and the police are in a standard context, he thinks that he will be understood as asserting what he says and they expect him to assert what he says, and thus there seem to be no reasons to think that he did not assert what he said. Importantly, Artie's blushing is not analogous to winking or finger-crossing: the function of winking is set by a convention that is a part of common ground (the context is not standard) and Artie's blushing is not. Therefore, we should think that Artie is understood as asserting what he says and his plan is not unwise: if the police are sufficiently attentive to detail, they will realise that he is lying and why he is lying. Nevertheless, there is an important concern that Tony was intentionally caused to believe falsely that Artie deceived the police. Tony was misled, that is; just as Artie planned.

In Krstić (2019, 656), I argue that the issue of whether Artie intended to deceive Tony is irrelevant for the question of whether lies must be aimed at deceiving—since Artie was addressing the police, not Tony. However, it may be that Artie was addressing Tony after all: Artie needed Tony to hear what he is saying. Hence, Tony does appear to be Artie's intended hearer and, it seems to follow, Artie did intend to deceive someone by lying—Tony.[5] The conclusion that Artie intends to deceive Tony by lying, however, does not follow. Tony expects and coerces Artie to say to the police that Tony is not in Artie's house and Artie says this because he knows that Tony expects him to do it. Therefore, Artie cannot reasonably intend to cause Tony to believe as true a proposition for which Artie knows that Tony knows is false. Thus, even if Artie addresses Tony, he cannot be *lying* to Tony in the sense in which this requires intending to deceive him. In fact, he does not seem to be lying at

---

4 In my original version of Artocchio (Krstić (2019), 655), "Artie *believes* that he always stutters when he lies," which leaves it open whether Artie may also believe that he always stutters when he utters something false in a way that does not count as lying (e.g., when being sarcastic). I avoid this ambiguity by saying that Artie believes that he blushes "only and always" when he lies. Also, rather than "telling" the truth, in this version, Artie lets the police know the truth.

5 Goffman (1981) divides hearers into ratified (official) and unratified. Ratified hearers can legitimately listen to the speaker whereas unratified cannot; they are bystanders.

all: he just says what Tony wants him to say.[6] Say that a company manager orders his assistant to present false data at the board meeting in exchange for a promotion. This assistant will lie to other board members but not to his boss. The same applies to Artie.

So, it seems as if Artie lies while not intending to deceive anyone by asserting what he says. However, whether ARTOCCHIO is enough to reject the deceptionist view depends on how broad our analysis of lying is. On the view that one may lie only by asserting something one believes is false, ARTOCCHIO vindicates the non-deceptionist definition. However, broader analyses exist (see footnote 3) and so one may argue that Artie's lie is intended to deceive Tony—only in a sense that does not involve asserting. Artie may, for instance, intend to deceive Tony by falsely implicating that he led the police off Tony's track. On this analysis, then, Artie both counts as asserting what he says and as intending to deceive by lying (i.e., by implicating false information to Tony). Therefore, while ARTOCCHIO does appear to put reasonable pressure on the deceptionist analysis, there still are some controversies about it. Most of these controversies can be eliminated with simple modifications. Consider the following case.

> WITNESSIO. A gruesome murder happened in Artie's bar. The police do not know who did it but Artie knows that Tony did it. Artie wants Tony off his back but he is afraid to testify against Tony. Artie believes that he blushes always and only when he lies, he believes that the police know this as well, and he decides to use this to let the police know that Tony is the murderer without actually testifying against him (thus avoiding the imminent retribution). Tony went to Polly's place to hide the murder weapon and Artie knows this. Therefore, Artie says to the police "Maybe you could talk to Tony? A minute ago, he rushed to Polly's house regarding a matter of great urgency. Tony definitely did *not* commit the crime." Artie hopes that he will start blushing while uttering the last sentence, that this will be a clear sign to the police that he is lying, and that they will realise from this that he wants to let them know both that Tony committed the crime and where they can find him. Artie hopes that they will catch Tony with the murder weapon.

Witnessio and Artocchio involve a rather plausible idea that some people may think that something will unmistakably show that they are lying and they are similar to the context of two recent movies. The first is *Knives Out* (Lionsgate, 2019) in which the character of a nurse, Marta Cabrera, cannot lie without vomiting ("Just the thought of lying [...] It makes me puke," she says), the detectives that are questioning her know this, she knows that they know, and indeed she does vomit every time she lies. The second is *Meet the Parents* (Universal Pictures, 2000), in which Jack Byrnes convinces his future son-in-law, Greg Focker, that he will unmistakably detect when Greg lies just by feeling Greg's pulse (the pulse is a lie-betraying sign). The main difference is that, on the one side, Marta avoids lying by giving true but incomplete answers to questions when asked and Greg simply goes for telling the truth, whereas, on the other side, Artie comes up with a plan to make lemonade when life gives him lemons: he decides to use what he thinks is his otherwise unfortunate reaction to his advantage and cause the police to learn the truth by lying to them. Another difference is that, if Artie blushes at the right moment, this will be not because blushing is somehow connected to whether he lies but rather because of the highly stressful situation he is in.

In Witnessio, Tony did not end up with a false belief; rather he was left without any belief regarding Artie's conversation with the police. More importantly, not only is Artie not addressing Tony, he hopes that Tony never learns about the conversation. Therefore, neither did Artie intend to deceive Tony, nor was he lying to Tony. Even if one can lie by making false promises, implicatures, or presuppositions, Artie does not lie to Tony in any of those senses; he is not addressing Tony. While Artie does not lie intending to deceive the police, misleading them about Tony's whereabouts or Artie's intentions is possible in Witnessio (they do not believe that Artie blushes always and only when he lies) but this is irrelevant for our discussion. We are not analysing whether epistemic harm will be caused but rather whether the liar *intends* to cause it or whether he *expects* that he might cause it (see Krstić 2020, sec.2.1) and Artie clearly lacks the intention to deceive them and he does not expect them to end up misled (he believes that they will understand why he blushes). Finally, because the belief that Artie blushes always and only when he lies is not common ground between Artie and the police, we cannot say that blushing implies that Artie does not assert what he says.

Witnessio avoids the most obvious objections, but it fails to avoid all objections. I anticipate two. The first is that, because he does not intend to give the police a reason to believe what he says, Artie does not assert what

he says and thus does not lie. The second is that Artie (indirectly) asserts the proposition he intends the police to infer from his behaviour and the literal meaning of the uttered proposition ("Tony committed the crime") and, because he believes this proposition to be true, he does not count as lying. I will resolve these objections in turn.

2 Objections

2.1 *Not an Assertion*

According to one influential analysis of assertion (Gricean in nature), I asserted that *p* by uttering *x* if and only if I uttered *x* intending to induce in you the belief that *p* or give you grounds for believing it by means of your recognition of the full content of my intention (e.g., Bach and Harnish 1979; Récanati 1987; similarly, Peirce 1935; Grice 1989).[7] Accounts of this sort are typically called *Gricean* or *Neo-Gricean*, Keiser (2016) calls them *epistemic*, and Harris (forthcoming) and Siebel (2020) call this approach to communication *intentionalism*. I will refer to the view as Neo-Gricean analysis of assertion.

According to one influential argument based on the Neo-Gricean analysis of assertion, because he does not intend to give his audience grounds for believing what he says because he says so, Artie does not count as asserting what he says in Witnessio and therefore—on the popular assumption that lies are a subset of assertion—he does not lie (e.g., Chisholm and Feehan 1977; Meibauer 2014a, 2014b, 2016; Keiser 2016; Harris forthcoming, 13, 15). That is to say, I assume that one would object that, while there is nothing in the context that could prevent Artie's utterance to count as an assertion, Artie did not act on an intention constitutive of asserting and *this failure* (rather than pragmatic considerations or linguistic conventions) disqualifies the utterance from counting as an assertion and a genuine lie (which requires asserting what you say).

Many influential philosophers think that this is a very serious problem for any non-deceptionist analysis of bald-faced lies. The idea is that, because Neo-Griceans think that the ordinary language concept of lying is too ambiguous,

---

7 Although Grice did not explicitly attempt to define assertion, Pagin and Marsili (2021) argue that his analysis of non-natural meaning can be straightforwardly applied to provide one. Peirce (1935, 547) writes that asserting involves giving a reason to believe what is said but his account also assumes that the asserter makes certain commitments.

they are looking for a definition of lying that fits neatly within their definition of assertion, which in turn fits neatly within a relevant broader theory of speech acts and communication. Therefore, they will maintain that bald-faced lies are not genuine lies because they do not involve genuine assertions. While Artie believes that his blushing signals that he is lying when he is lying, his conception of lying is too broad and he does not intend to lie in the relevant sense (which involves asserting what you say).

I will reject this objection in three steps, where each step gives my argument a premise. In step 1, I argue (by analogy) that the argument from the objection generates an unfalsifiable position. In step 2, I argue that it entails that competent language users unreliably track assertions, which is a very bold and empirically unsupported claim. The first two steps only show that the objection is much less serious than what initially seems; they do not show that it fails. The third step, however, shows that the argument fails. In step 3, I argue that, if understood as not allowing exceptions (such as bald-faced lies), the Neo-Gricean analysis of assertion misclassifies some sincere assertions as not assertions: some sincere asserters do not act on the intention to give their hearers grounds for believing what they say. I conclude that, considering steps 1–3, this specific argument claiming that bald-faced lies are not genuine assertions fails. I now proceed to step 1.

In his *Introductory Lectures on Psycho-analysis (1916–1917)*, Freud defends his dream theory from the concern that dreams do not reveal our unconscious mental life. His defence is strikingly similar to the comeback presented above: he dismisses counterexamples by reinterpreting them. The following is one such case.

> A woman dreamer says: "Am I supposed to wish that my husband were dead? Really that is outrageous nonsense! Not only is our married life very happy, though perhaps you won't believe that, but if he died I should lose everything I possess in the world." (Freud 1929, 121)

This woman directly challenges Freud's diagnosis in the same way Witnessio challenges the deceptionist account of lying (Artie thinks that he is lying; the woman thinks that she wants her husband alive). Freud's answer is very interesting.

> *Assuming that* unconscious tendencies do exist in mental life, the fact that the opposite tendencies predominate in conscious life

> *goes to prove nothing*. […] What does it matter if *you* [the woman] do find the results of dream-interpretation unpleasant, or even mortifying and repulsive? "*Ça n'empêche pas d'exister*" ["It doesn't prevent things from existing"]. [Freud (1929), 122; italics added]

Freud practically says that the woman incorrectly believes that she does not want her husband dead because her desire is unconscious. This reply raises two important problems (see Derksen 2001). According to the reply, Freud's theory is correct no matter what the woman says: the only difference is in whether the desire is conscious or unconscious. A theory defended in this way can never be disproved since any testimony becomes evidence in support of it. Therefore, the reply makes Freud's position unfalsifiable; the first flaw. The second flaw is that the comeback is viciously circular (i.e., it begs the question). Freud correctly says "*assuming that* unconscious tendencies do exist in mental life [i.e., that his theory is correct]," the fact that this woman's testimony contradicts his "goes to prove nothing." But, whether his theory is correct is exactly what is at stake. This testimony is a counterexample to his theory and the theory cannot be used as a reason to disregard it.

The argument to the conclusion that bald-faced lies do not involve genuine assertions involves the same fallacious line of reasoning. This is the analogy between the two replies: Freud says that, if the woman concurs with his diagnosis, her desire is conscious and, if she denies it, the desire is unconscious, she just thinks that she does not have the desire. Analogously, according to the given argument, if speakers intend to deceive by lying, then their lies involve genuine assertions and, if they do not intend to deceive, then their "lies" do not involve genuine assertions; these speakers just think that they are genuinely lying.

And here is the immediate problem with this comeback: if we cannot trust the speakers' judgements as to whether they are lying or not, then no testimony can be used as a counterexample to the given analysis of assertion. Therefore, the reply causes the view to become unfalsifiable. This reply uses the idea that intending to give grounds for beliefs is *necessary* for asserting to discredit cases of bald-faced lies but whether this intention really is necessary for asserting is exactly what is at stake in this debate. Therefore, the reply begs the question.

The circularity is actually very visible in this argument. What the reply is actually saying is that the deceptionist definition of lying fits neatly within the Neo-Gricean definition of assertion, which in turn fits neatly within a

relevant broader Gricean theory of speech acts and communication, but this is the same theory. In effect, then, the argument says that *assuming that Neo-Griceanism is correct*, the fact that other people (e.g., Artie, Pinocchio, Marta) see bald-faced lies as genuine lies goes to prove nothing; their testimonies do not matter. This is not to say that Neo-Griceanism is incorrect but this specific defence is problematic and, as philosophers, we should be basing our views on good arguments. Therefore, because it is viciously circular and unfalsifiable, this argument should not be accepted unconditionally. I now proceed to argue that it also makes a very daring claim; this is step 2 in my argument.

This dispute is not only about what people (Artie, Marta Cabrera, etc.) recognize as lying but also about what they recognize as asserting. Therefore, in discrediting bald-faced lies in this particular way, one is not just saying that common folks have a broader conception of lying, but also that they cannot recognize when a proposition is being asserted.[8] This strikes me as a rather bold position, which is the second premise in my argument, step 2. In a standard context, assertion simply seems to be a default interpretation of a declarative sentence, and this view seems to be common ground in the debate. Williamson (2000, 258), for instance, writes: "In natural language, the default use of declarative sentences is to make assertions." That being said, even though the idea is bold, it is not completely unwarranted. Consider the following argument made by Keiser (2016).

In *The Godfather 2* (Paramount, 1974), "Frankie Five Angels" Pantangelli is called in as a surprise witness in a Senate hearing against the mob boss Michael Corleone but, to everyone's utter shock, he goes against the agreement and claims under oath that he has no knowledge of any wrongdoings committed by Michael Corleone. Keiser (2016, 471) argues that, since he does not intend to give his audience grounds for believing what he says, Frankie does not assert what he says; rather, he is playing a "courtroom [language] game," a game in which a speaker can avoid asserting what they say but still achieve a specific intended effect (e.g., go for the record).

People standardly think that witnesses assert statements they make while testifying under oath—witnesses assume many assertoric commitments (they, e.g., guarantee that what they say is true)—and the terms "lying on the stand" or "lying under oath" are standardly understood as lying by asserting what

---

8 Harris (forthcoming, 7) writes that adherence to ordinary usage should be even less appealing in the case of assertion since the term is technical and the term rarely shows up in ordinary usage. Against this argument, see Krstić and Wiegmann (2024, sec.5).

the witness believes is false. According to Keiser, however, Franky was just making a move in a courtroom game. If he lied, this was not in the sense we are discussing here (it does not involve the default use of a declarative sentence) and our intuitions about the case are incorrect. Following this analysis of *Frankie*, one may say that, while we may allow that common folk can recognize asserting in standard contexts, the cases I discuss are all fictional and very unusual (nose growing, blushing). Therefore, even if it is true that people can reliably detect assertions in standard contexts, this is of limited value for my argument.

Keiser does not offer an unreasonable interpretation of Frankie, people's intuition may go in the wrong direction in the "courtroom" context, but the idea that bald-faced lies are not genuine assertions is not only bold, it is also controversial. For one, choosing *Frankie* to support a very general claim—namely, that *no* bald-faced lie involves a genuine assertion—is rather unfair: *Frankie* involves a specific, non-standard context that allows Keiser to apply her "game" analogy but people predominantly lie in standard contexts. Therefore, we cannot apply insights from Frankie to all cases of bald-faced lying: the analogy breaks. People would have been equally shocked if Frankie had said that he knew no Godfather in a standard context far away from the courtroom (e.g., in a private conversation), they would have been shocked because they would think that he is lying, and Keiser's argument would not apply here as easily as in the courtroom context.

Gambler, for example, involves a standard context and there are good reasons to say the same about Witnessio. After all, Artie is not in a courtroom, and the police did not arrest him or bring him in for questioning; they were just talking. Consider the following combination of *Gambler* and Witnessio.

> Gamblessio. Tony has a gambling addiction and Tony's wife, Carmela, knows this. Tony lies to Carmela by saying that he has quit gambling and he makes Artie keep his betting ticket. However, Artie accidentally drops it when he goes to Tony's place to pick up a thing and Carmella sees the ticket. Artie is worried about Tony but he does not want to openly tell Carmela that Tony did not stop gambling; he does not want to hurt Tony's feelings by betraying his trust. However, Artie believes that he blushes always and only when he lies, he believes that Carmela knows this as well, and he decides to use this to let her know that the ticket belongs to Tony without actually saying this out loud (which would give him an excuse in

front of Tony). Therefore, Artie says to her "Don't worry, Carmela, the ticket does not belong to Tony, it's mine" hoping that he will start to blush and that Carmela will infer from this that Artie is lying, that the betting ticket is Tony's, and that Artie is trying to preserve everybody's dignity by acting this way (it's a kind of a prosocial tell-tale sign bald-faced lie).

Gamblessio preserves the virtues of Witnessio while making the stakes lower. Therefore, even though the position according to which our intuitions about lying may be unreliable may make sense when applied to cases such as *Frankie*, it fails to easily generalise to all cases of bald-faced lying.

As I argued in step 2, the claim that no bald-faced lie is an assertion is very bold: it entails that many people are not competent speakers. Of course, there is nothing wrong with making bold claims *per se* but, other things being equal, we should go for less demanding claims. And other things are not equal: this bold argument not only begs the question and generates an unfalsifiable position, it also suffers from three additional problems. Two can be immediately noticed. One problem is that this position cannot be easily generalised to all cases of bald-faced lies. I discussed this problem here. It is not really obvious that we can say that Artie and Carmela do not know what it is to assert a proposition in the context of Gamblessio. Another problem is that Artie and Pinocchio assert according to many successful accounts of assertion and so we do not need to commit ourselves to very demanding positions.

Other views will say that Artie and Pinocchio assert what they say because they take themselves as being in a warranting context (Saul 2012), because they propose that what they say be added to official common ground (Stalnaker 1984, 1999, 2002; Stokke 2013, 2016, 2017, 2018), because they represent themselves as believing what they say (Black 1952; Davidson 1998; Fallis 2013) and even as knowing what they say (Unger 1975, 250–270; DeRose 2002, 185). They also count as asserting what they say because they make many assertoric commitments: they warrant the truth of what they say (Carson 2006, 2010), they undertake the responsibility of justifying their assertion and what follows from it (Brandom 1994, 173–175), they commit themselves to act in accordance with what they say (Dummett 1981) or that they will withdraw it if the proposition is shown to be untrue (MacFarlane 2005; similarly, Dummett

1991, 165), they commit themselves to the truth of what they say (Marsili 2021; similarly, Dummett 1981, 300), and so on.[9]

With all of this in mind, while one need not think that Neo-Griceanism delivers a failed analysis of assertion, one must wonder whether making bold claims just to keep a particular interpretation of Neo-Griceanism is justified. It may just be that Neo-Griceanism allows for exceptions.[10] If this is correct, we get to keep both the Neo-Gricean analysis of assertion and the idea that some lies not intended to deceive are genuine assertions. This is why, in my final step 3, I argue that the intention to give grounds for believing what you say cannot be necessary for asserting what you say; this is the additional third problem this bold position faces. Consider a real-life case involving a sincere speaker.

> ARANGIO. Stephen Miller puts credence 0 in the proposition that refugees benefit the American economy more than they cost. Jennifer Arangio, a lower-level aide who has looked at the relevant studies, has credence 1 that refugees benefit the American economy more than they cost. Arangio is well aware that, whatever she says, Miller's credence in this proposition will not be shifted one bit. Nevertheless, she tells Miller the truth and thereby risks her job.[11]

Because she is well aware that, whatever she says, Miller's credence in the proposition will not be shifted one bit, ARANGIO cannot reasonably intend to give Miller grounds for believing what she says based on her say-so. Therefore, according to the given argument, because she cannot rationally intend to give Miller grounds to believe what she says, ARANGIO cannot rationally assert to Miller a proposition she believes to be true. In other words, according

---

9 For more analyses of assertion, see Pagin (2015).

10 Krstić and Wiegmann (2024, sec.5) offer one plausible NEO-GRICEAN ANALYSIS OF ASSERTION that does not sit uneasily with the existence of bald-faced lies. The suggestion is that Neo-Griceans may simply hold that, just as a certain company systematically pays men higher salaries than women (some men will still be less paid than their female peers), asserters *systematically* intend to give grounds for believing that *p* by asserting this. Bald-faced lies now can count as genuine assertions because lies are systematically, rather than necessarily, intended to deceive.

11 Please do not confuse ARANGIO with a case discussed by Benton (2018). In ARANGIO, Miller's credence in *p* is 0 and nothing can change his mind simply because he is unresponsive to reasons whereas, in Benton's (third) case, B's credence in *p* is 1 and this is why A cannot make A more confident in *p*. However, in Benton's case, A can give B a reason not to become less confident in *p*. ARANGIO was suggested to me by an anonymous reviewer of one of my earlier papers that does not discuss the connection between lying and intending to deceive.

to a consistent application of the argument claiming that no bald-faced lie is an assertion, in this situation, it is *impossible* for Arangio to assert a proposition she believes is true. This result is surely counterintuitive: the issue of whether it is *possible* for me to sincerely assert something to you should depend on me (i.e., on whether I can utter the proposition, etc.), not on you (i.e., on whether you will believe me or not). Therefore, we should think that a consistent application of the deceptionist argument misclassifies some sincere assertions.

Given the arguments from steps 1–3, this particular argument cannot be used as a reason to say that bald-faced lies are not genuine assertions. In particular, the line of reasoning is such that it misclassifies some sincere assertions, because the argument begs the question, generates an unfalsifiable position, and is difficult to generalise to all bald-faced lies, we can safely assume that it is not a reason to think that bald-faced lies are not genuine assertions.

However, my argument needs to resolve one more issue: we need to see whether Artie asserted the proposition he uttered ($\neg p$) or the proposition he intended his hearers to infer (p). For, if he asserted the latter, then he asserted what he believed was true and thus did not lie. I discuss this interpretation below.

2.2 *Indirect Assertion*

The sentence "He's (She's) a friend of Dorothy" in the early 20th century US and British homosexual subculture made a claim about a person who was a homosexual.[12] Because expressing their sexual orientation was a criminal offence, homosexuals had to hide it. This sentence made it possible for people to say that a certain person is gay without uttering that proposition. With that in mind, consider this situation.

> Dorothy. Will and Grace, both familiar with the terminology of US and British homosexual sub-culture, are at a party where they meet Grace's friend Bill. Grace notices that Will fancies Bill. Thus, when Bill goes to order a drink, Grace says to Will: "Bill's a friend of Dorothy, you know. Why don't you buy him that drink?"

---

12 Possibly, "Dorothy" refers to Dorothy from *The Wizard of Oz*, who accepted those who are different.

Arguably, Grace asserts that Bill is gay because she means "Bill is gay" when she utters "Bill's a friend of Dorothy" and because they are in a context in which it is common ground that she asserts the proposition she means (this is what she says) rather than the proposition she utters (the literal meaning of the uttered sentence). Our intuitions seem to correspond with this interpretation: it is natural to think that Will will think that Grace lied to him if Bill turns out not to be gay but not if it turns out that Bill does not know a girl named Dorothy.

DOROTHY highlights a difficulty that may arise concerning my cases: in these cases, someone is intending to communicate proposition p by means of uttering $\neg p$ under certain circumstances, and hence it may be that they are asserting p, rather than $\neg p$. That is, it may be that Artie actually asserts that Tony committed the crime (p) by uttering "Tony definitely did *not* commit the crime" ($\neg p$). The means by which Artie asserts p (i.e., by uttering $\neg p$) is unusual, but it is hardly impossible to assert one proposition by uttering another.[13]

Indirect assertion is not an uncontroversial concept (see, e.g., LePore and Stone 2014; García-Carpintero 2018) but this interpretation of WITNESSIO is plausible and interesting enough to be seriously considered. I assume that the relevant analysis of my cases would go something like this. Just as Grace intends that her utterance "He's a friend of Dorothy," in the light of the relevant subculture's linguistic conventions, means "Bill is gay," Artie intends that his utterance "Tony did *not* commit the crime," in the light of his blushing, be understood as meaning "Tony committed the crime." And, because he asserts what he says (standard context) and he says what he means rather than what he utters, Artie is not lying—since he believes that the meant proposition is true.

It is pertinent to note that this interpretation is not consistent with the main idea behind the PINARTIO-style examples. Artie believes that his blushing signals that Artie *believes* the opposite of what he says (i.e., it signals that Artie lies). He does not believe that blushing signals that he *means* the opposite of the literal meaning of his utterance.[14] However, this is not a reason not to

---

13 I thank the anonymous reviewer for bringing this important concern to my attention.

14 According to Maynard Smith and Harper (2004), signals have evolved specifically to alter the receiver's behaviour, whereas cues are incidental sources of information detected by unintended receivers. Consider engaging in "cue mimicry" (mimicking a cue of another organism). The predatory jumping spider (*Portia fimbriata*) attracts orb-web spiders (*Zygiella x-notata and Zosis geniculatus*) by vibrating their web to resemble a fly struggling (Tarsitano, Jackson and Kirchner

consider this interpretation, my descriptions of the cases could be misguided. Let us therefore consider how the received analyses of assertion explain Artie's and Grace's behaviour.

Artie warrants the truth of *the uttered proposition* (the literal meaning of the uttered declarative sentence) rather than the truth of the proposition he intends the audience to infer from his behaviour. He proposes that the uttered proposition be added to official common ground, he represents himself as believing or knowing the uttered proposition ("Tony is not the murderer"), he commits himself both to the truth of this proposition and to act in accordance with this proposition. This is in clear contrast with Grace's behaviour. Grace warrants the truth of the "Bill is gay" proposition, she proposes that "Bill is gay" be added to official common ground, she represents herself as believing that Bill is gay, she commits herself both to the truth of this proposition and to acting in accordance with this proposition. Therefore, the fact that it does not seem odd to think that Dorothy indirectly asserts "Bill is gay" is not a reason to think that Artie indirectly asserts "Tony is the murderer."

We see that neither Pinocchio nor Artie is willing to accept any assertoric responsibility for the communicated propositions but only for the literal meaning of the uttered declarative sentences. Grace, however, does seem to be taking assertoric responsibility for the "Bill is gay" proposition. Therefore, while the idea that Artie indirectly asserts the proposition he intends the police to infer from his blushing and the uttered proposition is rather interesting, it does not seem to capture the relevant cases in the right way. Artie and Pinocchio intend to cause their hearers to realise that they are non-deceptively lying to them. This is vital for the success of their plan: the hearers should infer the truth from the fact that the blushing and the nose growing show that Artie and Pinocchio are lying.

3 Concluding Remarks

The idea that some lies can be intended to communicate the truth by having the hearer recognize that the speaker is lying is both plausible and important.

2000). The web vibrations of a struggling fly are cues, not signals: the fly is trying to set itself free rather than signal the orb-web spider to come down. Nevertheless, the predatory jumping spider is *using* the web vibrations to lure orb-web spiders in; therefore, this is a signal rather than a cue. Analogously, cues such as blushing when lying are not signals *per se*; however, when Artie *uses* blushing to send a certain message, it is a signal—since the idea is to alter the receiver's behaviour.

By showing that one can lie without intending to deceive anyone, tell-tale sign bald-faced lies also vindicate the non-deceptionist interpretation of the common knowledge bald-faced lies. Since we now know that lying without intending to deceive is possible, we can think that common-knowledge bald-faced liars—the gambler, for instance—could be genuine liars who did not intend to deceive. The tell-tale kind of bald-faced lies, thus, takes the debate out of the impasse and suggests that the deceptionist analysis of lying should be abandoned.

We can now move on and focus on other aspects of lying. For example, we can start analysing scenarios in which people typically lie to themselves and try to identify their motivation for such behaviour. Lying to myself is an *intra*personal analogue of *inter*personal tell-tale sign bald-faced lies: I will immediately know when I form the intention to lie to myself, I will know that I will know this, etc. Therefore, tell-tale sign bald-faced lies can help us to understand a much bigger class of human behaviour. Because a bald-faced self-liar will probably have similar motives as a bald-faced interpersonal liar, understanding other people's behaviour—namely, why others bald-faced lie to us—will help us to understand our own behaviour. And *vice versa*, we will be able to understand why other people lie better if we investigate our own motives for lying to ourselves. So, I suggest that this is the direction in which our analysis of lying should take.*

Vladimir Krstić
United Arab Emirates University
krstic.v@uaeu.ac.ae; drpop1@yahoo.com

References

ARICO, Adam J. and FALLIS, Don. 2013. "Lies, Damned Lies, and Statistics: An Empirical Investigation of the Concept of Lying." *Philosophical Psychology* 26(6): 790–816, doi:10.1080/09515089.2012.725977.

BACH, Kent and HARNISH, Robert M. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Massachusetts: The MIT Press.

BENTON, Matthew A. 2018. "Lying, Accuracy and Credence." *Analysis* 78(2): 195–198, doi:10.1093/analys/anx132.

Black, Max. 1952. "Saying and Disbelieving." *Analysis* 13(2): 25–33. Reprinted in Black (1954, 46–57), doi:10.1093/analys/13.2.25.

—. 1954. *Problems of Analysis: Philosophical Essays*. Ithaca, New York: Cornell University Press.

Brandom, Robert B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Massachusetts: Harvard University Press.

Carson, Thomas L. 2006. "The Definition of Lying." *Noûs* 40(2): 284–306, doi:10.1111/j.0029-4624.2006.00610.x.

—. 2010. *Lying and Deception: Theory and Practice*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199577415.001.0001.

Chisholm, Roderick M. 2001. *Ethics and Intrinsic Values*. Philosophy and Realist Phenomenology n. 12. Heidelberg: Carl Winter Universitätsverlag.

Chisholm, Roderick M. and Feehan, Thomas D. 1977. "The Intent to Deceive." *The Journal of Philosophy* 74(3): 143–159. Reprinted in Chisholm (2001, 141–160), doi:10.2307/2025605.

Davidson, Donald. 1998. "Who Is Fooled?" in *Self-Deception and Paradoxes of Rationality*, edited by Jean-Pierre Dupuy, pp. 1–18. Stanford, California: CSLI Publications. Reprinted in Davidson (2004, 213–230).

—. 2004. *Problems of Rationality. Philosophical Essays Volume 4*. Oxford: Oxford University Press, doi:10.1093/0198237545.001.0001.

Derksen, Anthony A. 2001. "The Seven Strategies of the Sophisticated Pseudo-Scientist: a Look into Freud's Rhetorical Tool Box." *Journal for General Philosophy of Science / Zeitschrift für allgemeine Wissenschaftstheorie* 32(2): 329–350, doi:10.1023/A:1013100717113.

DeRose, Keith. 2002. "Knowledge, Assertion, and Context." *The Philosophical Review* 111(2): 167–203, doi:10.1215/00318108-111-2-167.

Derrida, Jacques. 2002. *Without Alibi*. Stanford, California: Stanford University Press, doi:10.1515/9781503619890.

Dummett, Michael A. E. 1973. *Frege: Philosophy of Language*. London: Gerald Duckworth & Co.

—. 1981. *Frege: Philosophy of Language*. 2nd ed. Cambridge, Massachusetts: Harvard University Press. First edition: Dummett (1973).

—. 1991. *The Logical Basis of Metaphysics*. London: Gerald Duckworth & Co.

Dynel, Marta. 2011. "A Web of Deceit: A Neo-Gricean View on Types of Verbal Deception." *International Review of Pragmatics* 3(2): 139–167, doi:10.1163/187731011X597497.

—. 2015. "Intention to Deceive, Bald-Faced Lies, and Deceptive Implicature: Insights into Lying at the Semantics-Pragmatics Interface." *Intercultural Pragmatics* 12(3): 309–332, doi:10.1515/ip-2015-0016.

Fallis, Don. 2009. "What is Lying?" *The Journal of Philosophy* 106(1): 29–56, doi:10.5840/jphil200910612.

—. 2012. "Lying as a Violation of Grice's First Maxim of Quality." *Dialectica* 66(4): 563–581, doi:10.1111/1746-8361.12007 .

—. 2013. "Davidson was Almost Right about Lying." *Australasian Journal of Philosophy* 91(2): 337–353, doi:10.1080/00048402.2012.688980.

—. 2015. "Are Bald-Faced Lies Deceptive after All?" *Ratio* 28(1): 81–96, doi:10.1111/rati.12055.

FAULKNER, Paul. 2007. "What Is Wrong with Lying?" *Philosophy and Phenomenological Research* 75(3): 535–557, doi:10.1111/j.1933-1592.2007.00092.x.

—. 2013. "Lying and Deceit." in *The International Encyclopedia of Ethics*, edited by Hugh LAFOLLETTE, pp. 3101–3109. Oxford: Wiley-Blackwell, doi:10.1002/9781444367072.wbiee482.

FREUD, Sigmund. 1917. *Vorlesungen zur Einführung in die Psychoanalyse.* Leipzig: Heller.

—. 1922. *Introductory Lectures on Psycho-Analysis*. London: Allen & Unwin. Translation of Freud (1917) by Joan Riviere.

—. 1929. *Introductory Lectures on Psycho-Analysis*. 2nd ed. London: Allen & Unwin. Translation of Freud (1917) by Joan Riviere, first edition: Freud (1922).

GARCÍA-CARPINTERO, Manuel. 2018. "Sneaky Assertions." in *Philosophical Perspectives 32: Philosophy of Language*, edited by John HAWTHORNE and Jason TURNER, pp. 188–218. Hoboken, New Jersey: John Wiley; Sons, Inc., doi:10.1111/phpe.12116.

GOFFMAN, Erving. 1981. "Radio Talk: A Study of the Ways of our Errors." in *Forms of Talk*, pp. 197–330. Philadelphia, Pennsylvania: University of Pennsylvania Press.

GRICE, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, Massachusetts: Harvard University Press.

HARRIS, Daniel W. forthcoming. "Intentionalism and Bald-Faced Lies." *Inquiry*, doi:10.1080/0020174X.2020.1775381.

HAWLEY, Katherine. 2018. "Coercion and Lies." in *Lying: Language, Knowledge, Ethics, and Politics*, edited by Eliot MICHAELSON and Andreas STOKKE, pp. 229–245. Oxford: Oxford University Press, doi:10.1093/oso/9780198743965.001.0001.

JOHNSON, Samuel. 1755a. *A Dictionary of the English Language. Volume I*. London: W. Strahan. Facsimile edition: Johnson (1980).

—. 1755b. *A Dictionary of the English Language. Volume II*. London: W. Strahan. Facsimile edition: Johnson (1980).

—. 1980. *A Dictionary of the English Language*. London: Times Books. Facsimile edition of Johnson (1755a, 1755b).

KEISER, Jessica. 2016. "Bald-Faced Lies: How to Make a Move in a Language Game Without Making a Move in a Conversation." *Philosophical Studies* 173(2): 461–477, doi:10.1007/s11098-015-0502-5.

KENYON, Timothy. 2010. "Assertion and Capitulation." *Pacific Philosophical Quarterly* 91(3): 352–368, doi:10.1111/j.1468-0114.2010.01371.x.

Krstić, Vladimir. 2018. "Knowledge-Lies Re-Examined." *Ratio* 31(3): 312–320, doi:10.1111/rati.12184.

—. 2019. "Can You Lie Without Intending to Deceive?" *Pacific Philosophical Quarterly* 100(2): 642–660, doi:10.1111/papq.12241.

—. 2020. "On the Nature of Indifferent Lies, a Reply to Rutschmann and Wiegmann (2017)." *Philosophical Psychology* 33(5): 757–771, doi:10.1080/09515089.2020.1743255.

Krstić, Vladimir and Wiegmann, Alex. 2024. "Bald-Faced Lies, Blushing, and Noses that Grow: An Experimental Analysis." *Erkenntnis* 89(2): 479–502, doi:10.1007/s10670-022-00541-x.

Lackey, Jennifer. 2013. "Lies and Deception: an Unhappy Divorce." *Analysis* 73(2): 236–248, doi:10.1093/analys/ant006.

—. 2019. "Selfless Assertions." in *The Oxford Handbook of Lying*, edited by Jörg Meibauer, pp. 244–251. Oxford Handbooks. Oxford: Oxford University Press, doi:10.1093/oxfordhb/9780198736578.001.0001.

Leland, Patrick R. 2015. "Rational Responsibility and the Assertoric Character of Bald-Faced Lies." *Analysis* 75(4): 550–554, doi:10.1093/analys/anv080.

LePore, Ernest and Stone, Matthew. 2014. *Imagination and Convention. Distinguishing Grammar and Inference in Language.* Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198717188.001.0001.

MacFarlane, John. 2005. "Making Sense of Relative Truth." *Proceedings of the Aristotelian Society* 105: 305–323, doi:10.1111/j.0066-7373.2004.00116.x.

Mahon, James Edwin. 2015. "The Definition of Lying and Deception." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language; Information. Revision, December 25, 2015, of the version February 21, 2008, https://plato.stanford.edu/entries/lying-definition/.

Maitra, Ishani. 2018. "Lying, Acting, and Asserting." in *Lying: Language, Knowledge, Ethics, and Politics*, edited by Eliot Michaelson and Andreas Stokke, pp. 65–83. Oxford: Oxford University Press, doi:10.1093/oso/9780198743965.001.0001.

Marques, Teresa. 2020. "Disagreement with a Bald-Faced Liar." *Ratio* 33(4): 255–268, doi:10.1111/rati.12268.

Marsili, Neri. 2016. "Lying by Promising." *International Review of Pragmatics* 8(2): 271–313. Special Issue: Empirical Approaches to Lying and Deception, doi:10.1163/18773109-00802005.

—. 2021. "Review of Michaelson and Stokke (2018)." *Utilitas* 33(4): 502–505, doi:10.1017/s0953820821000182.

Maynard Smith, John and Harper, David. 2004. *Animal Signals.* Oxford: Oxford University Press.

Meibauer, Jörg. 2014a. "Bald-Faced Lies as Acts of Verbal Aggression ." *Journal of Language Aggression and Conflict* 2(1–2): 127–150, doi:10.1075/jlac.2.1.05mei .

—. 2014b. *Lying and the Semantics-Pragmatics Interface*. Berlin: Mouton de Gruyter.

—. 2016. "Understanding Bald-Faced Lies. An Experimental Approach." *International Review of Pragmatics* 8(2): 247–270. Special Issue: Empirical Approaches to Lying and Deception, doi:10.1163/18773109-00802004.

MICHAELSON, Eliot and STOKKE, Andreas, eds. 2018. *Lying: Language, Knowledge, Ethics, and Politics*. Oxford: Oxford University Press, doi:10.1093/oso/9780198743965.001.0001.

—. 2021. "Lying, Deception, and Epistemic Advantage." in *The Routledge Handbook of Social and Political Philosophy of Language*, edited by Justin KHOO and Rachel STERKEN, pp. 109–124. Routledge Handbooks in Philosophy. Milton, Abingdon: Routledge.

PAGIN, Peter. 2015. "Problems with Norms of Assertion." *Philosophy and Phenomenological Research* 93(1): 178–207, doi:10.1111/phpr.12209.

PAGIN, Peter and MARSILI, Neri. 2021. "Assertion." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language; Information. Revision, November 17, 2021, of the version of January 22, 2007, https://plato.stanford.edu/entries/assertion/.

PEIRCE, Charles Sanders. 1935. *Collected Papers of Charles Sanders Peirce, vol. 5: Pragmatism and Pragmaticism*. Cambridge, Massachusetts: Harvard University Press. Edited by Charles Hartshorne and Paul Weiss.

RÉCANATI, François. 1982. *Les énoncés performatifs. Contribution à la pragmatique*. Propositions. Paris: Éditions de Minuit.

—. 1987. *Meaning and Force: The Pragmatics of Performative Utterances*. Cambridge: Cambridge University Press. translation of Récanati (1982).

REINS, Louisa M. and WIEGMANN, Alex. 2021. "Is Lying Bound to Commitment? Empirically Investigating Deceptive Presuppositions, Implicatures, and Actions." *Cognitive Science* 45(2), doi:10.1111/cogs.12936.

RUTSCHMANN, Ronja and WIEGMANN, Alex. 2017. "No Need for an Intention to Deceive? Challenging the Traditional Definition of Lying." *Philosophical Psychology* 30(4): 438–457, doi:10.1080/09515089.2016.1277382.

SAUL, Jennifer Mather. 2012. *Lying, Misleading, and What is Said. An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199603688.001.0001.

SIEBEL, Mark. 2020. "The Belief View of Assertion." in *The Oxford Handbook of Assertion*, edited by Sanford C. GOLDBERG, pp. 99–118. Oxford Handbooks. Oxford: Oxford University Press, doi:10.1093/oxfordhb/9780190675233.013.4 .

SNEDDON, Andrew. 2021. "Alternative Motivation and Lies." *Analysis* 81(1): 46–52, doi:10.1093/analys/anaa027.

SØRENSEN, Roy A. 2007. "Bald-Faced Lies! Lying Without The Intent To Deceive." *Pacific Philosophical Quarterly* 88(2): 251–264, doi:10.1111/j.1468-0114.2007.00290.x.

—. 2010. "Knowledge-Lies." *Analysis* 70(4): 608–615, doi:10.1093/analys/anq072.

—. 2022. "Lie For Me: the Intent to Deceive Fails to Scale Up." *Synthese* 200(2), doi:10.1007/s11229-022-03603-3.

STALNAKER, Robert C. 1984. *Inquiry*. Cambridge, Massachusetts: The MIT Press.

—. 1999. *Context and Content: Essays on Intensionality, Speech and Thought*. Oxford: Oxford University Press, doi:10.1093/0198237073.001.0001.

—. 2002. "Common Ground." *Linguistics and Philosophy* 25(4–5): 701–721, doi:10.1023/a:1020867916902.

STOKKE, Andreas. 2013. "Lying and Asserting." *The Journal of Philosophy* 110(1): 33–60, doi:10.5840/jphil2013110144.

—. 2016. "Lying and Misleading in Discourse." *The Philosophical Review* 125(1): 83–134, doi:10.1215/00318108-3321731.

—. 2017. "Metaphors and Martinis: a Response to Keiser (2016)." *Philosophical Studies* 174(4): 853–859, doi:10.1007/s11098-016-0709-0.

—. 2018. *Lying and Insincerity*. Oxford: Oxford University Press, doi:10.1093/oso/9780198825968.001.0001.

TARSITANO, Michael, JACKSON, Robert and KIRCHNER, Wolfgang H. 2000. "Signals and Signal Choices Made by the Araneophagic Jumping Spider Portia Fimbriata while Hunting the Orb-Weaving Spiders Zygiella X-Notata and Zois Geniculatus." *Ethology* 106(7): 595–615, doi:10.1046/j.1439-0310.2000.00570.x .

THOMAS AQUINAS. 1888–1906. *Summa Theologica*. Leonina, IV-XII. Roma: Typographia poliglotta S.C. de Propaganda Fide.

—. 1912–1936. *Summa Theologica*. London: Burns, Oates & Washbourne. Translated by the Fathers of the English Dominican Province, http://dhspriory.org/thomas/summa/.

UNGER, Peter. 1975. *Ignorance: A Case for Scepticism*. Oxford: Oxford University Press, doi:10.1093/0198244177.001.0001.

VIEBAHN, Emanuel. 2019a. "Review of Michaelson and Stokke (2018)." *Notre Dame Philosophical Reviews*, https://ndpr.nd.edu/reviews/lying-language-knowledge-ethics-and-politics/.

—. 2019b. "Lying with Presuppositions." *Noûs* 54(3): 731–751, doi:10.1111/nous.12282.

WIEGMANN, Alex, WILLEMSEN, Pascale and EIBAUER, Jörg. 2022. "Lying, Deceptive Implicatures, and Commitment." *Ergo* 8(50): 709–740, doi:10.3998/ergo.2251.

WILLIAMS, Bernard Arthur Owen. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton, New Jersey: Princeton University Press.

WILLIAMSON, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press, doi:10.1093/019925656X.001.0001.

The Primitivist Response to the Inference Problem

Ashley Coates

While the inference problem is widely thought to be one of the most se-
rious problems facing non-Humean accounts of laws, Jonathan Schaffer
has argued that a primitivist response straightforwardly dissolves the
problem. On this basis, he claims that the inference problem is really a
pseudo-problem. Here I clarify the prospects of a primitivist response
to the inference problem and their implications for the philosophical
significance of the problem. I argue both that it is a substantial question
whether this sort of response ought to be accepted and that the inference
problem, contra Schaffer, remains a significant problem with important
implications for the non-Humean position. I also argue that this dis-
cussion indicates grounds to be wary about applying the Schaffer-style
strategy of straightforwardly dissolving problems by stipulation to other
philosophical problems.

Jonathan Schaffer has argued that taking it to be axiomatic that non-Humean
laws entail regularities "immediately dissolves" (2016b, 580) the well-known
inference problem for the non-Humean conception of the laws of nature. In
a slogan, the non-Humean should simply stipulate that "it is the business
of laws to govern" (2016b, 577). On this basis, Schaffer (2016b, 579) claims
that the inference problem is "no problem whatsoever" and that "whatever
problems the non-Humean about laws might have, the Inference Problem is
not among them."

There are two parts to Schaffer's proposal. The first is the idea that non-
Humeans ought to respond to the inference problem by taking the entailment
between laws and regularities to be primitive. The second is that this prim-
itivist response straightforwardly dissolves the inference problem with no
significant implications or costs for the non-Humean. So, Schaffer thinks
both that the non-Humean should give a primitivist response to the inference
problem and that doing so shows that the problem is really a pseudo-problem.

I argue here that, while there are potential primitivist responses to the inference problem, contra Schaffer, these responses come with substantial metaphysical commitments and argumentative burdens. This argument has important implications for both the prospects of primitivist responses to the inference problem and the philosophical significance of the problem. Regarding the former, it shows that it is a substantial question whether these responses can succeed and whether they are preferable to more traditional explanatory responses. Regarding the latter, it shows that the inference problem remains a serious philosophical problem that has important lessons to teach us about the implications of the non-Humean view.

I also draw some general lessons from this discussion for primitivist responses to philosophical problems in general. Schaffer (2016b, 586–587) thinks that his stipulative dissolution of the inference problem can be generalised to other philosophical problems, such as Bradley's regress. My argument here, though, indicates grounds for thinking that this strategy may often simply ignore the motivation for philosophical problems and, in so doing, obscure the lessons we ought to learn from those problems. Primitivist responses to philosophical problems, then, ought to proceed only with careful attention to the initial motivation for the problems and how that motivation might complicate any such response.

In section 1 and section 2 I outline Schaffer's proposed strategy and argue that it is based on a misinterpretation of the inference problem. In section 3, though, I outline an alternative primitivist strategy that avoids this problem. I then argue, in section 4, that this strategy comes with a substantial argumentative and theoretical burden and so does not simply dissolve the inference problem. In section 5 and section 6 I consider some potential objections to my argument and, in the process, identify a second potential primitivist strategy. Like the first strategy, though, this strategy involves substantial commitments and does not simply dissolve the inference problem. The result is that, while there are potential primitivist strategies for responding to the inference problem, they do not dissolve the problem and it is a substantial question whether they succeed. In section 7, I argue that this result also indicates general grounds to be wary of attempts to dissolve philosophical problems via stipulation.

1 Schaffer's Stipulative Response

Non-Humeans about the laws of nature hold that laws are distinct from regularities but, nonetheless, entail regularities. The inference problem, basically, is the problem of making good sense of this entailment. While it has generally been assumed that an adequate response to this problem would consist in a plausible explanation of the entailment, Schaffer defends a stipulative response to the problem on which the entailment is axiomatized rather than explained. In effect, Schaffer's recommendation is that non-Humeans take it as a primitive fact that laws entail regularities.

Schaffer bases his proposal on the point that any theory is entitled to primitives that do sufficient theoretical work to pay their way. In the case of primitive governing laws, the non-Humean can justify this posit by appealing to the standard motivation for non-Humean laws, on which such laws provide the best explanation of nomic regularities (2016b, sec.4.1). Schaffer, then, thinks that the inference problem is really "no problem whatsoever," because it can be dissolved by simply saying that "it is the business of laws" to govern (2016b, 579). So, Schaffer's proposal is to dissolve the inference problem by taking the entailment between laws and regularities to be primitive rather than attempting to explain it.

It will be helpful, though, to say more about the sense in which taking the entailment to be primitive avoids the need to explain it. Benovsky (2021) has recently proposed that metaphysical primitives are, in general, unexplained specifically in the sense that they are *ungrounded*. A theory's primitives, then, are those entities or facts that the theory takes to be fundamental. This idea fits well with Schaffer's (2016b, 581–583) repeated description of primitive laws as "fundamental posits" or "fundamental laws." It also provides a plausible account of explanatory responses to the inference problem as attempts to specify the grounds for the entailment from laws to regularities. Tooley's (1987) "speculative" response, for example, looks like an attempt to show that the entailment can be grounded in facts about conjunctive universals.

Understanding primitives just in terms of grounding, though, may be too narrow. On some views, there are forms of metaphysical explanation other than grounding. To take a view that I discuss further in section 6, Glazier (2017) thinks that "essentialist explanation" is a form of metaphysical explanation distinct from grounding. Essentialist explanations explain the fact that *a is F* in terms of the fact that *a is essentially F*. For instance, the fact that *Socrates is essentially human* can explain the fact that Socrates is human. If the fact

that Socrates is human is explained in this way, though, then it is surely not a primitive.

To accommodate these sorts of cases, I am going to understand explanation here in terms of the broader notion of metaphysical explanation rather than grounding. Traditional explanatory responses to the inference problem, then, are attempts to metaphysically explain how laws entail regularities. Schaffer's primitivist proposal, on the other hand, is to take the entailment to be a metaphysically unexplained, fundamental posit. This interpretation is, of course, consistent with the idea that axiomatizing the entailment provides a kind of epistemic explanation for the entailment by, for instance, clarifying its place in a theoretical system.

Schaffer's proposal, then, is based on the idea that the following sort of argument drives the inference problem:

(1) If governing laws involve a metaphysically unexplained connection between laws and regularities, then there are no governing laws.
(2) Governing laws involve a metaphysically unexplained connection between laws and regularities.
(3) Therefore, there are no governing laws.[1]

Given this motivation for the inference problem, the problem, as ordinarily understood, consists in showing that (2) is false by metaphysically explaining the relevant connection. Schaffer's argument, though, is based on the fact that (2) only counts against non-Humean laws once (1) is accepted. Given this point, Schaffer thinks that the problem can easily be blocked by taking the connection between laws and regularities to be primitive and, on that basis, rejecting (1).

Schaffer's approach, then, entails that the motivation for the inference problem depends on overlooking the mundane point that theories are, in general, entitled to invoke well-motivated primitives. Indeed, Schaffer makes this claim quite explicitly at various points in the paper. For instance, in discussing Lewis's claim that Armstrong's account of laws founders on the inference problem, Schaffer says that Lewis "has not understood that Armstrong can

---

1 A referee has pointed out that this presentation of the inference problem in terms of an argument is unusual. The problem is generally presented as a problem of *how* governing laws entail regularities rather than as an argument. Nonetheless, discussions of the problem generally take place against the backdrop of an implicit argument that failure to solve the problem provides grounds to reject the non-Humean view. The argument given in the main text is a reconstruction of Schaffer's apparent interpretation of that argument.

and should stipulate that *N* is a relation" (2016b, 580) for which the law to regularity entailment holds. Similarly, in his concluding paragraph, he writes (2016b, 587):

> It is a bad question—albeit one that has tempted excellent philosophers from Bradley through to van Fraassen and Lewis—to ask how a posit can do what its axioms say, for that work is simply the business of the posit. End of story.

At these points, Schaffer explicitly ascribes the motivation for the inference problem to a simple failure to understand that theories are entitled to posit primitives that do specified theoretical work.

At face value, though, it seems implausible that the inference problem could have been so widely taken to be a serious problem just on the basis of such a basic error. The details of Lewis's presentation of the problem do not make this interpretation any more plausible. Lewis's (1983, 366) objection to Armstrong's theory is:

> I find its necessary connections unintelligible. Whatever *N* may be, I cannot see how it could be absolutely impossible to have $N(F, G)$ and Fa without Ga. (Unless *N* just *is* constant conjunction, or constant conjunction plus something else, in which case Armstrong's theory turns into a form of the regularity theory he rejects.)

On Schaffer's interpretation, Lewis's objection to the claim that $N(F, G)$ entails that Fa only if Ga is premised just on an implicit rejection of primitives in general. As Lewis's objection is that the entailment is unintelligible, Schaffer's reading implies that Lewis is actually relying on the implicit claim that primitives are *unintelligible*. On this reading, Lewis accepts (1) just because he implicitly endorses:

(4) Any primitive is unintelligible.

It seems hard to believe that Lewis would be arguing, either explicitly or implicitly, on the basis of anything as implausible as (4). Indeed, as Schaffer (2016b, fn 2) notes, in an earlier section of the same paper, Lewis (1983, 352) himself points out that one way for any theory to accommodate a fact is by taking it to be primitive. Schaffer's reading, then, requires that Lewis, in the same paper, moves from this explicit defence of primitives to the unsupported

assumption that unexplained facts are *unintelligible*. Charity demands that we look for an alternative reading.

2   Understanding the Inference Problem

In expanding on his concern, Lewis (1983, 366) says:

> I am tempted to complain in Humean fashion of alleged necessary connections between distinct existences, especially when first-order states of affairs in the past supposedly join with second-order states of affairs to necessitate first-order states of affairs in the future. That complaint is not clearly right: the sharing of Universals detracts from the distinctness of the necessitating and the necessitated states of affairs. But I am not appeased. I conclude that necessary connections can be unintelligible even when they are supposed to obtain between existences that are not clearly and wholly distinct.

What drives Lewis's reasoning here is not an out-of-hand rejection of primitives, but rather general considerations about the sorts of necessary connections that are intelligible. Lewis's concern is that the Armstrongian law and the first-order state of affairs intuitively are not connected in a way that allows any two entities to stand in a necessitation relation. While Lewis is not clear on whether the putative entailment violates "Hume's dictum,"[2] he clearly thinks that it violates some closely related intuition or principle.

This interpretation of Lewis's objection sheds light on his earlier claim that he can understand the entailment only if N "just is constant conjunction, or constant conjunction plus something else." If $N(F, G)$ were identical with the fact that all Fs are Gs or had this fact as a constituent, then the two facts would plausibly be connected in a way that, in general, allows one fact to entail another. As Lewis points out, though, the facts cannot be connected in this way, because, if they were, Armstrong's theory would collapse into a Humean regularity theory.

Given these points, Lewis's argument can be reconstructed as follows:

---

2 Wilson (2010, 595) gives a standard contemporary statement of Hume's dictum as "there are no metaphysically necessary connections between distinct, intrinsically typed, entities." See Stoljar (2008) and Wilson (2010) for discussion of how to interpret the principle.

(5) An entity, Φ, necessitates an entity, Ψ, only if Φ stands in the sort of connection with Ψ that is necessary for any entity to necessitate another.

(6) Governing laws do not stand in the sort of connection with regularities that is necessary for any entity to necessitate another.[3]

(7) Therefore, governing laws do not necessitate regularities.

The key premise here is clearly (6). While Lewis is less clear than might be hoped about this, the intuition supporting (6) appears to be that, even if governing laws are not fully distinct from regularities, they are too distinct or different for the putative necessary connection between the two to be intelligible.

I think that this interpretation of Lewis is clearly preferable to the interpretation that follows from Schaffer's response to the inference problem. It is clearly the more charitable interpretation, as it avoids ascribing anything as implausible as (4) to Lewis and avoids the inconsistency that (4) would entail in Lewis's own views. Instead of, rather oddly, premising his argument on an unmotivated dismissal of theoretical primitives, Lewis is arguing on the basis of intuitions about the kinds of modal connections that make sense. This interpretation also makes sense of Lewis's appeal to general considerations about which necessary connections are intelligible, while Schaffer's interpretation ignores this part of Lewis's argument.[4]

A similar argument can be made for van Fraassen's (1989) discussion of the inference problem. Far from ignoring the possibility of a stipulative response to the problem, van Fraassen (1989, 97) explicitly argues against such a response. Like Lewis, his argument is based on the point that the regularity cannot be

---

3 Proponents of governing laws, of course, claim that such laws are not only necessary but also sufficient for a necessitation relation to obtain. However, showing that governing laws fail to satisfy some necessary condition for such a relation to obtain would show that they do not suffice for such a relation. I take it that this is the idea underlying Lewis's argument and, as I argue below, other influential discussions of the inference problem. Thanks to a referee for pushing me to clarify this point.

4 My interpretation also fits well with Lewis's well-known, closely related discussion of chance, which Schaffer (2016b, 586–587) gives as another example of an unmotivated rejection of working primitives. Lewis writes, "I don't begin to see [...] how knowledge that two universals stand in a special relation N^* could constrain rational credence about the future coinstantiation of those universals. Unless, of course, you can convince me that this special relation is a chancemaking relation: that the fact that $N^*(J, K)$ makes it so, for instance, that each J has 50% chance of being K." On my interpretation, what underlies Lewis's discussion here is his conviction that the facts that $N^*(J, K)$ and "that each J has 50% chance of being K" are too different to stand in a necessitation relation.

constitutive of the non-Humean law. The problem, then, is how non-Humean laws can entail regularities, given that they are "so distinctly different" (1989, 97) from each other. It is this point that van Fraassen appears to think rules out the stipulative response and motivates the demand—which he ultimately thinks cannot be met—for an explanatory response.

So, like Lewis, van Fraassen's presentation of the inference problem is not premised on an unmotivated rejection of primitives. Instead, also like Lewis, his argument is based on the idea that non-Humean laws and regularities are too distinct or different for the laws to entail the regularities, at least without a compelling explanation of the entailment. I take it, then, that the interpretation developed in this section does a better job than Schaffer's interpretation of capturing van Fraassen's reasoning in addition to Lewis's reasoning.

The interpretation also makes sense of Tooley's early "speculative" response to the inference problem, which puzzles Schaffer (2016b, 585). Schaffer is confused that Tooley (1987) feels the need to go beyond his own stipulative response and propose a speculative response that involves substantial claims about the metaphysics of universals.

To see how the current proposal dispels this confusion, we can begin with Tooley's interpretation of the inference problem. Tooley (1987, 110–111) understands the problem as follows:

> how, exactly, are we to think of the relationship which purport-edly obtains, on the present account of laws, between statements asserting that universals stand in certain nomological relations, and corresponding generalizations about the properties and/or relations of first-order particulars? The relation is to be one of logical entailment. But is it a formal relation, or does one have to postulate de re relations between distinct states of affairs?

The concern here is clearly whether the entailment from laws to regularities requires accepting necessary connections between distinct entities. Indeed, I think no other discussion of the inference problem is so explicitly cast in terms of a concern over Hume's dictum.

Tooley is equally explicit about what his speculative theory might offer to a solution to the inference problem. He says "it may provide an answer to the question […] of whether the present account of laws commits one to holding that there can be logical relations between distinct states of affairs" (1987, 123). He goes on to argue that his speculative theory provides a way to avoid this commitment (1987, 128–129).

Tooley, then, interprets the inference problem as the problem of how laws can entail regularities *without violating Hume's dictum*, and his speculative response is intended to show how this is possible. This interpretation and response are clearly in line with my interpretation of the inference problem, on which the problem is how laws can entail regularities without violating general modal principles. So, Schaffer's confusion at Tooley's proposed response is ultimately driven by Schaffer's failure to note how these general modal considerations motivate the problem.

My interpretation of the inference problem, then, fits better with the original presentations and discussions of the problem than Schaffer's interpretation. While general modal principles play a central role in these discussions, Schaffer's interpretation simply ignores this aspect of the discussions. I have argued that, as a consequence, he gives uncharitable and unconvincing interpretations of Lewis and van Fraassen and fails to make good sense of Tooley's discussion.

My interpretation, on the other hand, makes better sense of each of these discussions by accommodating the central role that general modal considerations play in them. Given this interpretation, the inference problem arises specifically as the need to show that (6) is false by showing that governing laws and regularities *are* connected in the manner required to stand in a necessitation relation. The motivation for the problem, then, comes from the kinds of general modal considerations that drive Lewis, van Fraassen, and Tooley's discussions rather than from an unmotivated rejection of primitives.

3 The Genuine Primitivist Alternative

The interpretation of the inference problem that I just defended entails not only that Schaffer's interpretation of the problem is misguided but also that his response to the problem is misguided. If Schaffer were right that the inference problem is motivated just by a general rejection of theoretical primitives, then he would be right that it could be solved by simply stipulating that non-Humean laws necessitate regularities. Given the interpretation of the problem that I just defended, though, this response begs the question. If the problem is motivated by the concern that non-Humean laws necessitating regularities has unacceptable modal implications, then stipulating that non-Humean laws *do* necessitate regularities simply assumes that those concerns are misguided.

This point can also be put in terms of general considerations about primitives. While any theory is entitled to invoke primitives to do theoretical work,

certain primitives may be independently problematic. While positing a primitive always comes at some theoretical cost, positing such primitives comes at an inflated cost. Indeed, if the posit is sufficiently problematic, it may be unacceptable regardless of the work that it does.

On the interpretation I have defended, though, the inference problem is motivated by the idea that the entailment between non-Humean laws and regularities—at least in the absence of a plausible explanation of the entailment—violates general principles or intuitions about necessary connections. This idea, however, also implies that the entailment is a problematic primitive that would either come at an inflated theoretical cost or actually be untenable. On Lewis's view, the general intuition that counts against the entailment renders it unintelligible, and, so, clearly entails that it is an *unacceptable* primitive.

The motivation for the inference problem, then, is also motivation for thinking that the entailment between non-Humean laws and regularities is an unacceptable or, at least, a problematic primitive. So, responding to the inference problem by simply positing this entailment as a primitive fails to address the problem. One does not show that the entailment is an acceptable primitive by positing it as a primitive.

Against this backdrop, it is also clear how a successful explanatory response to the inference problem would do the necessary work. As I proposed in section 1, the kind of explanation involved in attempted explanatory responses is metaphysical explanation. A successful explanatory response, then, would work by identifying metaphysical grounds—or some other metaphysical explanantia—for the entailment that do not violate the relevant general modal principles. In so doing, such an account would show how non-Humean laws respect the relevant principles in a way that simply stipulating that non-Humean laws entail regularities clearly does not.

This result, though, does not show that no primitivist response to the inference problem could succeed. Instead, it shows that such a response would have to come with an argument that the motivation for the inference problem does not, in fact, show that the entailment is an unacceptable primitive. Specifically, the response would have to be supported by an argument that general modal considerations actually fail to support (6). Given such an argument, rejecting (6) without explaining how laws entail regularities would be legitimate rather than simply begging the question.

So, the genuine primitivist alternative to explanatory responses to the inference problem is to argue that, even without an explanation of how non-

Humean laws necessitate regularities, general considerations about modality do not motivate (6). When made explicit, this line of reasoning might look quite plausible. At least given Lewis's relatively inchoate appeal to intuition, it seems that the non-Humean may quite reasonably deny the motivation for (6). This approach might be bolstered by arguing that the intuition in question is distinctively Humean, and, so, begs the question against the non-Humean.

To the degree that (6) is motivated specifically by Hume's dictum both the idea that the motivation is unconvincing and the idea that it begs the question against the non-Humean may look particularly plausible. Hume's dictum is generally thought to be a distinctively Humean principle and recently significant questions have been raised about whether there are any good grounds to accept the principle (Wilson 2010).

4 The Prospects of the Primitivist Response

On closer inspection, though, I think that widespread intuitions and significant modal principles provide non-circular support for (6). The consequence is not that a primitivist response to the inference problem is *impossible* but rather that such a response comes with a significant theoretical and argumentative burden.

In the first place, even if (6) is motivated by a distinctively Humean intuition, there would still be significant dialectical reasons for non-Humeans to attempt to adequately address it. A non-Humean who could make good sense of laws in a way that is consistent with as many Humean commitments or intuitions as possible would, after all, be in a better dialectical position. So, even if the motivation for the inference problem were, in important respects, based on Humean intuitions, there would still be substantial grounds for non-Humeans to attempt to avoid violating those intuitions.

Perhaps more significantly, though, it is not at all clear that the relevant intuition is distinctively Humean. At least some non-Humeans appear to endorse the concern that an unexplained necessary connection between laws and regularities would be highly problematic. For instance, Armstrong (1997, 226) contrasts his explanatory response to the inference problem with the "profoundly mysterious doctrine" that "[u]niversals, whether instantiated or uninstantiated, stand above the flux and certain relations between the universals 'govern' their instances, lay down the law to their instances." Tugby (2016, 1156), in developing his own explanatory dispositionalist response

to the problem, agrees that the position described here by Armstrong is "a difficult picture to comprehend."

Indeed, a plausible diagnosis for why so many non-Humeans have taken the inference problem to be a pressing problem is that they share this sort of intuition. There is certainly nothing obviously inconsistent about both thinking that there is important potential theoretical work for non-Humean laws, and being concerned that such laws involve a problematic necessary connection between laws and regularities.

Bird (2005; 2007, 91–97), moreover, has developed the inference problem against Armstrong specifically as the problem that $N(F, G)$'s entailing that all Fs are Gs violates Armstrong's *own* general modal commitments. In particular, he argues that the entailment is inconsistent with Armstrong's combinatorial approach to modality and his associated principle of Independence, on which there are no entailments between fully distinct states of affairs. In pursuing this argument, Bird argues, in a similar vein to Lewis and van Fraassen, that the Armstrongian law cannot have the regularity as a constituent.

So, Bird's presentation of the inference problem follows very closely the reconstruction that I gave in section 2. The idea is that, because $N(F, G)$ cannot have *all Fs are Gs* as a constituent, general principles concerning necessary connections mean that $N(F, G)$ cannot entail that all Fs are Gs. In this case, though, the argument is based on Armstrong's own quite precise modal claims, and, so, cannot be dismissed as being based on inchoate or distinctively Humean intuitions.

Furthermore, another precise and influential general modal principle that does not obviously beg the question against the non-Humean conception of laws straightforwardly rules out the primitivist approach. This is the principle that there are, in general, no brute necessary connections between entities.[5] While this principle is again consistent with thinking that there is important theoretical work for non-Humean laws to do, it is clearly inconsistent with taking the necessary connection between laws and regularities to be primitive.

---

5 Van Cleve (2018) provides a useful overview of the many different uses to which this principle has been put in recent metaphysics. Van Cleve's own conclusion is that it is a substantial question whether the principle ought to be accepted and whether the various uses to which it has been put are justified. The key point here, though, is that the principle is widely accepted and that denying it involves taking on a substantial general modal commitment.

So, given this principle, the only way to block (6) is via a plausible account of how laws "do their stuff."[6]

It is also worth noting that Wilsch (2018, 808–809; 2021, 916) has recently pointed to grounds for rejecting brute necessities that may look especially compelling to the non-Humean. He argues that "[d]istribution patterns across possibilities cry out for explanations in the way distribution patterns in the actual world cry out for explanation" (2021, 916). The idea that non-Humean laws are necessary to explain actual distribution patterns, though, is central to the case for non-Humean laws. So, if Wilsch is right, a non-Humean who takes it as primitive that laws necessitate regularities is in serious danger of undermining the original case for non-Humean laws.

The point of this section has been to show that widely accepted intuitions and principles about necessary connections support (6) without in any obvious way begging the question against the non-Humean. This conclusion, of course, does not rule out the possibility of a primitivist response to the inference problem. It remains possible to argue either that taking the entailment between laws and regularities to be primitive does not, in fact, violate significant modal principles or that, all things considered, any such violation is a price worth paying for non-Humean laws.

What the conclusion does indicate, though, is that a primitivist response to the inference problem cannot deliver a Schaffer-style stipulative response that "immediately dissolves" the problem and shows that it is "no problem at all." Instead, as the primitivist response assumes a substantial argumentative and theoretical burden, it leaves the inference problem in place as a significant problem that raises serious difficulties and potentially generates important commitments for the non-Humean.

Furthermore, in demonstrating the commitments and apparent costs that come with the primitivist response, the discussion here indicates that it is a substantial question whether the response can ultimately be made plausible or appealing. Certainly, the burden that attaches to this kind of response means that there remain significant initial grounds for favouring an explanatory response over such a response. So, my conclusion here is not only that the Schaffer-style stipulative dissolution of the inference problem fails but also that it is unclear how successful a more substantive primitivist response might be.

---

6 As Schaffer (2016b, 585) notes, Sider (1992, 262) uses this expression in this context. Schaffer expresses puzzlement at the demand for this sort of account.

5   Non-Humean Laws and General Modal Principles[7]

On the account of the inference problem that I have defended here, general modal considerations, such as Hume's dictum and the ban on brute necessities, are central to the problem. A resulting concern might be that the account simply collapses the inference problem into the distinct problem of whether these general modal principles ought to be accepted. Indeed, Hildebrand (2020, 6–7) has recently implied that the inference problem does disappear into these general modal questions. The idea is that, if these sorts of principles provide the only reason to deny that non-Humean laws primitively entail regularities, then the question really becomes whether we ought to accept these principles.

On my interpretation, though, the inference problem is not simply the problem of whether the relevant general principles ought to be accepted; rather, it is a problem that *presupposes* those principles. So, it is true that *one* possible response to the problem, the primitivist response, is to question the presupposition of those principles. However, a second possible response, the explanatory response, accepts the principles and attempts to show that non-Humean laws need not violate them.

Furthermore, as I argued in the previous section, while the primitivist response is a genuine option for the non-Humean, there is also significant initial motivation for pursuing the explanatory response. So, on my interpretation, one well-motivated response to the inference problem is to show how non-Humean laws can respect the relevant general modal principles. The interpretation, then, does not simply collapse the inference problem into the question of whether those modal principles ought to be accepted.

This point, though, leads to a second potential concern with my interpretation. In his recent survey article, Hildebrand (2020, 2) identifies non-Humean theories just as views that invoke modal primitives in accounting for nomic necessity. If this is right, then it seems that, irrespective of considerations about the inference problem, non-Humean theories will in general violate both Hume's dictum and the ban on brute necessities. The concern, then, is that my interpretation renders the inference problem redundant because non-Humean theories generally involve primitives that violate the relevant modal principles.

As I indicated earlier, though, prominent non-Humeans, like Armstrong and Tooley, endorse Hume's dictum and shun brute necessities. Furthermore, in their responses to the inference problem, both Tooley and Armstrong attempt to produce non-Humean theories that get by without modal primitives. This is especially clear in Tooley's case, as he says that his speculative theory is a view "according to which what laws of nature there are is capable of being unpacked simply in terms of what universals there are, together with part-whole relations between universals" (1987, 123). His goal here is clearly to provide a theory that does *not* involve any modal primitives.

While just how to understand Armstrong's (1997, 224–230) response to the inference problem is less clear, a similar interpretation seems plausible. On Armstrong's view, while N only contingently relates F and G, where $N(F, G)$ is the case it constitutes a structural universal (1997, 227). Armstrong's key idea appears to be that this fact ensures that, when a instantiates F, a also instantiates G. Whether this idea works is, I think, a substantial question, but the key point for now is that it does not appear to invoke modal primitives that violate Hume's dictum or involve brute necessities.

There are also more recent cases of non-Humeans explicitly rejecting brute necessities. In the previous section, I alluded to an argument by Wilsch against brute necessities that seems particularly appealing from an anti-Humean point of view. Wilsch (2021) proceeds to develop an anti-Humean view that eschews brute or fundamental necessities.[8] Kimpton-Nye (2021), in turn, has recently argued that invoking brute necessities fits poorly with dispositionalist views. Partly on this basis, he proposes a dispositionalist or power-theoretic view that grounds modal facts in instances of essentially qualitative properties.[9] As the modal facts in this theory are grounded in qualitative states of affairs, the theory does not appear to involve modal primitives or brute necessities.

There are, then, both prominent and recent cases of non-Humeans explicitly attempting to avoid any commitment to modal primitives that involve brute necessities. So, I do not think it should be assumed from the outset that non-Humeans are committed to these sorts of primitives. Given that the discussion in the previous section indicated significant initial reasons to be wary of such a commitment, this looks like good news for the non-Humean.

---

8 I discuss Wilsch's view further in the next section.

9 This sort of theory has recently received a fair amount of attention. In addition to Kimpton-Nye (2021), Coates (2021), Tugby (2021), and Azzano (2021) all discuss versions of it at length. Earlier discussions of a view along these lines are Jacobs (2011) and Tugby (2012).

Non-Humean theories, then, should not simply be assumed to involve brute necessities or to violate Hume's dictum, nor should the inference problem be thought to collapse into the problem of whether the relevant general modal principles should be accepted. Instead, the inference problem turns on the substantial question of whether non-Humean accounts of the laws of nature can respect these sorts of principles. As I argued in the previous section, this question has bite because the non-Humean would incur a significant argumentative and theoretical burden by rejecting these principles.

6 The Essentialist Primitivist Response

I have thus far interpreted the primitivist response to the inference problem as taking the necessitation between laws and regularities to be brute. Schaffer's understanding of the axioms with which primitive posits are outfitted, though, points to the possibility of an alternative interpretation. Schaffer (2016b, fn 1) interprets these axioms as "meaning postulates and so […] analytic to their terms," but he allows that they may also be thought of as essential truths. So, the entailment between laws and regularities is either analytic to the term "law" or essential to laws.

Both approaches might be thought to provide an explanation of the necessitation between laws and regularities rather than taking it to be brute. Given the analytic conception of axioms, the idea would be that laws entail regularities *because* doing so is part of what it *means* to be a law. Given the essentialist conception, on the other hand, the idea would be that laws entail regularities *because* doing so is part of what it is to *be* a law. In providing these explanations, though, these approaches might be thought to show how the non-Humean can reject (6) without violating general modal principles. The key idea would be that these explanations show how laws *are* related to regularities in the manner required for laws to entail regularities without violating these principles.[10]

While, as I just mentioned, Schaffer does suggest both the analytic and the essentialist interpretations of the axioms he proposes, I do not think that the strategy just outlined can be reasonably attributed to him. That is, I do not think he can reasonably be read as proposing that the axioms, in virtue of being analytic or essential to laws, can explain the entailment of regularities by non-Humean laws. Schaffer nowhere acknowledges the role that I have

---

10   I would like to thank a referee for *dialectica* for raising this possibility.

argued general modal considerations play in motivating the inference problem. Nor does he at any point allude to axioms, conceived either in the analytic or essentialist fashion, as being capable of explaining necessary facts. Instead, in line with my earlier interpretation of his proposal, he focuses on the *general* acceptability of axiomatizing rather than explaining facts, including in cases that do not involve any concerns about brute necessities.[11] Nonetheless, the strategy just outlined provides an alternative primitivist approach that is worth considering.

Given the analytic interpretation of axioms, though, the approach is not promising. In her discussion of Hume's dictum, Wilson (2010, 625) points out that the fact that a sentence is analytic does not answer the metaphysical question of why the entities referred to in the sentence stand in a necessary relation. Using the example of the sentence "necessarily, anything that is scarlet is red," she points out that, while the truth of the sentence

> may be established by attention to its constitutive words or concepts [...] [it remains an open question] what metaphysical facts about the entities at issue in [...] [the sentence] are such that expressions for or concepts applying to these entities incorporate their necessary connection (Wilson 2010, 625–626)

In the case of governing laws, there is no obstacle to defining the term "law" such that laws are distinct from, but entail, regularities. However, doing so provides no *metaphysical* explanation of how the entities, laws, and regularities, are related to each other in such a way that the former necessitates the existence of the latter. So, if axiomatizing the entailment is simply a matter of defining "law" in a certain way, then it does not address the concerns about brute necessities or Hume's dictum.

The proposal in terms of essences, on the other hand, is more promising. On this approach, to axiomatize the governing role of laws is to posit that it is essential to governing laws that they entail regularities. As I noted in section 2, Glazier (2017) has recently argued that, in general, the fact that *a is essentially F* can metaphysically explain the fact that *a is F*. If this is right, though, then the fact that *Law(Φ) essentially entails* Φ can explain the fact that

---

11 In a particularly clear example, Schaffer (2016b, 586) writes "I am saying that everyone needs their fundamental posits, and every posit needs to be outfitted with axioms (or else it is idle). One never needs to do anything further to explain the nature of these inferences beyond saying that they are axiomatic, and one never needs to say anything further about how the posit does its stuff beyond saying that it is the business of the posit to do so."

Law(Φ) entails Φ. Indeed, Wilsch (2021, 917) has recently proposed employing Glazier's argument to make precisely this essentialist move in response to the inference problem.

If one accepts Glazier's general claims about essentialist explanation, this approach clearly avoids the concern that non-Humean laws involve brute necessities. The approach may also avoid violating Hume's dictum, as on one interpretation entities are not distinct in the sense relevant to the dictum if they are essentially connected (Stoljar 2008). So, this essentialist approach appears to provide a potential primitivist approach that, unlike the modal primitivist approach discussed in the previous sections, blocks (6) by showing that non-Humean laws are consistent with the relevant general modal principles.

The essentialist approach, though, comes with a significant commitment not only to essentialism but to a particularly robust essentialism, on which objects have non-modal essences that metaphysically explain their essential properties. Indeed, for just this reason, this approach is not one that could be endorsed by Schaffer who is a skeptic about essence (2016a, 83).

The essentialist view also appears to raise difficulties for the idea that laws ground regularities (Emery 2019). The problem is that, if regularities are essential to laws, then laws ontologically depend on regularities. However, if laws ground regularities, then regularities also ontologically depend on laws. The apparent result would be an objectionable circularity in relations of ontological dependence.[12]

Perhaps more significantly, though, the approach appears to be inconsistent with the Dretske-Tooley-Armstrong view (DTA). The obvious way to extend the approach to this view is to claim that it is essential to N that $N(F, G)$ entails that all Fs are Gs. However, that N essentially stands in this non-trivial modal relation with distinct universals is inconsistent with the categoricalism about properties that is central to DTA.

I think, then, that the essentialist approach does provide a possible primitivist response to the inference problem. However, like the primitivist modal response, it comes with substantial metaphysical commitments and looks to be inconsistent with both a grounding conception of governing laws and DTA. So, this primitivist response also incurs a substantial burden and cannot deliver on Schaffer's straightforward dissolution of the inference problem.

---

12 Jaag (2014, 18) and Kimpton-Nye (2021, 3432) both raise closely related difficulties about dispositional-essentialist accounts of laws of nature.

7 Philosophical Problems and the Stipulative Strategy

My primary goal here has been to clarify the prospects and implications of a primitivist response to the inference problem. I identified two genuine primitivist options that involve taking it as primitive, respectively, that laws entail regularities and that laws *essentially* entail regularities. I argued that, while both approaches represent open possibilities for at least some non-Humeans, they both come with significant commitments and complexities. In so doing, I hope to have cleared the way for further consideration of whether either of these approaches ought ultimately to be accepted, and, if so, what the implications are for non-Humean theories of laws.

Whatever the ultimate verdict on these primitivist responses, though, I have argued that they do not deliver on Schaffer's idea that a primitivist response to the inference problem can straightforwardly dissolve the problem. Instead, these primitivist approaches leave the inference problem in place as a significant philosophical problem that has important implications for the prospects and commitments of non-Humean theories of laws.

I now want to clarify the implications of this result for Schaffer's attempt to generalise his stipulative strategy beyond the inference problem. Schaffer claims, for instance, that the strategy can be applied to Bradley's regress (2016b, sec.3.2), to the connection between chance and rational credence and to issues in the metaphysics of grounding (2016b, sec.5). In the case of Bradley's regress, he argues that the right response to the question of how relations relate is just to stipulate that "it is the business of relations to relate" (2016b, 586). My discussion thus far, though, indicates general grounds for being wary of this strategy.

To see why, it is useful to see how my account of the inference problem fits with Robert Nozick's account of the form of many central philosophical problems. According to Nozick (1981, 9), these problems have the form "how is one thing possible, given (or supposing) certain other things." Nozick refers to the "other things" here as "apparent excluders", as they are things that apparently exclude the possibility in question.

For instance, on Nozick's (1981, 8) interpretation, the problem of free will has the form:

> How is it possible for us to have free will, supposing that all actions are causally determined?

The problem is how we can have free will, *given that causal determinism appears to exclude free will*. Similarly, on my account, the inference problem has the form:

> How can governing laws necessitate regularities, *given that they do not appear to stand in the sort of general connection that is required for one thing to necessitate another*?

The problem is how laws can necessitate regularities, *given that general modal considerations appear to exclude this kind of necessitation*. As I have already argued, the response "laws *do* necessitate regularities" is no solution to this problem. Instead, an adequate response to the problem needs to show how the apparent excluder does not rule out the relevant necessitation. One could argue that the problem is entirely misguided by arguing that there is no reason to accept the excluder. However, simply ignoring the apparent excluder and stipulating that laws necessitate regularities is no response to the problem.

This result can be generalised. For any problem with the form:

> How is it possible that *p*, *given q*?

the simple stipulative response "*p*" is clearly unacceptable. An adequate response needs to acknowledge *q* as the motivation for the problem and attempt to show how *q* does not rule out *p*. This could be done by arguing against *q* or by arguing that *p* and *q* are, in fact, consistent. As I demonstrated in my discussion of the inference problem, these sorts of strategies are consistent with taking *p* to be primitive. However, simply ignoring *q* and stipulating *p* begs the question against the motivation for the problem rather than addressing it.

The general lesson here is that, prior to applying the Schaffer-style simple stipulative strategy to a philosophical problem, one ought to consider whether the problem is driven by apparent excluders that render that strategy misguided. In the case of Bradley's regress, for instance, one might think that the core problem is:

> How is it that *R* relates *a* to *b*, *given that it is possible for R, a and b to exist without R relating a to b*?[13]

---

13 See Maurin (2011) for this kind of interpretation of the regress.

On this interpretation, it is the apparent excluder that motivates the regress by indicating that *something more* than *R*, *a* and *b* is needed for *R* to relate *a* to *b*. Given this construal of the problem, an adequate response needs to give an account of what makes the difference between, on the one hand, *R*, *a* and *b* existing independently and, on the other hand, *aRb*. Simply stipulating that *R does* relate *a* to *b* does not solve this problem.

Nor does simply saying that "it is the business of relations to relate" clearly address the problem. Indeed, in this context it is not immediately clear what this claim would mean. It cannot mean that *R* necessarily, or essentially, relates *a* to *b*, because *R* might exist without relating *a* to *b*. It may mean that *R* cannot exist without relating *some* entities but, of course, this fact does not explain what distinguishes *aRb* from the independent existence of *R*, *a* and *b*.

As Maurin (2011) has indicated, though, this line of reasoning appears to rely on the assumption that relations are universals rather than tropes. If relations are tropes, then it may be possible to hold that *R* essentially relates *a* and *b*, and, so, that *R* exists only if *aRb*. So, *if one accepts the significant metaphysical claim that relations are tropes*, then it may be possible to respond to Bradley's regress by stipulating that relations essentially relate certain particulars.

Indeed, this point seems implicit in Schaffer's own discussion. He writes, "What it is to be a relation between *a* and *b* is to relate *a* to *b*" (2016b, 582). Here Schaffer appears to be implicitly treating relations precisely as tropes that are individuated by relating particular objects rather than others. If this is right, then Schaffer's proposed primitivist response to Bradley's regress smuggles in a highly significant ontological commitment, and, consequently, fails to deliver the advertised innocent, straightforward dissolution of the problem.

The upshot is that the situation regarding a primitivist response to Bradley's regress looks very similar to the situation regarding a primitivist response to the inference problem. In both cases, clarifying the excluders that motivate the problem indicates that the straightforward stipulative dissolution of the problem fails to engage with the motivation for the problem. Furthermore, clarifying this motivation indicates that the genuine primitivist options in responding to the problems come with substantial metaphysical commitments. This result means both that it is an open question whether these primitivist approaches ought to be accepted and that they constitute substantive metaphysical proposals rather than straightforward dissolutions of problems.

The general lesson is that any application of the simple, Schaffer-style stipulative response to a philosophical problem ought to be preceded by careful consideration of whether the problem at hand has the form identified by Nozick. Where problems do have this form, the Schaffer-style response simply ignores the motivation for the problem. In these cases, a primitivist response to the problem ought, instead, to involve a substantive argument that the excluders do not, in fact, rule out the primitivist approach. It is then an open question, to be addressed in each case, just how successful this argument is and which commitments come with it.*

Ashley Coates

0000-0001-8557-3163
University of the Witwatersrand
Johannesburg
ashley.coates@wits.ac.za

References

Armstrong, David M. 1997. *A World of States of Affairs*. Cambridge: Cambridge University Press.

Azzano, Lorenzo. 2021. "Dispositionality, Categoricity, and Where to Find Them." *Synthese* 199(1-2): 2949–2976, doi:10.1007/s11229-020-02917-4.

Barker, Stephen J. and Smart, Benjamin T. H. 2012. "The Ultimate Argument Against Dispositional Monist Account of Laws." *Analysis* 72(4): 714–722, doi:10.1093/analys/ans114.

Benovsky, Jiří. 2021. "Primitives." in *The Routledge Handbook of Metametaphysics*, edited by Ricki Leigh Bliss and J. T. M. Miller, pp. 387–394. Routledge Handbooks in Philosophy. Milton, Abingdon: Routledge.

Bird, Alexander. 2005. "The Ultimate Argument Against Armstrong's Contingent Necessitation View of Laws." *Analysis* 65(2): 147–155, doi:10.1111/j.1467-8284.2005.00542.x.

—. 2007. *Nature's Metaphysics. Laws and Properties*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199227013.001.0001.

Coates, Ashley. 2021. "Making Sense of Powerful Qualities." *Synthese* 198(9): 8347–8363, doi:10.1007/s11229-020-02576-5.

Emery, Nina. 2019. "Laws and their Instances." *Philosophical Studies* 176(6): 1535–1561, doi:10.1007/s11098-018-1077-8.

GLAZIER, Martin. 2017. "Essentialist Explanation." *Philosophical Studies* 174(11): 2871–2889, doi:10.1007/s11098-016-0815-z.

HILDEBRAND, Tyler. 2020. "Non-Humean Theories of Natural Necessity." *Philosophy Compass* 15(5), doi:10.1111/phc3.12662.

JAAG, Siegfried. 2014. "Dispositional Essentialism and the Grounding of Natural Modality." *Philosophers' Imprint* 14(34).

JACOBS, Jonathan D. 2011. "Powerful Qualities, not Pure Powers." *The Monist* 94(1): 81–102, doi:10.5840/monist20119415.

KIMPTON-NYE, Samuel. 2021. "Reconsidering the Dispositional Essentialist Canon." *Philosophical Studies* 178(10): 3421–3441, doi:10.1007/s11098-021-01607-2.

LEWIS, David. 1983. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61(4): 343–377. Reprinted in Lewis (1999, 8–55), doi:10.1080/00048408312341131.

—. 1999. *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press, doi:10.1017/CBO9780511625343.

MAURIN, Anna-Sofia. 2011. "An Argument for the Existence of Tropes." *Erkenntnis* 74(1): 69–79, doi:10.1007/s10670-010-9252-0.

NOZICK, Robert. 1981. *Philosophical Explanations*. Cambridge, Massachusetts: Harvard University Press.

SCHAFFER, Jonathan. 2016a. "Grounding in the Image of Causation." *Philosophical Studies* 173(1): 49–100, doi:10.1007/s11098-014-0438-1.

—. 2016b. "It is the Business of Laws to Govern." *Dialectica* 70(4): 577–588, doi:10.1111/1746-8361.12165.

SIDER, Theodore. 1992. "Tooley's Solution to the Inference Problem." *Philosophical Studies* 67(3): 261–275, doi:10.1007/bf00354539.

STOLJAR, Daniel. 2008. "Distinctions in Distinction." in *Being Reduced. New Essays on Reduction, Explanation, and Causation*, edited by Jakob HOHWY and Jesper KALLESTRUP, pp. 263–279. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199211531.001.0001.

TOOLEY, Michael. 1987. *Causation: A Realist Approach*. Oxford: Oxford University Press.

TUGBY, Matthew. 2012. "Rescuing Dispositionalism from the Ultimate Problem: Reply to Barker and Smart (2012)." *Analysis* 72(4): 723–731, doi:10.1093/analys/ans112.

—. 2016. "Universals, Laws, and Governance." *Philosophical Studies* 173(5): 1147–1163, doi:10.1007/s11098-015-0521-2.

—. 2021. "Grounding Theories of Powers." *Synthese* 198(12): 11187–11216, doi:10.1007/s11229-020-02781-2.

VAN CLEVE, James. 2018. "Brute Necessity." *Philosophy Compass* 13(9), doi:10.1111/phc3.12516.

VAN FRAASSEN, Bas C. 1989. *Laws and Symmetry*. Oxford: Oxford University Press, doi:10.1093/0198248601.001.0001.

Wilsch, Tobias. 2018. "Genuine Violations of Laws." *Australasian Journal of Philosophy* 96(4): 806–821, doi:10.1080/00048402.2017.1411963.

—. 2021. "The Governance of Laws of Nature: Guidance and Production." *Philosophical Studies* 178(3): 909–933, doi:10.1007/s11098-020-01465-4.

Wilson, Jessica M. 2010. "What is Hume's Dictum, and Why Believe It?" *Philosophy and Phenomenological Research* 80(3): 595–637, doi:10.1111/j.1933-1592.2010.00342.x.

How to Adopt a Logic

DANIEL COHNITZ & CARLO NICOLAI

What is commonly referred to as the *Adoption Problem* is a challenge to the idea that the principles of logic can be rationally revised. The argument is based on a reconstruction of unpublished work by Saul Kripke. As the reconstruction has it, Kripke extends the scope of Willard van Orman Quine's regress argument against conventionalism to the possibility of adopting new logical principles. In this paper we want to discuss the scope of this challenge. Are all revisions of logic subject to the Adoption Problem? If not, are there significant cases of logical revision that are subject to the Adoption Problem? We will argue that both questions should be answered negatively.

What is commonly referred to as the *Adoption Problem*[1] is often considered a challenge to the idea that the principles for logic can be rationally revised. The argument is based on a reconstruction of unpublished work by Saul Kripke.[2] As the reconstruction has it, Kripke essentially extends the scope of Willard Van Orman Quine's (1936) regress argument against conventionalism to the possibility of adopting new logical principles or rules. According to the reconstruction, the Adoption Problem is that new logical rules cannot be adopted unless one already can infer with these rules, in which case the adoption of the rules is unnecessary (Padro 2015, 18).

In this paper we want to discuss the scope of this challenge. Are all revisions of logic subject to the Adoption Problem? If not, are there significant cases of logical revision that are subject to the Adoption Problem? We will argue that both questions should be answered negatively. Kripke's regress does not arise for all rules of inference and not even for the adoption of those rules that are of relevance for the discussion of the rational revisability of logic.

---

1 This label for the problem is due to Padro (2015).

2 See Stairs (2006), Padro (2015), Finn (2019), Finn (2021), Kripke (2024) and Devitt and Roberts (2024). Since the basis of this discussion is an unpublished manuscript that is not authorized, we decided to refer to it in the following way: when providing textual evidence for Kripke's views, we quote from papers that are published and directly quote Kripke; in cases in which we want to give credit to Kripke for an observation or argument, we refer to Kripke (2024).

We will begin the paper in section 1 with a brief summary of the use that Quine made of the regress argument against a conventionalist conception of logic and sketch Quine's own view on the revisability of logic. Kripke seems to claim that the point that Quine makes against conventionalism should equally apply to Quine's own view on the rational revisability of logic. In section 2 we will look at which logical principles are at all subject to a potential regress or circularity problem and we will discuss whether the principles that are potentially subject to such problem are principles that are of relevance for the discussion of the rational revisability of logic.[3] Our arguments in section 2 will thereby follow the specific setup that Kripke introduced for the discussion of the Adoption Problem. In section 3 we will investigate actual cases of proposed logical revisions in order to show how the more abstract considerations of the previous sections may apply to "real life" examples.

Since we arrive at a negative answer to the two questions above, we will close the paper in section 4 by considering alternative targets for for the Adoption Problem. Perhaps it doesn't primarily target Quine's view on the revisability of logic but some other aspect of Quine's view on logic. However, as we will argue in that , also for these alternative targets Kripke's argument doesn't pose a real challenge.

The main claims of the paper are then that there is no adoption problem that would compromise rational revision of our logic, provided that we already possess some basic reasoning skills. This is the case both for the thought experiment considered by Padro and Kripke, and for more realistic scenarios of logical revision. Moreover, that basic reasoning skills are unadoptable is consistent with a Quinean philosophy of logic.

1   The Adoption Problem

According to Padro (2015), Kripke uses the following example to illustrate the problem of adoption:

> Let's try to think of someone—and let's forget any questions about whether he can really understand the concept of "all" and so on—who somehow just doesn't see that from a universal statement each instance follows. But he is quite willing to accept my author-

---

3 Since our discussion can't cover all possible revision to logic that one can come up with, we will limit our discussion to logics that are plausible alternatives to classical logic. We will motivate this choice in due course.

ity on these issues—at least, to try out or adopt or use provisionally any hypotheses that I give him. So I say to him, "Consider the hypothesis that from each universal statement, each instance follows." Now, previously to being told this, he believed it when I said that all ravens are black because I told him that too. But he was unable to infer that this raven, which is locked in a dark room, and he can't see it, is therefore black. And in fact, he doesn't see that that follows, or he doesn't see that that is actually true. So I say to him, "Oh, you don't see that? Well, let me tell you, from every universal statement each instance follows." He will say, "Okay, yes. I believe you." Now I say to him, "'All ravens are black' is a universal statement, and 'This raven is black' is an instance. Yes?" "Yes," he agrees. So I say, "Since all universal statements imply their instances, this particular universal statement, that all ravens are black, implies this particular instance." He responds: "Well, Hmm, I'm not entirely sure. I don't really think that I've got to accept that." (Padro 2015, fn.49)

1.1 *Quine against Conventionalism*

Lewis Carroll's (1895) similar dialogue between a tortoise and Achilles has famously been used by Quine (1936) in order to show that the logical positivists' conventionalism about logic is in trouble.[4] Conventionalism about logic (of the kind that Quine considers) explains why logic should have a special status: Logical principles are knowable *a priori* and necessarily true. According to conventionalism, we decide to maintain the statements of logic "independently of our observations of the world" and thus assign them a truth-value by convention. This accounts for their epistemic and modal status.

Although Quine expresses considerable sympathy for the view (granting that it is "perhaps neither empty nor uninteresting nor false"), he nevertheless sees it facing a difficulty that he summarizes as follows:

> Each of these conventions [Quine refers here to the schematic axioms of propositional logic] is general, announcing the truth of every one of an infinity of statements conforming to a certain

---

4 Who the target of Quine's paper "Truth by convention" eventually is, is not clear. Quine doesn't explicitly say that it is Carnap and there are reasons to think he targeted his own view (Ebbs 2011) and that of C.I. Lewis (Morris 2018).

description; derivation of the truth of any specific statement from
the general convention thus requires a logical inference, and this
involves us in an infinite regress. (1936, 103)

In Carroll's dialogue, the tortoise challenges Achilles to get it to infer in
accordance with Modus Ponens. Achilles fails to achieve this even though
the tortoise is ready to accept an explicit statement of Modus Ponens as a true
principle. For Quine, the upshot of that dialogue is that logic can't be based on
convention alone, since it seems that we need to have the ability to apply the
supposed conventions and derive consequences from them in order to follow
them. But then logic must be prior to such conventions (rather than the other
way around): "In a word, the difficulty is that if logic is to proceed *mediately*
from conventions, logic is needed for inferring logic from the conventions"
(1936, 104).

    Quine does see a way for the conventionalist to address this difficulty. What
if we can adopt a convention "through behaviour" (1936, 105) instead of
adopting it via explicitly announcing it first? Perhaps the explicit formulation
of these conventions can come later, once we have language and logic and
all that at our disposal. For Quine this is a live option, but not one that he is
still willing to describe as logic being based on "convention." From Quine's
behaviorist point of view, behavior that follows a conventional rule is indis-
tinguishable from behavior that displays firmly held beliefs.[5] Since the label
"convention" is then without explanatory power, we can drop it from our
account of logic.[6]

1.2 *Kripke against Quine*

As Padro (2015) explains, Kripke now turns the regress[7] argument against
Quine himself. Quine had famously suggested in "Two dogmas of empiri-

---

5 In fact, Quine only makes the much weaker observation that it would be "difficult to distinguish"
a behavioral adoption of conventions from behavior that displays firmly held beliefs.

6 See Azzouni (2014) for a discussion of conventionalism and Quinean arguments against it. Thanks
to the work of David Lewis and others we now have a much clearer idea of how behavior that is
based on firmly held belief can be distinguished from behavior that is guided by an implicitly
adopted convention.

7 In Carroll's original argument, the structure of the problem is a regress: the tortoise requires
always new meta-principles in order to apply Modus Ponens at a given level. The regress is
provoked, because the very rule that is supposed to be adopted is the rule that is necessary to
apply that new rule. In that sense, the regress obtains because of that *circularity*. In what follows
we will sometimes refer to that argument/problem as a regress or a circularity argument/problem.

cism" (1951) that not even logic is immune to revision. Empirico-pragmatic considerations may lead us to the adoption of a new logic. A view that is, of course, quite compatible with the idea that logic is nothing but firmly held belief in the first place. Perhaps—so Quine's own example—we may decide to adopt a logic that drops the principle of excluded middle because it may help to simplify quantum mechanics (1951). However, Kripke seems to believe that Quine's picture, viz. that we can treat principles of logic just like any other empirical hypothesis, is prone to the exact same objection that Quine mounted against conventionalism. Padro cites Kripke as follows:

> [...] the Carnapian tradition about logic maintained that one can adopt any kind of laws for the logical connectives that one pleases. This is a principle of tolerance, only some kind of scientific utility should make you prefer one to the other, but one is completely free to choose. Of course, a choice of a different logic is a choice of a different language form.
>
> Now, here we already have the notion of adopting a logic, which is what I directed my remarks against last time. As I said, I don't think you can adopt a logic. Quine also criticizes this point of view and for the very same reason I did. He said, as against Carnap and this kind of view, that one can't adopt a logic because if one tries and sets up the conventions for how one is going to operate, one needs already to use logic to deduce any consequences from the conventions, even to understand what these alleged conventions mean.
>
> This is all very familiar as a criticism of Carnap. Somehow people haven't realized how deep this kind of issue cuts. It seems to me, as I said last time, obviously to go just as strongly against Quine's own statements that logical laws are just hypotheses within the system which we accept just like any other laws, because then, too, how is one going to deduce anything from them? I cannot for the life of me, see how he criticizes this earlier view and then presents an alternative which seems to me to be subject to exactly the same difficulty. (2015, 113)

Stairs (2006) and Devitt and Roberts (2024) interpret the adoption problem as targeting in particular Quine's idea that logic is revisable and that we can adopt

a *new* logic.[8] Padro ([2015](#)) seems to see the adoption problem as a problem for adopting a logic in the first place and Kripke ([2024](#)) is vague about the target of the argument. Kripke discusses the adoption problem in a paper on Putnam's views on the possibility of revising logic for empirical reasons and clearly seems to think that the adoption problem should have some relevance for the revisability of logic. His main target is the use that Putman ([1969](#))—and others who follow Quine's views on the revisability of logic—makes of the phrase that we can "adopt" a new logic ([Stairs 2006](#), [2016](#)). Thus, we take it that discussing the adoption problem as a problem in the context of logic's revisability gets at what is ultimately at stake in Kripke's original argument. (However, as we will discuss in section [4](#), Kripke also makes some remarks in his paper that suggest that he, too, may have the adoption of a first logic in mind.)

We will begin this paper by considering the Adoption Problem as a problem for Quine's idea that logic is revisable but will discuss in section [4](#) whether that is the best interpretation of Kripke's attack on Quine. We hope that this brings some clarity into what the adoption problem possibly is and for which view this might be a problem. The idea that logic is rationally revisable is broader than the idea that a first logic can be adopted via the acceptance of logical principles, thus the former seems to be the natural starting point for our analysis.

According to this reconstruction of the argument, logic is not only not based on convention, but logic can't be rationally revised either, because whatever empirico-pragmatic reasons we may have for preferring some alternative logic, we can't adopt a new logic. Presumably the argument is then that the adoption of a new logical principle (as in Kripke's example) would already presuppose the logical competence that allows us to apply such principle. However, as in Kripke's example, if that competence is in fact the very rule we are supposed to adopt, then this can't work.

A *prima facie* reasonable reaction to the argument so understood—due to Devitt and Roberts ([2024](#)), for instance—is to distinguish the way in which we come to know the propositional form of a logical principle, its representation, such as "from a universal statement, each instance follows," and the way in which an agent can come to be *governed* by such logical principle, a state

---

8 Finn ([2019](#)) interprets Kripke to pose a problem for "anti-exceptionalism" about logic, but leaves it vague what aspect of anti-exceptionalism is the target. Revisability is, however, a central aspect of the anti-exceptionalist doctrine and clearly a potential target if there was a problem with *adopting* new rules.

that may not necessarily require a representational form. The first kind of knowledge may be dubbed *declarative*, the second *procedural*. According to this first reaction, therefore, the sort of revision involved in Carroll's example concerns the fact that declarative knowledge of a rule alone may not be sufficient to rationally revise one's logical beliefs. But this does not rule out the possibility of training someone in acquiring procedural knowledge of a new logical principle.

A similar position is assumed by Priest (2014), although framed in his distinction between the *logica docens*, *utens*, and *ens*. The logic we teach (*docens*) can be revised by means of a broadly abductive methodology. What is commonly called a "logic," for Priest, should in fact better be seen as a "logical theory," namely a substantial body of knowledge concerning some notion of logical consequence. Now, a logical theory can be rationally revised in the same way as other scientific theories can be revised, namely by comparing it with alternatives according to theory-choice criteria such as explanatory power, strength, adequacy to data, unifying power, and whatever else these may be. The logical theory we teach, therefore, can be rationally revised, and so can the logical theory we *use*. How? Simply by training oneself in a chosen *logica docens*. To connect Priest's approach to rational revisability of logic with the Carroll-Kripke example, what seems to be clear is that for Priest the process of acquisition of a rule is not a local procedure, but rather a global process of acceptance of a logical theory that goes well beyond the rules of a formal system. This point will be further expanded in section 3.

In the next three sections we leave aside these attempts to undermine the Adoption Problem that deny a significant role to the declarative knowledge of a rule or principle. We will work under the assumption that the declarative knowledge of a logical principle does indeed play a role in one's actual adoption, and consider in more detail how such process could actually work. This is indeed how Padro (2015, 31) understands "adoption": we adopt a way of inferring (for example, in accordance with Modus Ponens), if we pick it up on the basis of the acceptance of the corresponding logical principle alone.

As it will turn out in sections 2 and 3, there is no problem of adoption that would arise for the *revision* of logic (as Kripke seems to claim). It is true that one needs *some* basic reasoning skills in order to be able to adopt and apply new ones, but in pretty much all cases in which one has already a logic, these will be available.

1.3 *Logica Utens*

Although we will set aside Priest's solution to the problem of adoption, it will still be useful for our discussion to help ourselves to a distinction between *logica docens* and *logica utens*. The former is an explicit theory that may or may not be formalized in precise mathematical terms.

A *logica utens*, on the other hand, is—in our terminology—the logic that we reason with under suitably idealized circumstances. What matters is that the *logica utens* is not just a description of all of our actual inferences (including all inferences we would ourselves accept to be mistakes) but rather a reconstruction of the rules we recognize as normatively governing correct reasoning. While Aristotle is widely credited with having started the business of developing a logica docens, *homo sapiens* much earlier started to develop a *logica utens*.

Logica utens will play an important role in our analysis of the Adoption Problem. We will argue that Kripke's thought experiment is best understood as the attempt to revise one's *logica utens*, and we will pinpoint precisely when this task is bound to fail, and when it is instead unproblematic. Even the more general context of revision of one's logical theory can be thought of as an attempt to revise one's *logica utens*: in those cases revision of *logica utens* amounts to a revision of one's logical metatheory, and we will investigate whether this is a feasible task also in that context.

2 Patterns of Adoption

2.1 *What Can We Adopt?*

As noticed already in Cohnitz and Estrada-González (2019), when one looks carefully at the Carroll-Kripke example, it becomes clear that not all principles are equally problematic. To see this, let us frame our discussion in a logical formalism in which one has finitely many rules for introduction and elimination for a finite set of logical connectives (natural deduction or sequent calculi are both adequate options). Consider the following version of our original dialogue in which universal instantiation is now replaced by the introduction of the existential quantifier. It involves subjects A and B and we assume, for the sake of the argument, that B is not able to perform inferences according to *Existential Introduction*. As before, we assume that B is willing to cooperate in accepting and reasoning according to the hypotheses that A provides.

A: Consider the hypothesis that, if some predicate φ holds of t, then there is something that satisfies φ.

B: BOK, I am considering it.

A: This piece of paper is white, isn't it?

B: Yes.

A: Therefore, since if some predicate φ holds of an individual t, then there is something that satisfies φ, it follows that *there is something that is white*.

B: Sure, thanks!

In the above dialogue, unlike what happens in the Kripke case, nothing prevents B from following and accepting A's instructions. The reason is that *no prior* understanding of Existential Introduction is needed for B to follow the instructions given by A.

However, there is something else that *needs* to be presupposed by B. First of all they need the ability of inferring via Modus Ponens, as we learnt from Carroll's example. To be clear, we employ the label "Modus Ponens" for a rule of inference akin to the standard natural deduction rule, *or* the cut rule in a sequent calculus. A choice between one or the other may depend, for instance, on whether we conceive of the "if…, then…" in A's hypothesis as an entailment sign—in which case one needs cut—, or as an object linguistic conditional—in which case on needs a rule for the elimination of such a conditional. Of course we are not fixing a specific system in our discussion, and therefore these are at best structural analogies. We will come back to this point below.

In the light of Kripke's example, it would prima facie seem that also Universal Instantiation is required. However, both in Kripke's example and here we need much less than Universal Instantiation in full generality. Consider A's last sentence: it presupposes the capability of recognizing the validity of the step that goes from an argument of the form $\varphi(t/v) \therefore \exists v \varphi$, *for all φ*, to an argument of the form $P(t/v) \therefore \exists v P$ for a particular P. Similarly, in Kripke's example, the step that prevents the receiver of the instructions from agreeing on the desired conclusion is her incapability of recognizing the validity of the inference from an argument of the form $\forall v \varphi \therefore \varphi(t/v)$ to one of the form $\forall v P \therefore P(t/v)$. In both cases, it is a form of universal instantiation that is at stake. But at a closer look, the inferences under considerations are in fact of the form:

SCS. For any *formula* φ, if $\Phi(\varphi)$, then $\Phi(P/\varphi)$, for some fixed argument pattern Φ.

SCS is a very distinguished form of Universal Instantiation. First, quantifiers range over a fixed *set* of formulae of the language under consideration. Under the natural assumption that the languages we speak are countable, the size of such set is then countable too, whereas no such assumption is required for the general form of Universal Instantiation. Moreover, SCS has a form that is well-known to logicians: it is a *schematic substitution principle*—whence the label SCS—, according to which, by accepting the schema, one accepts all its specific instances in the language under consideration.

This discussion can be generalized by formulating a more abstract recipe for adoption in the box below.

RECIPE FOR ADOPTION.

1. One starts with a schematic logical principle $\Phi(X_1, \ldots, X_n; z_1, \ldots, z_m)$ of the form

   (1) if $\Phi_1(\vec{X}; \vec{z})$ and ... and $\Phi_k(\vec{X}; \vec{z})$, then $\Psi(\vec{X}; \vec{z})$,

   with $\vec{X}$ and $\vec{z}$ possibly empty strings of variables of finite length. Here the X_i's are one sort of variables to be replaced with formulae, and the z_j's are meta-variables for terms possibly including a different sort of variables for objects. Some machinery for renaming variables is also assumed.

2. One is then given a schematic instance of the antecedent of the conditional

   $\Phi_1(\vec{A}; \vec{t})$ and ... and $\Phi_k(\vec{A}; \vec{t})$

   for $\vec{A}$ formulae of the language and $\vec{t}$ actual terms in the language.[MISSING FOOTNOTE]

3. SCS enables one to go from (1) to

   if $\Phi_1(\vec{A}; \vec{t})$ and ... and $\Phi_k(\vec{A}; \vec{t})$, then $\Psi(\vec{A}; \vec{t})$,

4. by Modus Ponens applied to (2) and (3), one concludes $\Psi(\vec{A}; \vec{t})$, thereby inferring according to (1).

A few comments to the RECIPE FOR ADOPTION are in order. First, we are analysing Kripke's pattern for adoption. As such, the intended application of our pattern is the scenario envisaged by Kripke: we are not putting forward a recipe to adopt *any possible logical principle*, but a list of notable examples. That being said, the recipe possesses some degree of flexibility intended to deliver fruitful applications under several specific formalisms. As anticipated, a first (deliberate) scope of manoeuvre is given by the way in which premisses of inferences are gathered in (1). The most straightforward way to understand "and" is as a metatheoretic juxtaposition sign, very much like commas in a sequent calculus formulation.[9] In this way, the final detaching step that we call "Modus Ponens" becomes akin to an application of the structural rule of cut. One then easily sees that, under this reading, the principle of conjunction introduction "if φ and ψ, infer $\varphi \wedge \psi$" is unaffected by the adoption problem.

As noticed by Kripke (2024) himself, if instead one identifies "and" with the object linguistic conjunction, conjunction introduction might acquire a status analogous to Modus Ponens and Schematic Substitution, because gathering premisses via conjunction presupposes the rule of conjunction introduction. An alternative may be to dispense with conjunction, and consider the operation of gathering premisses via nested conditionals (e.g., "if φ and ψ, then χ" is turned into "if φ, then ψ only if χ"). Under this assumption, other principles will become unadoptable, such as the principle of conditional introduction.[10]

The extent to which SCS is a logical rule can be debated at length: it can even be argued that it is *the* logical rule, as it is possible to axiomatize, say, classical logic, by resorting to axioms involving specific predicate letters—and not axiom schemata or rule schemata—and some principle akin to SCS. For our concerns, however, what matters is that the form of universal instantiation that Kripke suggests is presupposed by our capability of acquiring Universal Instantiation is not as strong. Rather, it is a very specific form of universal instantiation that has much to do with our ability of recognizing and combining syntactic patterns.

---

9 Again, some vagueness concerning different implementations of this idea is assumed: we do not take a stance on whether commas should be understood as distinguishing elements in a set, a multiset, or a sequence.

10 We would like to thank an anonymous referee for asking for a clarification of the status of what we call "Modus Ponens." Given our purpose, any choice that is more specific than our current proposal would lead to specific choices that are not compatible with the general analysis of Kripke's project that is the main aim of the paper.

The problems encountered with the adoption of a logical rule—as far as Kripke's example is concerned—boil down, therefore, to the necessity of certain presuppositions to the process. Under a plausible reading of the pattern isolated by Kripke, such presuppositions amount to competence with Modus Ponens and the validity and a very specific form of universal instantiation SCS.[11]

2.2 *Where Can We Adopt?*

In Kripke's example, the receiver of the instructions may not be able to perform any inference. The scenario is compatible with a *tabula rasa* adoption. Let us now consider a more realistic, although still highly idealized, scenario in which an agent is in possession of *some* inferential abilities that are in need of revision. In general, revisions can reasonably involve either (i) dropping some principle from the set of one's logical beliefs, or (ii) adding principles to it.[12] We call the former process Drop, and the latter Add.

Most cases of proposed logical revision at the heart of modern and contemporary debates involve Drop. Starting with classical reasoning, intuitionists proposed to drop the law of excluded middle or, equivalently, to weaken one of the rules for negation. Paracomplete and paraconsistent logicians also propose to drop one of the rules for negation, although their weakening of classical negation is more severe than the one proposed by the intuitionists. Some subtler proposals are also possible. Supervaluationists, for instance, agree with all inferences of classical logic of the form $\langle \Gamma, \varphi \rangle$, but disagree on inferences with multiple conclusions.[13]

Let us start with Drop. There are various scenarios compatible with dropping a logical principle. In the trivial case, revision simply amounts to disre-

---

11 A recent paper by Finn (2021) makes use of the same idea, but erroneously assumes that the ingredients of this "recipe" are Modus Ponens and Universal Instantiation and that both of these rules are individually necessary and jointly sufficient for the adoption of any other logical rule. As we argue here, the recipe doesn't require Universal Instantiation in full generality but only a very restricted form. Also, depending on the logical rule in question, Modus Ponens is not always necessary either (just consider rules that allow adding theorems to any step in the reasoning). As explained, those two rules may also be not jointly sufficient.

12 Of course it is possible that the proposed adoption in question leads from a set of logical beliefs to another which is inconsistent with the previous one, but in the reasonable cases in which this happens one can always describe this process as the result of first dropping some rule and then adding to the remaining principles some other principles.

13 For instance, they drop the classical inference $\langle \{\varphi \lor \neg\varphi\}, \{\varphi, \neg\varphi\} \rangle$.

garding some principle, previously regarded as logical. There is no adoption involved in the revision, and *a fortiori* there is no adoption problem. In a slightly less trivial case, the rule that needs to be adopted is not one of the principles that fall under our understanding of Modus Ponens.[14] In this case, it is clear that the patter of adoption straightforwardly applies (modulo some adjustments required by the specific formalism employed and discussed above). For instance, an agent who is able to infer according to Modus Ponens and SCS is in the position to adopt the familiar principles involving conjunction, disjunction, negation.

Another case of DROP may concern the adoption of a new rule by restricting the scope of previously acquired rules. The crucial (and non-trivial) case involves adoption of restricted versions of Modus Ponens. Some paraconsistent logics, Priest's LP for instance, result from classical logic via the restriction of the elimination rule for the conditional to formulae that are not truth value gluts (2008). Similarly, non-transitive logics restrict the meta-inference of Ripley (2015), by allowing it only for some non-pathological sentences. In such a scenario, a crucial issue concerns whether the pattern of adoption should be itself revised to feature such restricted detachment principles instead on the original form of Modus Ponens. Luckily, the answer is positive. If one wanted to apply the pattern for adoption to the restricted Modus Ponens, schematic substitution and the restricted form of Modus Ponens would suffice. The (re)adoption of *other* principles by means of restricted Modus Ponens may be more problematic. For instance, paraconsistent logics such as LP feature unrestricted principles governing conjunction and disjunction, and therefore the adoption of such principles will not involve only sentences with a classical truth value.

Problems can occur only, if the logical resources become to weak to apply the principle even with suitably restricted rules. Whether there are interesting cases of that kind, will be explored below.

Let us now turn to ADD. *Prima facie* there are good reasons to doubt the significance of ADD, if one assumes that the process of adoption has classical logic as its starting point and restricts oneself to the propositional case. The Post completeness of classical propositional logic tells us that the only consequence relation that properly extends it is the trivial one. On the other hand, when we move to first-order classical logic, which isn't Post-complete, it is also clear that Modus Ponens and Universal Instantiation are already

---

14 We are leaving out SCS from the picture, because of its special status.

in place. Therefore, any revision that follows our schema for adoption is also unproblematic—new rules can be adopted and applied by following the pattern for adoption isolated above. For instance, we might consider a higher-order version of the rule of existential introduction:

(2) from $\varphi(R)$, infer $\exists X \varphi(X)$

with R a set variable which is free for X in φ. As before, the adoption of such a rule would require the capability of applying SCS. In the specific case of (2), the schematic variable needs to be of a suitable type; it should be capable of taking variables like X as arguments. This process, however, is still carried out once a suitable language is fixed. The substitution involved in the adoption of (2) does not require any substantial decision on the semantic status of the different types of variables. Similarly, a higher-order version of the rule of (monadic) Universal Instantiation

(3) from $\forall X \varphi(X)$, infer $\varphi(P/X)$

can be accommodated in our framework via SCS once a suitable language is fixed. What is only required is that the schematic variable φ can be instantiated to a specific formula of the higher-order language one is considering. In other words, in the pattern of adoption for (2) and (3), one always assumes a specific domain of syntactic entities on which SCS operates. And this is all that seems to be required.

As expected, the only problematic candidates in the context of ADD are logics that either don't have what we called Modus Ponens or do not have SCS. It is fair to say that, if one operates in Kripke's idealized scenario of a *tabula rasa* adoption, our analysis deems the unrestricted rule of Modus Ponens as unadoptable. However, it is equally fair to say that the debate is still open on whether logics that do not feature Modus Ponens satisfy some fundamental adequacy requirements for playing the role of a *logica utens*, i.e. whether a logic without such a rule could be an adequate formal model for any possible form of natural reasoning. We rest content with the claim that, for the overwhelming majority of case studies, the last step (3 to 4) of our pattern of adoption applies.

What about SCS? It is a common assumption in much of contemporary semantics that natural languages must (in some way, Cohnitz 2005) be *compositional*. How else could it be explained that we can use and understand new sentences with novel meanings? However, compositionality requires some

form of systematic syntactic decomposition and of keeping track of how, for example, argument places of predicates are filled. It is hard to see why such capacity shouldn't already be sufficient for the kind of schematic substitution that Kripke's example requires. Compositionality by itself guarantees that competence with a sentence like "Sam kisses Martin" entails competence with "Martin kisses Sam," "Reinold kisses Julie"—this fact is behind the systematicity argument for compositionality (Szabó 2000). But then the basic skills involved in processing a compositional language (treating linguistic items as schematic and (re)combinable with other linguistic items of certain syntactic categories) already allow one to reason in accordance with SCS. This skill doesn't seem to be in need of "adoption."[15]

SCS is weaker than the rule of Universal Instantiation. It is a basic (logical or linguistic) skill that is presupposed by reasoning of any kind. Not just any logical rule we learn, but learning any new compositional phrase requires mastery of schematic substitution.[16] A fortiori, any logic that is supposed to model an actual *logica utens* will have to contain SCS then.

Again, there can be formal systems that are weaker than classical logic and that do not contain Modus Ponens or SCS. But the real question is whether there is any formal system that models a *logica utens* but fails to enable the reasoner to adopt a new rule. If any application of logical rules requires some (suitably restricted form of) Modus Ponens and SCS, and if from that a reasoner can obtain a (suitably generalized) form of Modus Ponens and SCS that is sufficient for grasping the application conditions for a new rule, then every logic that is a possible *logica utens* will allow upwards adoption (as well as downwards adoption to any logic that is a possible *logica utens*). If this is right, then Kripke's "adoption problem" does not actually pose a problem for the adoption of a new logic.

But Kripke's scenario is anyway highly artificial. No one adopts a logic simply because some oracle told them that the principle behind it is logically valid. We may come to reason in new ways, because we adopted a new theoretical perspective on matters of validity.

---

15 To be precise, for the application of SCS in reasoning, we need not only the ability to compose new expressions, but also to decompose them. This requires compositionality, as well as *inverse compositionality* (Pagin 2003).

16 And, as we argued above, schematic substitution is implicit in our mastery of composing and decomposing complex expressions in general.

3   Adoption in a Logical Theory

We have argued that the revision of logic by adoption of a new logical principle is best understood as a revision of one's *logica utens*. In this section we consider the patterns of adoption isolated earlier in the arguably more realistic context of a *logical theory*, typically defined as a collection of principles governing the core notions involved in one's specific account of logical consequence: truth-preservation, predication, negation, implication, assertion, formality, consistency, provability and so on. Therefore, giving a full account of one's preferred logical theory is often a highly non-trivial matter. That the Adoption Problem discussed by Kripke should carry over to these more realistic contexts is clear from the discussion of empirically motivated logical revision found in Kripke (2024).

3.1   *Deflationary Views of Logical Theories*

The preliminary characterization of logical theories just given is not the only one considered in the literature. It more or less aligns to what Hjortland (2017) calls *non-deflationary* logical theories. Following this terminology, a typically deflationary account is the one articulated in Williamson (2017), which holds that the ultimate task of logical theories is to unravel general claims about the world. Meta-linguistic notions such as truth and validity are not the primary concern of logic, which is essentially a *non-metalinguistic* enterprise pointed at discovering absolutely general laws of reality. In this, logic does not differ from physics, or from metaphysics; it only proceeds at a much higher level of abstraction.

    Williamson suggests that a logical theory is a collection of nonmetalinguistic generalizations corresponding to logical truths. This picture is motivated by the following process: Williamson starts from valid inferences in some logic S in a language $\mathcal{L}_S$—e.g., $\neg\neg\varphi \therefore \varphi$. It proceeds by extending $\mathcal{L}_S$ with new, higher-order variables of the same type as formulae of $\mathcal{L}_S$ and by replacing the entailment relation with a conditional—in our example, this turns $\neg\neg\varphi \therefore \varphi$ into $\neg\neg X \rightarrow X$. The process is then completed by universally quantifying over the free higher-order variables of the translation of the logical claim under considerations. A logic, in this view, is a collection of claims such as $\forall X(\neg\neg X \rightarrow X)$. Endorsing a logic is endorsing a collection of universally quantified claims: since there is no reason to consider higher-order quantification as more metalinguistic than first-order quantification (Williamson 2017,

329), a logical theory is no more metalinguistic than any other theoretical enterprise seeking universal laws, such as physics itself.

Given our analysis, the problem of adoption in a deflationary logical theory of the kind just sketched does not arise. Already the process of turning a purported valid inference into a universal generalization of the appropriate type requires a prior understanding of quantification. It is hard to see how this understanding may not involve something as basic as SCS: this is especially clear in the step that requires the expansion of one's language with variables of the appropriate type. The very adequacy of this process seems to rest on the capability of instantiating such variables with formulae of $\mathcal{L}_s$, as required by SCS. Moreover, the substitution of the entailment sign with a suitable conditional certainly presupposes a conditional that satisfies Modus Ponens. How can the reduction be put to use, if one cannot retrieve the original inference by assuming an instance of the antecedent of the law-like conditional and conclude its consequent via Modus Ponens? The structural assumptions required by Williamson's view of logical theories therefore presuppose both SCS and Modus Ponens; our analysis of the pattern for adoption entails that the circularity involved for the adoption of a new rule does not arise in the presence of such principles.[17]

3.2 *Logical Theories*

Logical theories, in the abstract—and more substantial—sense considered in this section, can be seen as the formal counterpart of *logicae utenses*. In the same way as a *logica utens* encodes the agent's dispositions towards a class of inferences (or meta-inferences), a logical theory enriches this acceptance of a class of validities with a collection of meta-theoretic claims concerning semantic and proof-theoretic notions associated with such inferences. For instance, the logical theory of intuitionistic logic includes an account of what is a canonical or direct method of verification, as opposed to an indirect one. Similarly, the logical theory of paraconsistent logic involves a characterization of negation and falsity (and truth) that substantially differs from the classical exclusive approach to negation. Taken at face value, claims of the sort just

---

17 Williamson ultimately rejects this Tarski-Bolzano procedure of bringing inferences to their normal form *as a tool to compare logical consequences*. This is because the procedure requires a strong conditional, and many of the logics involved in the comparison will not have it. What we said however still stands: on this view of logical theories Modus Ponens and SCS are essential requirements.

described belong to the metatheory of one's logic. And such a metatheory typically amounts to a fragment of classical or intuitionistic mathematics.

There are at least two possibilities to formulate the adoption problem in this richer framework, depending on what one considers to be the class of logical principles that can be adopted/revised. On the one hand, one might consider revision and adoption of the purely logical part of one's metatheory, which may not align with the object-theoretic logical principles. On the other, one can extend the status of logical principle to core metatheoretic principles such as consequence and truth, and consider their adoption and revision.[18]

Let us consider the scenario in which one wishes to revise/adopt *logical principles* of one's overall logical theory, including the logic of metalinguistic concepts. In the abstract case, it is clear that this is no more nor less problematic than allowing for a revision of object-linguistic logical principles: the logical component of one's logical theory is simply a collection of inference patterns that one recognizes as valid *in one's metatheory*. There seem to be no substantial differences between the analysis of the local adoption problem above and the present case: again, the only problematic cases might be cases of ADD, in which from a weaker metatheory one moves to a stronger metatheory. For instance, one might ask whether the intuitionistic logician is able to adopt a classical perspective on validity. In the current setting, this can simply be reduced to the problem of whether one can instruct an intuitionist to infer according to, say, double negation elimination $\neg\neg\varphi \therefore \varphi$. But in the presence of SCS and Modus Ponens, we have seen that this is unproblematic: one starts with exhibiting a specific doubly negated instance $\neg\neg A$ of $\neg\neg\varphi$; by SCS, one provides the intuitionist with the concrete instance of—a suitable translation of—the original principle "if $\neg\neg A$, then A." From $\neg\neg A$ and "if $\neg\neg A$, then A,"

---

18 One might also think about a third option, in which one's logical theory plays a purely instrumental role. In this scenario, one would keep all metatheoretic principles fixed, consider them in a purely instrumental role, and take into account only adoption and revision for the object-theoretic logical inferences. The discussion of the previous section would then largely transfer to this case, with possibly a further complication. Suppose we are in the crucial case of the absence of Modus Ponens on one's object-theoretic logical toolbox. In this case the instrumentalist about metatheory may find herself in the position of not accepting (yet) object-theoretic claims of the form $\varphi, \varphi \rightarrow \psi \therefore \psi$, but accepting—given a standard set-theoretic semantics:

   (4) If "φ," is true and (if "φ" is true, then "ψ" is true), then "ψ" is true.

where "is true" is a standard Tarskian truth predicate for the object language. Therefore, the instrumentalist would have to argue that, even though she is able to infer on the basis of principles such as (4), she is in no position to adopt Modus Ponens at the object linguistic level.

the agent that possesses the general capability of inferring by Modus Ponens can immediately conclude *A*. Under the assumption that intuitionistic or classical foundations are the only reasonable candidates for the logic of the metalinguistic components of one's logical theory, we can safely conclude that no worries of circularity can arise in this second reading of logical theories.

The assumption that one's logical metatheory is framed in classical or intuitionistic set theory may be questioned. There have been interesting attempts, in the context of some approaches to the semantic paradoxes, to align a weaker nonclassical approach—generally substantially weaker than intuitionistic logic, since semantic paradoxes affect classical and intuitionistic logic alike—in the object theory with a nonclassical metatheory (Leitgeb 2007; Bacon 2013; Weber, Badia and Girard 2016). Such attempts, however, are at best at an initial stage and cannot yet be considered to be actual rivals of a classical or intuitionistic metatheory. For instance, most of these meta-theoretic results heavily rely on a classical meta-meta-theory. What would be required is a non-classical set theory (or of an alternative foundational framework) in which all metatheoretic reasoning could be performed.

The status of non-classical set theories, however, is controversial. Let us consider for instance on some paracomplete and paraconsistent options. Partial set theories have been developed by Gilmore (1974), Aczel and Feferman (1980), Feferman (1984): the naive comprehension principle is built on top of a three valued logic such as Strong Kleene logic. Consistency is obtained by showing (in a classical metatheory) that membership can be interpreted by means of a positive inductive definition. The main drawback of such attempts consists in their deductive weakness: the theories are able to recover only a fragment of predicative mathematics.

Paraconsistent set theories have also been extensively studied in recent years. Several combinations of set-theoretic and logical principles are possible. One option is to formulate naive comprehension on top of the LP (Restall 1992; Priest 2006). Due to the weakness of the conditional of LP, it is not clear whether this option can deliver standard set-theoretic results such as Cantor's theorem, or even the existence of two objects (Weir 2004). An alternative is to replace the conditional of the paraconsistent logic with a relevant conditional. In this way, a substantial amount of standard results of classical set theory can be obtained (Weber 2012). However, doubts still remain about the adequacy of such an option: as argued in Incurvati (2020, chap. 4), the relevant conditional is insufficiently motivated, and the fundamental extensional nature of the set

concept is compromised of in such approaches—there are sets that have the same members but that are not identical.

We are left with the possibility of adopting/revising quasi-logical principles such as truth and falsity. This is, arguably, the option that is closest to actual cases of revision of one's logical assumptions. Paraconsistent and paracomplete logicians motivated by semantic or logical paradoxes, for instance, aim at a revision also of foundational tools, such as comprehension axioms, that are needed to define their notion of logical consequence. In this context, one considers not only a collection of logical inferences, but also the principles of quasi-logical notions such as truth, property predication, and consequence as possible candidates for revision. Can the worries of circularity/regress adumbrated in the local case of adoption in the previous sections have some bearing on such cases of revision?

If the adoption/revision process is a local process involving some specific quasi-logical rules and follows the blueprint of Kripke's setup, our analysis in section 2 can be transferred with only little modifications. For instance, if one's logical theory makes essential use of the notion of truth, one might want to adopt/revise suitable principles for the truth predicate, e.g., a disquotational rule of the form "from φ, infer True($\ulcorner\varphi\urcorner$)."[19] If Modus Ponens and SCS are available, one can essentially follow the pattern outlined above for the case of adopting a logical rule such as double negation in an intuitionistic logical theory. The only step that requires care is the selection of a suitable range instance of instances of SCS. In the case unrestricted schemata such as double negation, in fact, specifying a range of instances of SCS is a trivial affair: all sentences of the language are allowed. By contrast, due to the Liar Paradox, selecting a suitable range for the instances of φ in the truth rules might prove to be involve resources that are very complex in computational terms. We cannot choose all instances whatsoever to avoid inconsistency, and a more sophisticated procedure is needed. Now, if this procedure is purely syntactic, it can be easily implemented in the pattern for adoption stated above without any ad hoc move. For instance, if one intends to adopt the rule "from φ, infer True($\ulcorner\varphi\urcorner$)" for instances of φ that do not contain "True," the relevant specification of the range of SCS is a fairly simple procedure—at most primitive recursive—and can be reasonably taken to be part of the conceptual

---

19 A couple of qualifications about the example: first, the rule should be intended to apply also to φ that we have assumed, and not only proved. Secondly, this rule should be intended to be adopted together with other truth rules. These qualifications are needed to ensure that the rule characterizes truth, and not weaker notions such as provability.

toolbox of anyone that understands the syntax of the language of their logical theory.

If the specification of the relevant instances of SCS is not syntactic, it may result in a more complex procedure. If, for instance, this involves selecting the grounded sentences in the sense of Kripke (1975), or the set of stable truths in the sense of the revision theory of truth (Gupta and Belnap 1993), this would involve a highly non-computable process (McGee 1988; Burgess 1986). Therefore, we might have a situation in which there is no Kripke-style circularity in adopting "from φ, infer True($\ulcorner\varphi\urcorner$)," but simply the absence of a suitable schematic substitution rule to implement in the pattern of adoption for such rule. It should be clear, however, that this scenario is perfectly compatible with our analysis of the problem of adoption/revision. Whereas the adoption problem concerns the one's (seeming) impossibility of inferring according to a rule that is available to her, in the scenario under consideration the agent *does not* have at her disposal a suitable version of SCS to perform inferences, because its range may be too complex too be specified.

We are then left with the familiar scenario in which one would like to adopt/revise quasi-logical rules but does not possess Modus Ponens. We have already cast some doubts on the availability of a workable *logica utens* in the absence of Modus Ponens. In the context considered here, this is even more so, since a logical theory may involve complex semantic constructions couched in classical mathematics, which require a substantial use of classical logic.

4 Alternative Quinean Targets for Kripke's Argument

For all we have argued so far it seems that there is no adoption problem that would pose an obstacle or challenge to the idea that we can rationally revise our *logica utens*, provided that prior to the revision we already possess some basic reasoning skills and that our revision is supposed to preserve these. Neither in the abstract scenario that Kripke presents nor in more realistic cases is it plausible to assume that we lack the resources to apply new logical rules in reasoning.

As we explained in section 1, we started with discussing the case of revision, since that seemed to us the broadest target for the adoption problem. In this last section, we will look at other aspects of a broadly Quinean philosophy of logic that could be potential targets of an adoption problem.

We could identify four possible alternative targets that are part of Quine's conception of logic and may, at least *prima facie*, be affected by the proposed

regress. The candidates are in turn the adoption of a first logic, the transition from the acceptance of a principle to the adoption of certain behavior, the problem of the missing normative force of purely descriptive logical principles, and the *knowledge that/knowledge how*-distinction. We will discuss the candidates in this order.

4.1 *The Adoption of the First Logic*

So far we have considered the adoption problem as as a challenge for Quine's idea that we can adopt a *new* logic. So it was legitimate in our argument to suppose that some logic and some language is already in place and that an individual has on the basis of some *reasoning* arrived at the conviction that she should adopt a different way of reasoning, that she should adopt a new logic.

But perhaps is best understood in close similarity to Quine's original point against conventionalism and concerns the question how—on Quine's view— logic could have ever gotten off the ground (1936). After all, also on the conception that logic is just general, firmly held belief (1951), there seems to be the issue that firmly believing Modus Ponens does not yet allow you to reason with it, if you don't yet have that capacity. Thus, as a general theory of what logic is, Quine's theory isn't better than conventionalism, since it still is open to the challenge that it can't explain how the first logical principles could have been adopted in absence of an already existing logic.[20]

Although this well may be so, it is not clear that this is a challenge that Quine needs to address. Or, in other words, it seems to us that Quine, quite clearly, does not have to address it. Quine (1936) presents a picture according to which the first principles of logic are not adopted as a result of engaging with some explicit formulation of the principles (as conventionalism has it), but where they get adopted in behavior and only later are reconstructed in terms of explicit reasoning principles or rules. This adoption in behavior does not require that Quine's theory of belief revision applies to it, so he does not at all need to explain how *homo sapiens* managed to develop structured reasoning that is describable in terms of schematic inference principles. This should be part of a general naturalistic account of how higher cognition and reasoning in general developed. To require that Quine's conception of logic provides some detailed explanation of this process is entirely inadequate.

---

20 This seems to be how Padro (2015) understands the adoption problem.

It is worth emphasizing that conceding that the acceptance of logical principles cannot explain how reasoning got off the ground takes nothing away from the idea that principles of logic are (as far as epistemology is concerned) just like other hypotheses.[21] By developing a *logica docens* (as a formal representation of our most general ways of reasoning) we can critically study the way we think about most general matters (or matters most generally) and maybe decide to make revisions to those central aspects of our web of beliefs. Just as we would do with other hypotheses. How we could then "adopt" the so revised logic, we have described above.

4.2 *From Belief to Behavior*

A second potential target for the regress argument is Quine's idea of the status of logic in the web of belief. Quine (1951) considers logic to be nothing but firmly held *belief*, statements that are just like any other statements in the web of belief, with the only difference, that they are more central than others, and thus less likely to be given up. But adopting a logic is not just adopting some belief. It is adopting a way of reasoning. There are two ways to make that challenge. The first would be to see this as a critique of Quine's behaviorism.[22] For a behaviorist, having a certain belief (for example, the belief that Modus Ponens is valid) just means to show certain forms of behavior (for example to reason in ways that are licensed by Modus Ponens). But perhaps that's too short-sighted. As Kripke's thought experiment shows (on this interpretation), one may accept a belief (viz. that Modus Ponens is valid) and yet fail to show the appropriate behavior (e.g., to assent to implications that are licensed by Modus Ponens). The thought experiment then doesn't show that there indeed is a regress or circularity problem, but that there may be a problem of a certain kind of "stubborness": someone may count as having grasped and accepted

---

21 To see this, maybe it helps to consider an analogy with scepticism about our senses. Maybe we need to have already default trust in our senses in order to be able to learn anything from them. That may make the hypothesesis "I can trust my senses" special (in comparison to other empirical hypothesis) insofar as my knowledge of the world wouldn't get off the ground without it. But even if that were so, *this* would not make this hypothesis immune to revision, not even immune to revision via information that I receive through my senses.

22 Quine's behaviorism is a well-known aspect of much of his work. We already encountered it in Quine (1936), when Quine argues that there is no difference between firm belief and implicit and spontaneous acceptance of a convention. Most famously, Quine's behaviorism shows in his arguments in Quine (1960).

a certain belief, but just doesn't act in a way that may be canonical for the ascription of that belief.

This may be a reasonable challenge to the idea that "*S* believes that *p*" can be analysed as "*S* is disposed to assent to this and that under conditions such and such." But this doesn't seem to be a specific problem for Quine's theory of logic than rather a problem for Quine's theory of belief. However, while the regress/circularity argument *displays* the problem, it doesn't actually establish anything that could seriously be regarded as an argument for the claim that such an analysis must fail. It seems still perfectly reasonable to just respond to such argument that it merely shows that the person in the dialog who doesn't reason in accordance with, for example, Modus Ponens has not yet actually adopted the relevant belief.

4.3 *The Normative Force of Logical Principles*

A closely related challenge (one that actually makes use of a regress) is to interpret Kripke's argument as revealing that Quine overlooked the *normative* nature of logic (if one believes that is has such a normative nature). Logic, on this view, tells us how we ought to reason. However, the general principles that are featured in Kripke's thought experiment are not norms or imperatives. They don't say anything about how anyone *should* reason. Therefore there is a gap between adopting the belief that a certain logical principle is true and adopting the *norm* that one ought to reason in a certain way. Quine, who takes logical principles to be just like any other general scientific hypotheses, overlooks this.

There are two reasons why this is not a plausible target for Kripke's argument. First, not all logical norms or imperatives will hold unconditionally. But if they are norms that apply under certain conditions, then also a conditional norm could do nothing about the regress. The reasoner in the Kripke scenario would still need to be already following that norm in order to apply it under the current conditions. Thus, just adding deontic force to a rule doesn't help with the regress at all, our hypothetical reasoner would still have to instantiate the general norm to the current case and then detach a consequence concerning what they now ought to infer.

The second reason is that the plausible normative force of logical principles is in fact to weak to be of any help in Kripke's thought experiment. As Besson (2018) explains, the recent discussion of the normative force of logic strongly suggests that in order for the argument to go through, we'd need an imperative

or a rule that would "move" a subject to reason in accordance with the logical principle at issue. However, as we have learned from Harman (1986) and others, logical principles can't give rise to such rules. It simply isn't always rational to use Modus Ponens and endorse *q* whenever you believe *p* and $p \supset q$ for some *p* and *q*. However, a weaker principle that would, say, allow that it is rationally permissible to believe *q* whenever you believe *p* and $p \supset q$ for some *p* and *q* is plausible, but would not lead to a plausible regress (see Besson 2018 for details). Once you know the principle

(5) Given your beliefs *P* and (if *P*, then *Q*), you are rationally permitted to reason to *Q*.

we can explain why you should be rationally permitted to reason with Modus Ponens. If the regress argument is supposed to make a point about normativity, it simply operates with the wrong deontic force.

4.4 *Knowledge That and Knowledge How*

This leaves us with a last candidate which again tries to explain the problem of the regress by a certain insufficiency of the merely propositional knowledge that we acquire, when we accept the claim that Modus Ponens is valid. We mentioned in the beginning in section 1 that Priest as well as Devitt and Roberts both see the problem of adoption as primarily an issue of acquiring certain *knowledge how* after one has convinced oneself of the relevant *knowledge that*. Stairs (2006) also seems to understand Kripke in this way.

Take a familiar analogy: from reading a book about how one rides a bike, one doesn't know yet how to ride a bike in the sense that one won't be able (yet) to ride a bike. The latter will require certain practical competence, a skill, that can not be acquired by simply reading a description of what that skill involves. Instead, the acquisition of that skill might require training.[23] In the regress argument, the subject accepts Modus Ponens but doesn't have the skill to apply it, she thus gets a new bit of propositional knowledge which she doesn't know how to apply either, and so forth.

Priest as well as Devitt and Roberts seem to think that also the adoption of a new logic requires that we train ourselves in the application of a rule

---

23 We don't distinguish here between knowledge how and a skill, for the purpose of our argument it is sufficient to note that there are skills for which it is true that they can't be acquired by just understanding an instruction.

in order to be able to apply it. However, as our discussion above shows, the competence that rule application of logical principles requires is merely the competence with basic inferences like Modus Ponens or SCS for the examples that we considered. For instance, one could infer according to Conjunction Elimination by just plugging the conditional rule into our recipe above. The relevant knowledge how, in these cases, is a certain basic capacity to reason in the first place. Adoption of a new rule thus does not require training in new rules.

Another question may be what it takes to "see" new implications that one didn't see as implications with the "old" logic, or how one can get to stop seeing implications that aren't implications according to a new logic. This seems to be what Kripke has in mind when he is complaining that a merely formal account of logic would not be the same as an intuitive form of reasoning:

> What I mean is this: you can't undermine intuitive reasoning in the case of logic and try to get everything on a much more rigorous basis. One has just to think not in terms of some formal set of postulates but intuitively. That is, one has to *reason*. [...] One can only reason as we always did, independently of any special set of rules called "logic," in setting up a formal system or in doing anything else. (Stairs 2016) [24]

This version of the adoption problem seems to be what Kripke originally had in mind, but it neither leads to a regress, nor is it very convincing. The regress is irrelevant, since the problem is not that a logical rule is missing and requires the introduction by some explicit statement of the rule (the application of which again requires the rule, and so on ad infinitum). The problem is rather that any formal statement of logical laws is not the same as a way of reasoning. Thus, whether such a formal account is stronger or weaker than our actual way of reasoning, or in our terminology, whether revision goes via DROP or ADD, is irrelevant; if a formal logic does not agree with our intuitive way of reasoning, we will not be able to adopt such logic. Seeing that a consequence follows is as impossible to adopt as unseeing that a consequence follows, according to that view.

The point is then not that we need training to be able to apply a new rule (i.e. to be able to apply a new general rule to a new concrete case). As we

---

24 Stairs (2016) and Stairs (2006) are also discussions of Kripke's lectures, but focus primarily on his case against quantum mechanics and less on a reconstruction of the adoption problem.

argued above, application of the rules is easy once you have the skill necessary to follow our recipe. The problem is rather that such a form of application of an explicit rule does not count as *reasoning*.

But why should it not? Why should the habituation of a logic have any special status? Kripke presumably does not want to give the same value to all our dispositions to draw inferences intuitively. We often make mistakes in our intuitive reasoning. Maybe reasoning is a complex cluster of dispositions for Kripke; dispositions to draw inferences as well as dispositions to retract them after reflection. But if reasoning is such a wider cluster, then reasoning is malleable. And once reasoning is malleable in light of new information about our inferences not being valid (maybe on the basis of a formal representation of that inference), why stop there? Why only consider reasoning as a set of dispositions stable under such a narrow equilibrium, rather than stable under a wider equilibrium that considers more general principles of theory choice, e.g., fruitfulness, etc. The latter is just the anti-exceptionalist, Quinean view.

Carroll/Quine-type considerations do not provide support for excluding a wide equilibrium view, nor an argument against the possibility of habituation or the malleability of reasoning.*

Daniel Cohnitz

0000-0001-5958-8572

Utrecht University

d.cohnitz@uu.nl

Carlo Nicolai

0009-0003-1027-5294

King's College London

carlo.nicolai@kcl.ac.uk

References

ACZEL, Peter and FEFERMAN, Solomon. 1980. "Consistency of the Unrestricted Abstraction Principle using an Intensional Equivalence Operator." in *To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, edited by Jonathan P. SELDIN and J. Roger HINDLEY, pp. 67–98. London: Academic Press.

AZZOUNI, Jody. 2014. "A Defense of Logical Conventionalism." in *The Metaphysics of Logic*, edited by Penelope RUSH, pp. 32–48. New York: Cambridge University Press, doi:10.1017/CBO9781139626279.

BACON, Andrew. 2013. "Non-Classical Metatheory for Non-Classical Logics." *The Journal of Philosophical Logic* 42(2): 335–355, doi:10.1007/s10992-012-9223-9.

BESSON, Corine. 2018. "Norms, Reasons, and Reasoning: A Guide Through Lewis Carroll's Regress Argument." in *The Oxford Handbook of Reasons and Normativity*, edited by Daniel STAR, pp. 504–528. Oxford Handbooks. New York: Oxford University Press, doi:10.1093/oxfordhb/9780199657889.001.0001.

BURGESS, John P. 1986. "The Truth is Never Simple." *The Journal of Symbolic Logic* 51(3): 663–681, doi:10.2307/2274021.

CARROLL, Lewis. 1895. "What the Tortoise Said to Achilles." *Mind* 4(14): 278–280, doi:10.1093/mind/iv.14.278.

COHNITZ, Daniel. 2005. "Is Compositionality an A Priori Principle?" in *The Compositionality of Meaning and Content. Volume I: Foundational Issues*, edited by Markus WERNING, Edouard MACHERY, and Gerhard SCHURZ, pp. 23–58. Linguistics and Philosophy n. 1. Heusenstamm b. Frankfurt: Ontos Verlag, doi:10.1515/9783110323627.

COHNITZ, Daniel and ESTRADA-GONZÁLEZ, Luis. 2019. *An Introduction to the Philosophy of Logic*. Cambridge Introductions to Philosophy. Cambridge: Cambridge University Press, doi:10.1017/9781316275573.

DEVITT, Michael and ROBERTS, Jillian Rose. 2024. "Changing Our Logic: A Quinean Perspective." *Mind* 133(529): 61–85, doi:10.1093/mind/fzad010.

EBBS, Gary. 2011. "Carnap and Quine on Truth by Convention." *Mind* 120(478): 193–237, doi:10.1093/mind/fzr020.

FEFERMAN, Solomon. 1984. "Toward Useful Type-Free Theories I." *The Journal of Symbolic Logic* 49(1): 75–111, doi:10.2307/2274093.

FEIGL, Herbert and SELLARS, Wilfrid, eds. 1949. *Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts.

FIELD, Hartry. 2007. "Solving the Paradoxes, Escaping Revenge." in *Revenge of the Liar. New Essays on the Paradox*, edited by J. C. BEALL, pp. 78–144. Oxford: Oxford University Press.

FINN, Suki. 2019. "The Adoption Problem and Anti-Exceptionalism About Logic." *Australasian Journal of Logic* 16(7): 231–249, doi:10.26686/ajl.v16i7.5916.

—. 2021. "Limiting Logical Pluralism." *Synthese* 198(suppl. 20): 4905–4923, doi:10.1007/s11229-019-02134-8.

GILMORE, Paul C. 1974. "The Consistency of Partial Set Theory without Extensionality." in *Axiomatic Set Theory*, edited by Thomas J. JECH, pp. 147–154. Proceedings of Symposia in Pure Mathematics n. 13.2. Providence, Rhode Island: American Mathematical Society. Proceedings of the Symposium held at the University of California, Los Angeles, July 10 – August 5, 1967.

GUPTA, Anil and BELNAP, Nuel D., Jr. 1993. *The Revision Theory of Truth*. Cambridge, Massachusetts: The MIT Press.

HARMAN, Gilbert H. 1986. *Change in View*. Cambridge, Massachusetts: The MIT Press.

HJORTLAND, Ole Thomassen. 2017. "Anti-Exceptionalism About Logic." *Philosophical Studies* 174(3): 631–658, doi:10.1007/s11098-016-0701-8.

INCURVATI, Luca. 2020. *Conceptions of Set and the Foundations of Mathematics*. Cambridge: Cambridge University Press, doi:10.1017/9781108596961.

KRIPKE, Saul A. 1975. "Outline of a Theory of Truth." *The Journal of Philosophy* 72(19): 690–716. Reprinted in Kripke (2011, 75–98), doi:10.2307/2024634.

—. 2011. *Philosophical Troubles*. Collected Papers n. 1. Oxford: Oxford University Press.

—. 2024. "The Question of Logic." *Mind* 133(529): 1–36, doi:10.1093/mind/fzad008.

LEITGEB, Hannes. 2007. "On the Metatheory of Field's 'Solving the Paradoxes, Escaping Revenge' [on Field (2007)]." in *Revenge of the Liar. New Essays on the Paradox*, edited by J. C. BEALL, pp. 159–183. Oxford: Oxford University Press.

MCGEE, Vann. 1988. *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*. Indianapolis, Indiana: Hackett Publishing Co.

MORRIS, Sean. 2018. "Quine Against Lewis (and Carnap) on Truth by Convention." *Pacific Philosophical Quarterly* 99(3): 366–391, doi:10.1111/papq.12185.

PADRO, Romina. 2015. "What the Tortoise Said to Kripke: the Adoption Problem and the Epistemology of Logic." PhD dissertation, New York: City University of New York, https://academicworks.cuny.edu/gc_etds/603/.

PAGIN, Peter. 2003. "Communication and Strong Compositionality." *The Journal of Philosophical Logic* 32(3): 287–322, doi:10.1023/a:1024258529030.

PRIEST, Graham. 1987. *In Contradiction. A Study of the Transconsistent*. Nijhoff International Philosophy Series n. 39. Den Haag: Martinus Nijhoff Publishers.

—. 2001. *An Introduction to Non-Classical Logic*. 1st ed. Cambridge: Cambridge University Press.

—. 2006. *In Contradiction. A Study of the Transconsistent*. Oxford: Oxford University Press. First edition: Priest (1987), doi:10.1093/acprof:oso/9780199263301.001.0001.

—. 2008. *An Introduction to Non-Classical Logic. From* If *to* Is. 2nd ed. Cambridge: Cambridge University Press. First edition: Priest (2001), doi:10.1017/CBO9780511801174.

—. 2014. "Revising Logic." in *The Metaphysics of Logic*, edited by Penelope Rush, pp. 211–223. New York: Cambridge University Press, doi:10.1017/CBO9781139626279.016.

Putnam, Hilary. 1969. "Is Logic Empirical?" in *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*, edited by Robert S. Cohen and Marx W. Wartofsky, pp. 216–241. Boston Studies in the Philosophy of Science n. 5. Dordrecht: D. Reidel Publishing Co. Reprinted as "The Logic of Quantum Mechanics" in Putnam (1975, 174–197).

—. 1975. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. Cambridge: Cambridge University Press. Second edition: Putnam (1979).

—. 1979. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. 2nd ed. Cambridge: Cambridge University Press. First edition: Putnam (1975).

Quine, Willard van Orman. 1936. "Truth by Convention." in *Philosophical Essays for Alfred North Whitehead*, edited by Filmer Stuart Cuckow Northrop, Charles Hartshorne, Raphael Demos, Paul Weiss, Willard van Quine, Scott Buchanan, Henry S. Leonard, S. Kerby-Miller, and Otis H. Lee, pp. 90–124. London: Longmans, Green & Co. Reprinted in Feigl and Sellars (1949, 250–273) and in Quine (1966, 77–106).

—. 1951. "Two Dogmas of Empiricism." *The Philosophical Review* 60(1): 20–43. Reprinted in Quine (1953, 20–46), doi:10.2307/2266637.

—. 1953. *From a Logical Point of View: 9 Logico-Philosophical Essays*. Cambridge, Massachusetts: Harvard University Press. Cited after the revised edition: Quine (1961).

—. 1960. *Word and Object*. Cambridge, Massachusetts: The MIT Press. New edition: Quine (2013).

—. 1961. *From a Logical Point of View: 9 Logico-Philosophical Essays*. 2nd ed. Cambridge, Massachusetts: Harvard University Press. Revised edition of Quine (1953), reprinted 1980.

—. 1966. *The Ways of Paradox and Other Essays*. Cambridge, Massachusetts: Harvard University Press. Second, enlarged, edition: Quine (1976).

—. 1976. *The Ways of Paradox and Other Essays*. 2nd ed. Cambridge, Massachusetts: Harvard University Press. First edition: Quine (1966).

—. 2013. *Word and Object*. Cambridge, Massachusetts: The MIT Press. First edition: Quine (1960).

Restall, Greg. 1992. "A Note On Naive Set Theory in LP ." *Notre Dame Journal of Formal Logic* 33(3): 422–432, doi:10.1305/ndjfl/1093634406.

Ripley, David. 2015. "Naive Set Theory and Nontransitive Logic ." *The Review of Symbolic Logic* 8(3): 553–571, doi:10.1017/s1755020314000501.

Stairs, Allen. 2006. "Kriske, Tupman and Quantum Logic: The Quantum Logician's Conundrum." in *Physical Theory and its Interpretation. Essays in Honor of Jeffrey Bub*, edited by William Demopoulos and Itamar Pitowsky, pp. 253–272. The University of Western Ontario Series in Philosophy of Science n. 72. Dordrecht: Springer Verlag.

—. 2016. "Could Logic Be Empirical? The Putnam-Kripke Debate ." in *Logic and Algebraic Structures in Quantum Computing*, edited by Jennifer Chubb, Ali Eskandarian, and Valentina Harizanov, pp. 23–41. Cambridge: Cambridge University Press, doi:10.1017/CBO9781139519687.004.

Szabó, Zoltán Gendler. 2000. *Problems of Compositionality*. New York: Garland Publishing Co.

Weber, Zach. 2012. "Transfinite Cardinals in Paraconsistent Set Theory." *The Review of Symbolic Logic* 5(2): 269–293, doi:10.1017/s1755020312000019.

Weber, Zach, Badia, Guillermo and Girard, Patrick. 2016. "What Is an Inconsistent Truth Table?" *Australasian Journal of Philosophy* 94(3): 533–548, doi:10.1080/00048402.2015.1093010.

Weir, Alan. 2004. "There Are No True Contradictions." in *The Law of Non-Contradiction: New Philosophical Essays*, edited by Graham Priest, J. C. Beall, and Bradley Armour-Garb, pp. 385–417. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199265176.001.0001.

Williamson, Timothy. 2017. "Semantic Paradoxes and Abductive Methodology." in *Reflections on the Liar*, edited by Bradley Armour-Garb, pp. 325–346. New York: Oxford University Press, doi:10.1093/oso/9780199896042.001.0001.

Animalism with Psychology

Robert Francescotti

Here I develop an account of our persistence that accommodates each of the following compelling intuitions: (i) that we are animals, (ii) that we existed prior to the onset of whatever psychological capacities are necessary for personhood, and we can continue to exist with the loss of those and other psychological capacities, (iii) that with suitable psychological continuity, the person goes with the brain/cerebrum in remnant person and brain/cerebrum transplant cases, and (iv) that it is possible for us to survive gradual large-scale replacement of organic with inorganic parts. With the help of a couple of recent "hybrid" animalist accounts, I develop an analysis of our persistence that entails (ii)–(iv) while being consistent with (i).

Here I develop an account that captures each of four compelling and widely discussed intuitions regarding what we (human persons) are and the changes we can undergo while continuing to exist. What makes this project significant is that each of these intuitions is strong enough that it would be desirable to find an account that allows the truth of all four. Yet, these four initially appear to form an inconsistent set, and they really do seem quite difficult to reconcile. Here are the intuitions.

We are members of the species *Homo sapiens*, which is a species of animal. It would seem to follow that we are animals. The claim here is not merely that each of us is causally or otherwise intimately associated with an animal, as even those who believe that we are immaterial souls can accept. The animalist claim is that each of us *is* an animal in the strict sense of identity. That is:

Each of us is numerically identical with an animal.[1]

---

[1] Rejecting this identity claim does not require believing that there are immaterial souls. One might endorse the *constitution* approach (e.g. Baker 2000, 2007) according to which a human person is constituted by but not identical with the spatially coincident animal. Or one might accept the *brainist* view that we are proper parts of animals, specifically, the functioning brain or the part of it responsible for the psychological states that make us persons. (See, for example, McMahan 2002; Parfit 2012; Campbell and McMahan 2016).

It is also tempting to believe that no special psychological states or capacities are necessary for our persistence. As Olson (1997b, 1997a) pointed out, it seems that each of us once existed as a fetus, before the onset of the psychological capacities commonly associated with personhood (such as rationality and self-awareness). It is also tempting to think that we can survive the loss of those and various other psychological functions, as arguably happens, for example, when one enters a persistent vegetative state.[2] So in addition to the animalist identity claim, there is the following intuition:

> We existed prior to the onset of whatever psychological capacities are necessary for personhood, and we can continue to exist with the loss of those and other psychological capacities.

Of course, this belief is not independent of the view that we are animals. It seems that what it takes for an animal to continue to exist is a biological affair and not a matter of psychology.

However, suppose that the brain of some person is removed and placed in the cranium of a different body. Suppose, also, that the brain transplant is a success, with higher-level neural functioning retained in such a way that the animal with the brain after the operation is perfectly psychologically continuous with the animal that had the brain before the operation. It is tempting to think that the person goes with the brain and therefore has switched bodies. Or suppose that rather than being placed in a different body, the excised brain is kept disembodied but with cerebral functions sustained in such a way that psychological capacities definitive of personhood remain. There is the intuition that in this case the brain is a person, what Johnston (2007) calls a "remnant" person, and the same person as the earlier embodied individual, if psychological continuity is maintained.

We can imagine that instead of removing the entire brain, just the cerebrum is removed and sustained in such a way that all of the psychological capacities definitive of personhood are retained, and with ample psychological continuity. Imagine also that lower brain functions are preserved in the body left behind so that the body remains alive. Despite the fact that the cerebrumless body remains alive, there is the strong intuition in this case, again because of the psychological continuity, that the person goes with the cerebrum.[3] This

---

2 Olson (e.g. 1997a, 111–114) expresses the intuition that one would persist in a vegetative state.
3 This thought-experiment assumes that the psychological activity/capacities definitive of personhood are confined to the cerebrum. If you think they extend to other regions of the brain, then

belief seems at odds with the animalist thesis that we are animals, for it appears that the animal goes with the cerebrumless body and not the cerebrum. Also, if our persistence does not require psychological continuity, if some sort of non-psychological (e.g. purely biological) continuity is sufficient, then there would seem to be no reason to deny that the person persists as the living animal left behind. So there would appear to be a conflict between the first two intuitions and the following belief:

> With the right sort of psychological continuity retained, the person goes with the brain/cerebrum in remnant person and brain/cerebrum transplant cases.

If we go with the brain or cerebrum, then assuming that a brain or a cerebrum is not itself an animal, we can persist without being an animal.[4] Given that animals are organisms and that to be an organism one must be at least largely organic, another way for us to persist without being an animal is to have all or most of our organic parts replaced with inorganic bits. If the replacement is done gradually enough with physical continuity intact, it is tempting to believe that we would persist in that inorganic form. So, there is also the following intuition:

> It is possible for us to survive gradual large-scale replacement of organic with inorganic parts.

That seems to conflict with the view that we are identical with animals, on the assumption that animals cannot persist without being sufficiently organic.[5]

---

imagine that a larger portion of the brain is removed, so long as the portions that control vital functions are left behind.

4 One might support the view that the detached whole brain, with activity suitably sustained, counts as an animal or an organism at least. See, for example, van Inwagen (1990, sec.15) and Olson (1997a, 133). But it seems much less plausible to regard a mere cerebrum as an animal/organism; "a detached cerebrum is not an animal, or a living organism of any other sort" (Olson 1997a, 115). So, as a potential threat to animalism, the transplant case is more effective when imagining that only the cerebrum is transplanted. (Although, see Madden (2016b, fn. 32) who raises doubt for the view that a functioning detached cerebrum is not an animal; also, Madden's (2016a) account of our persistence discussed here in section 1 might be used to support the idea that an animal can persist as a cerebrum and while remaining an animal.)

5 Baker (2016) presents the possibility of surviving inorganic replacement as a reason to reject animalism.

If we wish to develop an account of our persistence that allows the truth of all four of the intuitions mentioned, then it seems we will need to emphasize both psychological and non-psychological continuity. In section 1, I discuss two "hybrid" accounts that have been offered according to which the animals that we are have persistence conditions that are partly psychological and partly non-psychological.[6] These proposals go a long way toward reconciling the four intuitions, but as shown in section 1 they do not fully succeed. Building on these accounts, I develop a hybrid proposal in sections 2 and 3 that more successfully captures the four intuitions.

1 The Hybrid Approach

"*The biological continuity approach* (BCA)," Langford reports, tells us that in non-branching cases "we persist iff we have a biological continuer" whereas "[t]he *psychological continuity approach* (PCA) affirms that, again restricting our attention to non-branching cases, we persist iff we have a psychological continuer" (Langford 2014, 356–357). With "iff" BCA and PCA are characterized as views about what is necessary for our persistence and what is sufficient. Langford proposes retaining the sufficiency claims of BCA and PCA and rejecting the necessity claims, with a disjunctive account according to which we fall under the substance concept, *bio-psycho-continuer*, where "[s]omething counts as a bio-psycho-continuer only if (in non-branching cases) it can persist by way of either biological continuity or psychological continuity" (Langford 2014, 361).[7]

Langford offers reasons for expressing the notion of a bio-psycho-continuer in terms of a necessary condition alone, with "only if" instead of "if and only

---

6 The description "hybrid" for this sort of view is not uncommon. Noonan uses the label "the Hybrid Approach" for the view that "we are animals the persistence conditions of which are partly biological and partly psychological" (2003, 205), with examples of support including Wiggins (1996, 246) and McDowell (1997, 237). Noonan develops a hybrid account, using the label "the hybrid view" for "the complex view that takes psychological continuity as a sometimes sufficient but not a necessary condition for personal persistence" (2019, 196) and mentioning (Noonan 2019, 226, 229) Langford (2014) and Madden (2016a) as advocates. Also see Noonan's (2021) defense of the hybrid approach. Olson (1997b) uses "the hybrid proposal" for the sort of account that describes our persistence in terms of both psychological continuity and biological continuity, while entailing that neither is necessary for our persistence. In this paper I describe as hybrid any account of our persistence that includes psychological and non-psychological conditions.

7 Compare with Sharpe's "psychologically-serious animalism," which is a *conjunctive* approach; "biological and psychological continuity are individually necessary but only jointly sufficient for the persistence of human persons" (2015, 65).

if" (2014, 361). But even with just a necessary condition, the account accommodates three of the four intuitions mentioned earlier. Since the account is disjunctive, it allows that we can persist without psychological continuity, and even in the total absence of psychological capacities, provided there is non-branching biological continuity. So the account allows that each of us once existed as a fetus and as an embryo, and that we can continue in a persistent vegetative state. Being disjunctive, the view also grants that we can survive major or even total inorganic replacement where biological continuity is lost; non-branching psychological continuity suffices. Moreover, Langford's view that we persist by way of either biological or psychological continuity is compatible with the claim that we are animals, animals that are bio-psycho-continuers.

Given that biological continuity is required to continue to be an animal, on Langford's view, even though we are animals, we are not animals essentially. While "animalism" is sometimes used for the view that we are animals essentially, it is often used only for the view that we are animals, with the essential claim regarded as an additional component.[8] With Langford's disjunctive account, one can be an animalist in the latter sense while also accepting that we can survive inorganic replacement and thereby become non-animals. Denying animal essentialism is also a way for animalists to believe that we follow the brain/cerebrum in remnant person and transplant cases even while denying that the cerebrum or the whole brain counts as an animal.

But is Langford's view really compatible with the cerebrum intuition? Suppose the cerebrum is removed with psychological functioning maintained, and suppose the cerebrumless body is destroyed immediately after. The disjunctive view tells us that the person goes with the cerebrum, provided that psychological continuity is sustained. But suppose that the cerebrumless body is not destroyed. Suppose, also, that it remains alive with lower brain functions intact. In this case, we have a biological continuer in addition to a psychological continuer. So where does the person go on Langford's account? Recall his proposal that "[s]omething counts as a bio-psycho-continuer only if (in non-branching cases) it can persist by way of either biological continuity or psychological continuity" (2014, 361). By explicitly applying only to

---

8 Olson (2015a) gives the title "weak animalism" to the bare claim that we are animals, and he uses the description "new animalism" for the conjunction of weak animalism and the denial of any further claims animalists often make, e.g. that animals are animals essentially or fundamentally. Olson also uses the labels "accidental animalism" (2015b) and "accidentalism" (2016) for the view that we are accidentally and not essentially animals.

non-branching cases, the analysis yields no verdict in the branching case we are imagining.

Is there a way of extending the account to branching cases? The most straightforward way of doing so is with the claim that we persist in branching and non-branching cases by way of either psychological or biological continuity. But, as Langford (2014, 365) mentions, we then get the result that in the scenario imagined the person survives as both the cerebrum and the cerebrumless animal, and so the person ends up in two places at once, which seems implausible. Yet, as Langford points out, non-disjunctivists, including proponents of PCA and BCA, also have to deal with branching cases (e.g. where there is more than one psychological continuer or more than one biological continuer), and their options for doing so are also available to the disjunctivist (Langford 2014, 365). One of the options Langford mentions (besides holding that the person ends up in two places at once) is to accept the view that in the case described, there were two persons within one body, two spatially coincident persons, before the operation, and one of them goes with the cerebrum while the other goes with the cerebrumless body. (See Noonan's (2019, 2021) way of developing a "multiple occupancy" version of the disjunctive approach.)[9] However, there is the strong intuition that in the case imagined the person goes with the cerebrum and not the living body left behind. So I think a more plausible option for the disjunctivist is to add Nozick's (1981) notion of a "closest continuer" to Langford's account, and then develop the account in a way that ensures that the cerebrum is the closest continuer in the branching case we are imagining. I will present what I consider the best way to do that in sections 2 and 3. The solution proposed is not simply a matter of claiming that in branching cases, psychological continuity always wins out over physical/biological continuity, and the solution offered is not a disjunctive one. But before we turn to my proposal, let's get a clearer idea of how to proceed with a brief discussion of Madden's (2016a) hybrid account.

With the belief that we are fundamentally biological organisms of the kind *human animal*, Madden claims that "[o]ne of us persists if and only if a sufficient number of capacities for human-animal-characteristic activity

---

9 Lewis (1976) and Robinson (1985) defend the view that with fission there is more than one individual there all along. See how Noonan (2019, 140–144, 225–231; 2021) defends this "multiple occupancy" view, within the framework of a disjunctive hybrid account, in order to preserve the only *x* and *y* principle that, roughly, whether *x* is identical with *y* can depend only on facts about *x* and *y* and how they interrelate.

are continuously preserved (along a dominant path)" (2016a, 6).[10] Madden provides a list of some of the activities characteristic of the kind *human animal*. The list includes biological activity of various sorts (such as digesting, ageing, and fighting infection) as well as activities characteristic of non-biological material aggregates (e.g. resisting penetration and blocking light). In the list of activities characteristic of the human animal kind, Madden also mentions various psychological functions, including planning, remembering, visually attending, and problem-solving. So, on his account, "psychological capacities are relevant to our persistence" (Madden 2016a, 6).

Requiring only that a *sufficient* number of capacities are preserved allows that we can exist without psychological capacities, either before their presence or after their loss—provided that enough of the biological or purely physical abilities are preserved (continuously along a dominant path). Also, on Madden's view, we can persist even with the loss of biological capacities, assuming a sufficient number of our purely physical or psychological ones are retained. So it seems that on his account, we can survive replacement of organic with inorganic parts.

What about persisting as a remnant person? The account allows for that possibility by including in the list of activities characteristic of the human animal kind various psychological functions. With the inclusion of psychological activity, the person goes with the excised brain/cerebrum on Madden's view, provided that enough psychological activity or the capacity for it is continuously preserved (along a dominant path).

But suppose that a living body were left behind. Where would the person go in that case? If the cerebrum were removed, with all person-making psychological features sustained, then certain capacities that Madden considers characteristic of the human animal kind would follow the cerebrum. Of course, other capacities would go with the remainder of the body, and with life sustained via lower brain functions, it seems that many more capacities would follow the remainder of the body than would follow the cerebrum.

Wishing to maintain that the person goes with the suitably functioning cerebrum even with a living body left behind, Madden (2016a, 7) points out that the "single term 'thinking' grossly underestimates the number and diver-

---

10 Madden is following Wiggins' development of the broadly Aristotelian view that, as Madden puts it, "[a] macroscopic continuant is, most fundamentally, a locus of law-like activity characteristic of its general kind" (2016a, 4). Langford (2014) also appeals to Wiggins's view of persistence in terms of continuation of an object's principle of activity.

sity of human-organism-characteristic capacities preserved" by the cerebrum, including among others:

> [the capacity for] colour discrimination, grammatical string detection, social hierarchy navigation, duration sense at different temporal scales, vertical-horizontal line discrimination, face recognition, place recognition, practical know-how, auditory phoneme individuation, predictive naïve physics, story-telling, [and] episodic memory (Madden 2016a, 7)

If the number and diversity of psychological capacities such as these make the path of the cerebrum the dominant path in the brain removal case being imagined, then we can say with Madden's account of our persistence what many of us are inclined to say—that the person goes with the cerebrum and not the cerebrumless body even if the latter remains alive with the capacity to control vital functions.

Granted, the number and diversity of psychological capacities preserved by the cerebrum is much greater than what the single term "thinking" implies. However, as Kotak (2018) mentions, it is doubtful that those functions controlled by the cerebrum outweigh the host of other functions (biological, chemical, and physical) that remain in the cerebrumless individual. Those other functions on Madden's list of activities characteristic of the human animal kind include biological activities found in most kinds of organism (e.g. growing, excreting, aging, digesting, sweating, and dying); they also include "activities characteristic of simple material concretions," such as resisting penetration and blocking light (Madden 2016a, 6). It is unclear at best that the path of psychological continuity would be the more dominant one if the body left behind had the great wealth of non-psychological capacities and activity that the life of a human organism requires. Moreover, as Hershenov points out:

> There won't clearly be psychological dominance if we imagine the transplant of a very unsophisticated human mind due to damage or developmental immaturity […]; [yet] most readers would assume that the subject of experience has moved with the minimally sentient human mind (Hershenov 2020, 91)

So it is doubtful that Madden's account adequately captures the intuition that we go with the psychologically continuous cerebrum in cases where the cerebrumless body remains alive.

I take it that it's a merit of Madden's account, and Langford's, that psychological continuity is not considered necessary for our persistence, for this allows that we predate and can outlast psychological capacities. It is also a merit of their proposals that they don't regard biological continuity as necessary for our persistence, which is consistent with our surviving inorganic replacement. There is the additional virtue that their accounts, while not requiring biological continuity, allow that each of us is identical with an animal. The proposal developed in the next two sections maintains these merits, while more effectively capturing the intuition that we go with the brain/cerebrum in remnant person and transplant cases.

2 Towards a New Hybrid Account

There might have been, and perhaps there are, persons who are immaterial substances (gods, angels, or Cartesian egos perhaps) or have immaterial substances as proper parts. What I am calling "physically realized" persons, by contrast, are those who are confined to physical space and at some level of composition are wholly comprised of physical parts.[11] Working on the assumption that we are physically realized, the question arises: What does it take for any physically realized person, human or otherwise, to persist?

Let's use "P" to indicate whatever psychological activity and capacities are definitive of personhood (such as self-awareness and rationality) and let "P-continuity" designate whatever continuity of P-activity/capacities grounds the persistence of persons. Also, "x at t" is used here to abbreviate "x as x is at t," and "y at t*" is used to abbreviate "y as y is at t*," without presupposing that "x at t" and "y at t*" name temporal parts.[12] As a first approximation, then, the proposal is:

> H. Necessarily, if x is a physically realized person at time t, then for any y at time t*, $x = y$ if and only if

> (i) y at t* is physically continuous with x at t, and
> (ii) y satisfies the following requirement:

---

11 I talk of *physically realized* persons rather than physical persons to avoid the assumption that all of the properties of a person are reducible to physical properties. Talk of physical realization allows the truth of a non-reductive physicalist account of our mental and, say, biological properties.

12 The discussion here is meant to be neutral on whether endurantism is true. This formulation and any of the points expressed in this paper in endurantist terminology can be rephrased in 4-dimensionalist terms.

its P-continuity at t^* with x at t is not exceeded by the P-continuity with x at t of anything else at t^* that is physically continuous with x at t.

H applies to all physically realized persons, including any non-human physically realized persons there might be. According to H, we and any other physically realized persons persist as *the physical continuer that preserves an unsurpassed degree of continuity of psychological activity/capacities definitive of personhood*. I use "H" as short for "Hybrid Account," indicating that like Langford's and Madden's accounts, H emphasizes both psychological and non-psychological continuity. Also, H is only a first approximation. Some modifications are needed, as might already be apparent, and these will be added in section 3. For the remainder of this section let's consider how the core analysis, H, accommodates the various intuitions regarding our persistence.

H does not specify which psychological activity or capacities count as definitive of being a person, and H is also neutral on the type of psychological continuity that matters to our persistence (overlapping chains of direct memory links, suitable causal connections, sharing a first-person perspective, or similarity of personality traits). H is neutral as well on which factors make it the case that one instance of P-continuity outweighs another. Thus, as intended, the analysis is compatible with a wide range of differing opinions on what type of psychological continuity, and how much and what degree, is a factor in our persistence.

H also leaves unspecified how best to construe physical continuity. One might choose to think of it in terms of spatio-temporal continuity of a sufficient portion of matter, with different possible views on how much is sufficient (e.g. more than half or perhaps some more sizable majority). Or one might wish to understand physical continuity in terms of some type and degree of causal continuity of physical processes, either in addition to or instead of the emphasis on spatio-temporal continuity.[13] So H may be adopted by those with differing views on what counts as physical continuity. However, while H leaves much latitude on how to understand physical continuity, not just anything goes. The analysis is not meant to require any sort of biological continuity—so as to allow that persons like us who are biological might survive large-scale replacement of organic with inorganic parts. The physical continuity mentioned in H is also to be understood, as one would expect, in

---

13 For better insight on different ways to understand physical continuity, see Sauchelli's description of different ways to view the persistence conditions of bodies of matter (Sauchelli 2017, 212–213).

such a way that one is physically continuous with the earlier fetus and the cognitively deprived individual one might become; and it is to be understood in such a way that the excised cerebrum is physically continuous with the person from which it was removed and the person receiving the implant. This allows that the person existed before and can survive the loss of P-capacities, and it allows that the person goes with the brain/cerebrum in remnant person and transplant cases. And condition (ii) helps ensure these results.

Suppose some person x is awaiting cerebrum-removal. The cerebrum is removed with P-activity/capacities sustained so that we end up with a remnant person, y, who is P-continuous with x. Since y is small relative to the cerebrumless body left behind, y is not the item at the time that is most physically continuous with x. The cerebrumless body is more physically continuous with x than y is. Still, y is physically continuous with x to some degree, and on the assumption that P-capacities are housed in the cerebrum,[14] y is a physically continuous item with unsurpassed P-continuity—unsurpassed by the P-continuity with x of anything else at the time that is physically continuous with x. So with the emphasis in (ii) on being a physical continuer with unsurpassed P-continuity, H gives the intuitively correct result that x goes with the cerebrum and not the rest of the body even if the cerebrumless portion remains alive. (For ease of exposition here and in some of what follows I abbreviate. When I say that y is continuous with x, I mean that *y as y is at the one time* is continuous with *x as x is at the other time*. For example, saying above that y is physically continuous and P-continuous with x is shorthand for saying that y as y is in the remnant person state is physically continuous and P-continuous with x as x is/was before surgery.)

To satisfy the requirement specified in condition (ii) of H, y's P-continuity with x cannot be surpassed, i.e. cannot be surpassed by the P-continuity with x of anything else at the time that is physically continuous with x. However, this does not require that y actually is P-continuous with x. Suppose that person x begins to lose all P-capacities, and suppose that what is left is a living body, y, without any P-capacities. Since y has no P-capacities, y is not P-continuous with x. But, we may suppose, nothing else at the time that is physically continuous with x has any more P-capacities than y has. In this sense, y's P-continuity with x is unsurpassed. So H gives the result that the

---

14 As mentioned in footnote 3, if you think the psychological activity/capacities definitive of personhood extend to other regions of the brain, then imagine that a larger portion of the brain is removed. Also, extended mind considerations are set aside here solely for the sake of simplicity.

person persists in this case (though presumably not as a person) despite the loss of P-capacities.

Also consider some adult person, x, and the former fetus, y, prior to the gain of P-capacities. While x and y are physically continuous, there is no P-continuity between x and y. However, nothing at the time that is physically continuous with x has any more P-capacity than y has. So y's P-continuity with x, albeit lacking, is unsurpassed (by the P-continuity with x of anything else at the time that is physically continuous with x). Thus, H gives the result that x and y are the same individual in this case, thereby capturing the intuition that we existed before any P-capacities were acquired.

H also allows that we existed prior to and can survive the loss of *all* psychological capacities. Suppose that y is physically continuous with x, but y has no psychological capacities. Then y is not psychologically continuous with x, and since P-continuity is a type of psychological continuity, y is not P-continuous with x. But suppose that nothing else at the time that is physically continuous with x has any psychological capacities. Then nothing at the time that is physically continuous with x has any more P-continuity with x than y has. Assuming that x is a physically realized person, $x = y$, according to H.

H does not entail that we are animals. One might accept the analysis and argue that we are constituted by animals and not identical with them. One might also accept H and believe that the brain is the nonderivative locus of mentality, and one might conclude that the person is the brain or some part of it even when it is inside an animal. But, clearly, one can consistently accept H while believing that we are animals, for one can consistently believe that the persistence conditions of those human animals who are persons (and any other animals that might qualify as persons) are as described by conditions (i) and (ii). So suppose that we conjoin H with the belief that we are identical with animals. The result is an animalist account that has as a strong point in its favor the fact that it honors each of the other intuitions—that we existed prior to and can survive the loss of person-making and other psychological capacities, that we go with the brain/cerebrum in remnant person and transplant scenarios, and that we can survive inorganic replacement.

However, there are some concerns with H that show that a few additions are required.

3 Modifying H

Imagine that a person, *x*, splits in half, leaving two individuals, *y* and *z*, both of whom are physically continuous with *x* (as *x* was before the division), and equally so. Also suppose that *y* and *z* are equally P-continuous with *x*, and nothing else at the time that is physically continuous with *x* is more P-continuous with *x* than *y* and *z* are. According to H, both *y* and *z* are identical with *x*, which cannot be given that *y* is not identical with *z*. Unless we wish to maintain that there were two persons present all along, some addition to H is needed.

We might add as a third condition that the person persists only as *the most dominant* physical continuer with unsurpassed P-continuity. With this addition, we get the desired result that in the case imagined where there is no dominant physical continuer, the pre-fission person does not survive the division. But imagine the case is slightly different. Person *x* divides, leaving *y* and *z*, and *y* at the time contains just a little more of *x*'s matter before the division than *z* does—suppose only two ounces more. Also suppose, as before, that *y* and *z* are equally P-continuous with *x*, each with unsurpassed P-continuity. Many would be reluctant to accept that a difference in just a couple of ounces could make a difference in where the person goes.[15] So it seems that the third condition added to H's (i) and (ii) should require not just greater physical continuity, but *significantly* greater physical continuity—the most dominant physical continuer (with unsurpassed P-continuity) and significantly so. For the purpose of the analysis, let's leave it open what significantly greater physical continuity amounts to, thereby allowing for different views on how much additional physical continuity is significant enough to make a difference in whether the person survives in cases of physical branching.

In section 2, the physical continuity mentioned in H was left unspecified so that H may be accepted by those with differing views on how best to construe physical continuity. Yet, what counts as physical continuity by some standards might not be the sort of physical continuity that is necessary for our persistence. So in modifying H we might wish to require *suitable* physical continuity. For example, we probably would want to place limits on how far along we persist in the process of death and decomposition. One might be inclined to accept the *Termination Thesis*, the view that we cease to exist when

---

we die.[16] The proponent of this thesis might believe that the type of physical continuity required for us to persist involves continuity of life. Although, if we demand continuity of biological life, then it seems we cannot retain the idea that we might survive massive inorganic replacement (given that something is a biological entity only to the extent that it is largely organic). Rather than requiring continuity of biological life, those attracted to the Termination Thesis might instead require continuity of a certain sort of *internal complexity*, where the requisite internal complexity might be explained in terms of the manner and degree to which one's internal processes are causally interdependent. So long as it is a type of internal complexity that corpses lack, then the revised analysis would be acceptable to proponents of the Termination Thesis. And as long as it is a type of internal complexity had by the typical human fetus, a human in a persistent vegetative state, a remnant person, and a human survivor of inorganic replacement, the analysis continues to honor the intuition that we persist in those cases.

Even those who reject the Termination Thesis, believing that we survive death as corpses, would probably maintain that our persistence requires continuity of some degree or type of internal complexity to preclude our surviving to the very end of the decomposition process.[17] It is not clear how the details of the internal complexity are best specified, although it is clear that there are different types and degrees of internal complexity had by a freshly dead human corpse but lacked by, say, a human skeleton.

Suppose, then, that we preface the two occurrences of "physically continuous" in H with the word "suitably," where "suitably" serves as a placeholder for whatever restrictions one might wish to place on the physical continuity mentioned in the analysis. Also suppose we add a third condition that requires being the most dominant suitable physical continuer (with unsurpassed P-continuity) and significantly so.

> H'. Necessarily, if x is a physically realized person at time t, then for any y at time t^*, $x = y$ if and only if
>
> (i) y at t^* is suitably physically continuous with x at t,

---

16 See Feldman's (1992) classic discussion and rejection of the Termination Thesis.

17 Mackie, who rejects the Termination Thesis, proposes that "[a]n organism persists for as long as it retains enough of its parts, in a sufficiently similar state of organisation" (Mackie 1999, 238).

(ii) *y* satisfies the following requirement: its P-continuity at *t** with *x* at *t* is not exceeded by the P-continuity with *x* at *t* of anything else at *t** that is suitably physically continuous with *x* at *t*, and

(iii) *y* exhibits significantly greater suitable physical continuity with *x* at *t* than does anything else at *t** that satisfies the requirement of unsurpassed P-continuity specified in (ii).[18]

Provided that suitable physical continuity is not understood in a way that prevents any one of us from being suitably physically continuous with a fetus, someone in a vegetative state, a cerebrum, or someone with an inorganic composition, then like H, H' honors the intuition that we pre-dated the acquisition of and can survive the loss of P-capacities, that the person goes with the P-continuous brain/cerebrum in remnant person and transplant cases, and that we can survive inorganic replacement. Also, nothing about the additions to H rules out our being identical with animals. One can consistently believe that we are animals whose persistence conditions are those described by (i)–(iii).

To better understand the *suitable* physical continuity requirement in H', let's see how the notion of suitable physical continuity might be used to handle some other potential problem cases. Suppose that person *x* starts to experience major psychological changes, so that the person *y* at a later time who is highly physically continuous with *x* before the changes is only modestly P-continuous with *x*. Also imagine that just before the disruption of P-continuity, an exact psychological duplicate of *x* is created out of entirely new matter except for a few of *x*'s atoms. Given the few atoms retained, the psychological duplicate, *z*, is physically continuous with *x* to a minimal degree. Also, if P-continuity does not require *causal* continuity of psychological features, then given the perfect psychological duplication *z* is more P-continuous with *x* than *y* is, in which case, H' seems to give the result that *x* survives as *z*. Some might consider it implausible to believe that *x* and *z* are the same person in this case especially given that *y* is highly physically continuous with *x* and P-continuous with *x* to some degree. If we wish to avoid the result that *x* and *z* are the same person, one obvious way to do so is to insist that more than just a relatively minimal amount of physical continuity is required for a person to persist (where the sum of only a few atoms is minimal relative to the whole

---

18 Like H, H' is expressed in endurantist terminology (especially with talk of *x* being identical with y), but the formulation can be rephrased in 4-dimensionalist terms and with reference to temporal parts made explicit (with "*x*-at-*t*" and "*y*-at-*t**") if one wishes.

body while a brain or cerebrum is not). By including in the idea of suitable physical continuity that there is more than just a relatively minimal amount of physical continuity, we avoid the result in the case described that x and z are the same person.

Suppose, instead, that without any major psychological changes, the cerebrum is taken from person x and successfully transplanted. Also imagine that upon removal, the cerebrum is immediately replaced with a duplicate cerebrum, quickly enough that there is no disruption of psychological continuity. At the end of the procedure, there are two persons: person y who is the recipient of the original cerebrum and person z with the original body and duplicate cerebrum. If P-continuity does not require causal continuity, then we may suppose that in this case, y and z are equally P-continuous with x, and both with unsurpassed P-continuity. Also, because z has all of the original body except for the cerebrum, z is highly physically continuous with x and significantly more so than y is. H' entails that the original person x survives as z (the person with the original body and duplicate cerebrum) and not as y (the recipient of the original cerebrum).

It's not clear to me what to think about this result. But those who find the result counterintuitive can avoid it while still accepting H'. One option is to construe suitable physical continuity as requiring *physical continuity of the locus of* P-*features* (person-making psychological features). Assuming that the cerebrum is the locus of P-features, if suitable physical continuity demands physical continuity of the locus of P-features, then z (the person with the duplicate cerebrum) is not suitably physically continuous with the original person x. So by construing the suitable physical continuity mentioned in H' as requiring physical continuity of the locus of P-features, we can endorse H' without holding that z is the same person as x. H' would instead give the result that x follows the original cerebrum and survives as y. There is another way to secure this result. We might construe P-continuity as requiring causal continuity of P-features. If P-continuity requires causal continuity of P-features, then H' does not entail that x survives as z. The result, instead, is that x survives as person y with the original cerebrum. (Of course, the appeal to causal continuity might also be used to avoid the result in the previous case that the person survives as the psychological duplicate.)

Here's another concern. H' gives the result that if P-continuity is sustained, then the person persists as the cerebrum when it is removed. Now suppose the cerebrum is successfully placed into another body. After surgery, there is the individual, y, that received the cerebrum, and there is the cerebrum

itself, *z*. Assuming that P-capacities are confined to the cerebrum, it seems that *y* and *z* are equally P-continuous with the pre-implant cerebrum, the remnant person, *x*. Many would be inclined to think that in this case after successful implantation *x* persists as the entire individual, *y*, and not just the cerebrum, *z*, contained within. But given H', whether *x* persists as *y* or as *z* depends on which is more physically continuous with *x*. It would seem that on most plausible construals of physical continuity, *y* is not more physically continuous with *x* than *z* is. In fact, it is arguable that *z* is more physically continuous with *x* given that there is so much of *y* that is not physically continuous with *x*. So there is the worry that, with H', *x* remains cerebrum-sized even after implantation—or perhaps worse, *x* does not persist at all assuming neither *y* nor *z* is significantly more physically continuous with *x* than the other is.

If we wish to secure the result that the person persists as the whole animal after implantation and not just the cerebrum, we might include in the idea of suitable physical continuity that a suitable physical continuer is a *maximal* physical continuer, where a maximal physical continuer is one that does not have any suitable physical continuers as proper parts. Given that the whole animal after implantation is a suitable physical continuer of remnant person *x*, the maximality constraint entails that the cerebrum after implantation is not a suitable physical continuer of *x*. So if the whole animal is a suitable physical continuer and if being a suitable physical continuer entails being a maximal physical continuer, then with H' we get the desired result that the person coincides with the whole animal after implantation (and is the whole animal given animalism).[19] Although, perhaps we can secure this result without adding the maximality constraint. One might have some reason, independently of maximality considerations, to believe that when the cerebrum is housed in an animal, the bearer of psychological properties is the animal itself, and not the cerebrum. For example, if we had reason to believe that the person is the genuine bearer of psychological features and that the person is the animal when the cerebrum is contained within, then we would have some reason to believe that the animal, and not the cerebrum, is the bearer of psychological features. If the implanted cerebrum, *z*, is not the bearer of psychological features, then it is not P-continuous with the cerebrum, *x*, before

---

19 Adding this maximality constraint to the notion of suitable physical continuity in H' does not commit one to the view that personhood is maximal. It places no restrictions on whether persons can have persons as proper parts. Although, it does entail that if *y* is a suitable physical continuer of person *x*, then *x* does not persist as any proper part of *y*.

implantation. In that case, *z*'s P-continuity with *x* would be surpassed by the animal's P-continuity with *x*. So then we would get the result, even without a maximality constraint, that the person does not remain cerebrum-sized after implantation.[20]

In further clarification and defense of H', let us consider some of Olson's objections to combining animalism with persistence conditions that are partly psychological.

4 Some of Olson's Objections

Olson mentions a way to formulate a hybrid account of the persistence of human organisms and shows that it has implausible consequences. One might wonder whether an animalist who endorses H' faces the same sort of objection. Olson considers the following proposal:

> if *x* is a human organism at time *t* and *y* exists at time *t**, *x* = *y* *iff* *x* is (uniquely) psychologically continuous, at *t*, with *y* as it is at *t** or no being is psychologically continuous at *t** with *x* as it is at *t* and *y* has the appropriate sort of brute-physical continuity, at *t**, with *x* as it is at *t*" (Olson 2015a, fn. 16).

This is obviously a disjunctive account according to which human organisms, and we (given the animalist belief that we are human organisms), persist with either psychological continuity or physical continuity in the absence of psychological continuity. The problem with this analysis is that it has the following implausible result:

> if your cerebrum were removed from your head and then destroyed, while the brainless animal left behind survived in a vegetative state, you would first go with the cerebrum and then discontinuously 'jump' to the brainless organism, even though there would be neither psychological nor biological continuity across the jump (Olson 2015a, fn. 16).

Consider any disjunctive account on which we persist with either psychological or physical (including biological) continuity. To capture the cerebrum intuition, we might be inclined to add that psychological continuity always

---

20 Thanks, again, to anonymous referees for *Dialectica* for bringing the problems cases discussed in this section to my attention.

wins out when both types of continuity are present, as the disjunctive account Olson considers entails. But then we get the implausible result that Olson describes: the person follows the path of psychological continuity and goes with the cerebrum, and when the cerebrum is destroyed and there is only physical/biological continuity, the person goes with the cerebrumless body. However, like H, H' requires physical continuity (which needn't be biological continuity), and therefore precludes a person discontinuously jumping from one region to another. Since, according to H', the person goes with the cerebrum in its disembodied state and since the naked cerebrum is not physically continuous with the cerebrumless individual, H' entails that the person does not persist as the latter when the cerebrum is destroyed. On H' (and H), the person ceases to exist when the cerebrum is destroyed.

Of course, by requiring physical continuity, H' does not capture all of what some believe about our persistence. There is the not uncommon belief that psychological continuity of the right sort is sufficient for our persistence, for example, that if a person's body were completely destroyed and replaced with a psychological duplicate, then the person would persist as the duplicate despite the absence of physical continuity. Yet, it is not at all clear that mere psychological continuity, even continuity of person-making psychological features, is sufficient for the persistence of a physically realized person, and requiring physical continuity seems the most effective way to reconcile animalism with the cerebrum intuition while also avoiding the implausible result that Olson mentions of an animal discontinuously jumping from one region to another.

Now let's turn to some objections Olson raises that target any attempt to combine animalism with psychological persistence conditions to preserve the cerebrum intuition.[21] The objections to "new animalism," as Olson (e.g. 2015a) calls it, include the following.

(a) Suppose the cerebrum is removed, but biological life continues in the cerebrumless animal left behind. It would seem that the cerebrumless animal is the same animal, the same organism, as the animal before surgery. But if we are animalists who believe that the person goes with the cerebrum, then we are led to believe that the animal left behind is not identical with the animal before surgery. This result is implausible given that the same biological life is present.

---

21 See, for instance, Olson (1997a, 111–123; 2015b, sec.4; and 2015a, sec.9).

Also, (b) if the cerebrumless animal is not the same animal as the one before surgery, then a new animal has come into existence with the removal of the cerebrum. "But can you really create an animal merely by cutting away an organ belonging to another animal—one not even necessary for life?" (Olson 2015a, 105) It seems not.

Furthermore, (c) if we are animalists and believe that the person follows the cerebrum, then we are led to believe that in cases where the cerebrum is successfully transplanted into a different body, the animal that has received the cerebrum is a different animal than the cerebrumless animal there before that was awaiting transplant. Assuming there aren't two human animals in that region after implantation, it follows that the cerebrumless animal has gone out of existence. Yet, it does seem odd to suppose that an animal could be destroyed with the addition of a cerebrum. "[H]ow could you destroy an animal merely by supplying it with the organ—again, not even a vital organ—that it was missing?" (Olson 2015a, 105).

In partial response to these objections, it should be noted that there are worries for traditional animalists who deny that the person goes with the brain/cerebrum, believing that our persistence has nothing to do with psychology. There are the *remnant person* worries for the traditional animalist which are analogous to worries (b) and (c) above that Olson raises for non-traditional ("new") animalists who believe that the person goes with the brain/cerebrum. Johnston (2007) points out that if we were to remove someone's brain and sustain it in such a way that the cerebral activity continues to yield personhood, then given that the person does not go with the brain, a new person (the remnant person) would be brought into existence when the brain is removed. Johnston reminds us that "[y]ou can't bring a person into being simply by removing tissue from something […] unless that tissue was functioning to suppress mental life or the capacity for mental life" (Johnston 2007, 47).[22]

Olson (1997a, 121; 2015b; 2016) describes an additional remnant person worry when the brain or cerebrum (just the cerebrum, suppose) is implanted in a new head. For the traditional animalist, the cerebrum does not become an animal when it is implanted; rather the animal awaiting transplant simply acquires a new part. Since this new part carries with it psychological states sufficient for personhood, the animal becomes a person when the cerebrum is implanted. So, assuming that there aren't two persons in the same skin after

---

22 Olson (1997a, 120) also mentions this problem.

the cerebrum is implanted, the result seems to be that the remnant person ceases to exist upon implantation. This result, Olson points out, conflicts with the *destruction principle* (which is just as plausible as the creation principle Johnston mentions) that you cannot destroy a person merely by surrounding the person with sustaining tissues.

Olson (2015b, sec.4) points out that by denying that we are essentially animals, an animalist can avoid these remnant person concerns by insisting that the person goes with the cerebrum. These non-traditional animalists who support the cerebrum intuition, including any animalists who endorse H', avoid the remnant person worries by denying that a person comes into existence with the removal of a cerebrum and that a person ceases to exist with its implantation. It is true, as Olson adds, that the remnant person problems are avoided only with the cost of the corresponding "remnant-animal" worries, (b) and (c), mentioned above. Still, given the remnant person worries, it is not clear at least from (b) and (c) alone that non-traditional animalists who support the cerebrum intuition, including those who accept H' or certain other hybrid accounts of our persistence, are in any worse shape than the traditional animalists who reject the cerebrum intuition.[23]

Recall objection (a). If we are animalists who believe that the person goes with the cerebrum, then we are led to believe that the animal left behind is not identical with the animal before surgery. But, the objection goes, this result is implausible given that what is involved is the same biological life. Regarding the cerebrumless animal, Olson asks, "Would it not be the organism from which the cerebrum was removed? It would apparently have the same *life*, in Locke's sense of the word [...] that the original animal had" (2015a, 104). However, it is not clear why one should deny that continuity of life can be imparted to different organisms. Biological fission cases show that continuity of life is not sufficient for the persistence of an organism; if post-fission y and z are equally biologically continuous with x (with continuity of vital functions), then given that y is not identical with z, at least one of y and z is not identical with x. It might be insisted that *non-branching* biological continuity, with continuity of vital functions, is sufficient for the persistence of an organism. But whether this is true is precisely what's at issue in the debate over whether the persistence conditions of organisms who are persons are partly psychological. It is true that "if an organism's biological life carries on,

---

23 Olson (1997a, 120–121) mentions this possible defense by one who supports a psychological account of the persistence of animals who are persons.

we should expect it to continue to be the life of that same organism" (2015a, 104). This is what we should expect given that it usually is the case that only one individual partakes of any one life. But that this is what we should expect allows that in certain highly unusual cases, where something arguably relevant to the persistence of a person (such as unsurpassed P-continuity) vies with biological continuity, distinct individuals can partake of the same life. So continuity of life, even if non-branching, does not itself seem a strong enough reason to conclude that it is the life of the same organism.

It is also worth noting that given the distinction between being the same organism and having the same life, objections (b) and (c) make an animalist's support of the cerebrum intuition seem at least a bit more implausible than it really is. Olson describes the non-traditional animalist's view as entailing that animals are "created" and "destroyed" in the transplant case.[24] This description is somewhat misleading. Talk of an organism being created brings to mind a new biological life, a new set of vital processes, coming into existence; and talk of an organism being destroyed suggests that a set of vital processes has come to an end. But this is not what is happening in the transplant case, even for an animalist who believes that the person goes with the cerebrum. The animal before the removal of the cerebrum and the cerebrumless animal left behind are biologically continuous with continuity of vital functions, and when the cerebrum is implanted, the newly equipped animal and the animal before awaiting transplant are biologically continuous with continuity of vital functions. In either case, it is the same life, i.e. the same set of vital processes. This is not creation or destruction of lives. So for the non-traditional animalist who believes that the person goes with the cerebrum, the number of animal *lives* involved in the transplant scenario is just as we would have expected: one on the donor side and one on the recipient side.

It is true, as Olson mentions (e.g. 1997a, 116–117; 2015a, 104–105), that on the non-traditional animalist account, the number of *animals* in the transplant case is more than what one might have thought. According to the "new" animalist who believes that the person goes with the cerebrum, two individuals partake of the same animal life on the donor side (the one before cerebrum removal and the cerebrumless animal left behind). Given that there are two distinct individuals and both are animals, there are two animals. On the recipient side, there is the individual before the cerebrum was removed to make room for a new cerebrum. That's a third animal. And if enough of

---

24 See, e.g. Olson (2015a).

the right sort of cerebral activity is sustained after removal, then with the non-traditional animalist view, the person follows the cerebrum here too, and the cerebrumless animal would be a different individual from the one there before, and a different animal since it is an animal. So, as Olson indicates, on the non-traditional animalist view the total number of animals involved in the case is *four* (the animal after implantation being the same individual as the animal on the donor side before cerebrum removal). That there are four animals involved might seem too implausible to accept. However, if we recognize the possibility of more than one organism partaking of the same life, and if we are also open to there being some good reason to believe that this is what happens in the transplant case (as an animalist who endorses H' would insist), then the result that there are four animals involved is likely to seem not as implausible as it might initially appear.

So, in response to Olson's objections, there are the following points to consider. Animalists can avoid worries (b) and (c) by denying that the person goes with the cerebrum, but only at the expense of incurring the remnant person concerns analogous to (b) and (c), concerns which are avoided by non-traditional animalists who accept H' (or various other hybrid accounts of our persistence). Also, regarding (a): given the conceptual distinction between being the same organism and sharing the same life, it is not clear why one should deny that continuity of life, even non-branching continuity of life, can be imparted to different animals. An animalist who endorses H' would believe that there is good reason not to deny that continuity of life can be shared by different animals. Moreover, if we remain open to the possibility that there might really be some good reason to believe that continuity of life is imparted to different animals in the transplant case, and if we also keep in mind that on a non-traditional animalist account, life is not created or destroyed with the removal or implantation of the cerebrum, then the commitments of the view will perhaps seem not as implausible as Olson's descriptions suggest.[25]

Also, and very importantly, in weighing the pros and cons of conjoining animalism with H', let's not forget that adding H' to the view that we are animals

---

25 Also see Sauchelli's (2017, 213–214) explanation of how we can describe what happens in transplant cases in a way consistent with the cerebrum intuition but without any mysterious creation or destruction of animals. There is also Madden's (2016a) mention of plant cutting to show that it is not metaphysically mysterious for an organism to persist as a relatively small portion of its original size with the much larger portion becoming a new organism, and Madden also mentions plant grafting cases to show that it is not odd to suppose that one biomass can fuse with a much larger one, with the smaller mass persisting as the product of the fusion.

captures more of what many of us are inclined to believe than does a tradi-
tional, non-hybrid animalist view, i.e. each of the four intuitions mentioned
at the start.

5 In Sum

Recall the following three compelling claims:

> We existed prior to the onset of whatever psychological capacities
> are necessary for personhood, and we can continue to exist with the
> loss of those and other psychological capacities.

> With the right sort of psychological continuity retained, the person
> goes with the brain/cerebrum in remnant person and brain/cere-
> brum transplant cases.

And:

> It is possible for us to survive gradual large-scale replacement of
> organic with inorganic parts.

Since each of these is a strong intuition, it is desirable to find a theory that en-
tails all three. To reconcile the first two, we need a hybrid theory in which our
persistence conditions are partly psychological and partly non-psychological.
  If we also wish to maintain the following principle with the animalist,

> Each of us is numerically identical with an animal,

then we need to hold that the persistence of the animals with which we
are identical is a partly psychological affair. Also, to allow the possibility of
inorganic replacement, our analysis cannot entail that biological continuity is
necessary for the persistence of the animals that we are. Yet, with some sort
of physical continuity required, the hybrid account can avoid the implausible
consequence Olson (2015a, fn. 16) described, and mentioned in section 4,
of a person discontinuously jumping from the removed cerebrum after it is
destroyed to the cerebrumless animal left behind.
  So, inspired by the hybrid accounts of Langford (2014), Madden (2016a),
and others, I proposed H which includes the idea of *the physical continuer
with unsurpassed* P-*continuity.* The physical continuer with unsurpassed P-

continuity can be wholly inorganic and it is the P-continuous brain/cerebrum in remnant person and transplant cases. And since unsurpassed P-continuity might be none at all, H allows that we existed prior to P-capacities and can survive their loss. Also, H allows that we are identical with animals (but not that we are essentially animals given that we can persist in an inorganic state or as a brain/cerebrum).

The modification of H, H', ensures that we follow the most dominant physical continuer in branching cases where there is a tie in P-continuity. "Significantly" was introduced to ease concerns about our persisting as more than just a very slightly dominant physical continuer, and the requirement that one is a *suitable* physical continuer was added to allow, among other things, limits on how minimal the physical continuity might be and limits on how far along we survive in the process of death or after. H' is compatible with various opinions on how dominant the physical continuer should be and what sort and degree of physical continuity is suitable. Talk of P-continuity in H' (and H) is also unspecified to allow differing views on what sort and degree of psychological continuity is relevant to our persistence. Thus, H' is a framework within which to accommodate each of the four intuitions mentioned with the basic insight that we persist as the appropriate physical continuer with unsurpassed P-continuity.*

Robert Francescotti
0000-0002-6399-7210
San Diego State University
rfrances@sdsu.edu

Baker, Lynne Rudder. 2000. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press, doi:10.1017/CBO9781139173124.

—. 2007. *The Metaphysics of Everyday Life: An Essay in Practical Realism*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press, doi:10.1017/CBO9780511487545.

—. 2016. "Animalism vs. Constitutionalism." in *Animalism: New Essays on Persons, Animals, and Identity*, edited by Stephan Blatti and Paul F. Snowdon, pp. 50–63. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199608751.001.0001.

Campbell, Tim and McMahan, Jeff. 2016. "Animalism and the Varieties of Conjoined Twinning." in *Animalism: New Essays on Persons, Animals, and Identity*, edited by Stephan Blatti and Paul F. Snowdon, pp. 229–252. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199608751.001.0001.

FELDMAN, Fred. 1992. *Confrontations with the Reaper: A Philosophical Study of the Nature and Value of Death*. New York: Oxford University Press, doi:10.1093/acprof:oso/9780195089288.001.0001.

HERSHENOV, David B. 2020. "Evaluating Hylomorphism as a Hybrid Account of Personal Identity." *Quaestiones Disputatae* 10(2): 86–105, doi:10.5840/qd20201025.

JOHNSTON, Mark. 2007. " 'Human Beings' Revisited: My Body is Not an Animal." in *Oxford Studies in Metaphysics*, volume III, edited by Dean W. ZIMMERMAN, pp. 33–76. Oxford: Oxford University Press.

KOTAK, Aakash. 2018. "The Hybrid Theory of Personal Identity." MPhil thesis, Cambridge: University of Cambridge, Philosophy, https://www.academia.edu/36323799/The_Hybrid_Theory_of_Personal_Identity.

LANGFORD, Simon. 2014. "On What We Are and How We Persist." *Pacific Philosophical Quarterly* 95(3): 356–371, doi:10.1111/papq.12035.

LEWIS, David. 1976. "Survival and Identity." in *The Identities of Persons*, edited by Amélie Oksenberg RORTY, pp. 17–40. Berkeley, California: University of California Press. Reprinted, with a postscript (Lewis 1983b), in Lewis (1983a, 55–73).

—. 1983a. *Philosophical Papers, Volume 1*. Oxford: Oxford University Press, doi:10.1093/0195032047.001.0001.

—. 1983b. "Postscript to Lewis (1976)." in *Philosophical Papers, Volume 1*, pp. 73–77. Oxford: Oxford University Press, doi:10.1093/0195032047.001.0001.

MACKIE, David. 1999. "Personal Identity and Dead People." *Philosophical Studies* 95(3): 219–242, doi:10.1023/a:1004239225105.

MADDEN, Rory. 2016a. "Human Persistence." *Philosophers' Imprint* 16(17).

—. 2016b. "Thinking Parts." in *Animalism: New Essays on Persons, Animals, and Identity*, edited by Stephan BLATTI and Paul F. SNOWDON, pp. 180–207. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199608751.001.0001.

MCDOWELL, John Henry. 1997. "Reductionism and the First Person." in *Reading Parfit*, edited by Jonathan DANCY, pp. 230–250. Philosophers and Their Critics. Oxford: Basil Blackwell Publishers. Reprinted in McDowell (1998, 359–382).

—. 1998. *Meaning, Knowledge, and Reality*. Cambridge, Massachusetts: Harvard University Press.

MCMAHAN, Jefferson. 2002. *The Ethics of Killing. Problems at the Margins of Life*. Oxford: Oxford University Press, doi:10.1093/0195079981.001.0001.

NICHOLS, Ryan, SMITH, Nicholas D. and MILLER, Fred, eds. 2009. *Philosophy Through Science Fiction. A Coursebook with Readings*. London: Routledge.

NOONAN, Harold W. 1989. *Personal Identity*. London: Routledge.

—. 2003. *Personal Identity*. 2nd ed. London: Routledge. First edition: Noonan (1989).

—. 2019. *Personal Identity*. 3rd ed. London: Routledge. First edition: Noonan (1989).

—. 2021. "Personal Identity and the Hybrid View: A Middle Way." *Metaphysica* 22(2): 263–283, doi:10.1515/mp-2020-0007.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, Massachusetts: Harvard University Press.

Olson, Eric T. 1997a. *The Human Animal: Personal Identity without Psychology*. Oxford: Oxford University Press, doi:10.1093/0195134230.001.0001.

—. 1997b. "Was I Ever a Fetus?" *Philosophy and Phenomenological Research* 57(1): 95–110. Reprinted, in revised form, in Nichols, Smith and Miller (2009, 339–348), doi:10.2307/2953779.

—. 2015a. "What Does it Mean to Say That We Are Animals?" *Journal of Consciousness Studies* 22(11–12): 84–107.

—. 2015b. "Animalism and the Remnant-Person Problem." in *Philosophical Perspectives on the Self*, edited by João Fonseca and Jorge Gonçalves, pp. 21–40. Bern: Peter Lang Publishing Co.

—. 2016. "The Remnant-Person Problem." in *Animalism: New Essays on Persons, Animals, and Identity*, edited by Stephan Blatti and Paul F. Snowdon, pp. 145–161. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199608751.001.0001.

Parfit, Derek. 2012. "We Are Not Human Beings." *Philosophy* 87(339): 5–28, doi:10.1017/s0031819111000520.

Robinson, Denis. 1985. "Can Amoebae Divide without Multiplying?" *Australasian Journal of Philosophy* 63(3): 299–319, doi:10.1080/00048408512341901.

Sauchelli, Andrea. 2017. "The Animal, the Corpse, and the Remnant-Person." *Philosophical Studies* 174(1): 205–218, doi:10.1007/s11098-016-0677-4.

Sharpe, Kevin W. 2015. "Animalism and Person Essentialism." *Metaphysica* 16(1): 53–72, doi:10.1515/mp-2015-0003.

van Inwagen, Peter. 1990. *Material Beings*. Ithaca, New York: Cornell University Press.

Wiggins, David. 1996. "Replies." in *Essays for David Wiggins – Identity, Truth and Value*, edited by Sabina Lovibond and Stephen G. Williams, pp. 219–230. Aristotelian Society Series n. 16. Oxford: Basil Blackwell Publishers.

The Attitudinalist Challenge to Perceptualism about Emotion

Michael Milona

Perceptualists maintain that emotions essentially involve perceptual experiences of value. This view pressures advocates to individuate emotion types (e.g., anger, fear) by their respective evaluative contents. This paper explores the Attitudinalist Challenge to perceptualism. According to the challenge, everyday ways of talking and thinking about emotions conflict with the thesis that emotions are individuated by, or even have, evaluative content; the attitudinalist proposes instead that emotions are evaluative at the level of attitude. Faced with this challenge, perceptualists should deepen their analogy with sensory experience; they should distinguish types of emotions by their content much as we can plausibly distinguish types of sensory experience (e.g., visual, auditory) by theirs. A second lesson is that perceptualists should distinguish an emotion's representational guise (uniform across emotions) from its formal object (which varies).

Perceptualists maintain that emotions essentially involve perceptual experiences of value. On this approach, anger might be thought to involve an experience of offense, pride an experience of one's own achievement, and so on. The perceptual approach has enjoyed significant support in emotion theory (Roberts 2013; Tappolet 2016, *inter alia*). Theorists have also relied on it in value epistemology (Milona 2016), action theory (Döring 2007), and normative ethics (Stockdale 2017). To be sure, perceptualist theories vary in the details, including important ways that I canvass below. But despite such differences, perceptualists are unified in taking emotions to have evaluative content in much the way that visual, auditory, etc. experiences have empirical content.

At first glance, perceptualism looks like a promising starting point for analyzing emotions. Many philosophers today maintain that emotions are not (mere) bodily sensations; they are *evaluations*. It was once popular to treat

these evaluations as forms of judgment (Solomon 1976; Nussbaum 2004). But many have since migrated from *judgmentalism*, as it is often called, to perceptualism. One major reason for this trend is simple. When we are over-come with fear, to take a familiar example, we sometimes explicitly judge that what we fear isn't dangerous. But such cases are not experienced as similar to making contradictory judgments (see D'Arms and Jacobson 2003; Naar 2020). They instead seem more akin to perceptual illusions, whereby things appear other than we believe them to be (Tappolet 2016). So if we accept that emotions are evaluations, then a perceptual model looks like a promising starting point.

However, Julien Deonna and Fabrice Teroni (2012, 2015) forcefully argue that perceptualism isn't a great starting point after all. While they agree that emotions are evaluative experiences, they maintain that perceptualism goes awry in treating all emotions as being the same type of attitude. This leads per-ceptualists to distinguish emotion types by virtue of their supposedly differing evaluative content. But, Deonna and Teroni argue, there are several ways in which ordinary, pretheoretical ways of talking and thinking about emotions conflict with emotions being distinguished by their evaluative content, or even having such content at all. I refer to these objections as the *Attitudinalist Objections*, or jointly as the *Attitudinalist Challenge*. They maintain instead that the evaluative dimension of an emotion is a feature of the attitude rather than its content; and because this evaluative dimension is different for each emotion, each emotion is a different type of attitude. Their theory is thus a version of *attitudinalism*, according to which emotions are evaluative attitudes but do not have evaluative content.[1]

This paper defends perceptualism in the face of the Attitudinalist Chal-lenge.[2] I argue that the objections either rely on subtle mistakes about what

---

1 Deonna and Teroni's view is the most widely discussed version of attitudinalism. There are important alternatives, however (e.g., Müller 2017). See section 2 and section 6 below.

2 Perceptualism faces numerous other objections. My own view is that they can be answered, though I haven't space to do so here. For example, Deonna and Teroni argue that emotions having cognitive bases makes them importantly different from perceptual experiences (2012, 69). Perceptualist responses include Tappolet (2016, 24–31) and Milona and Naar (2020). Some likewise argue that the phenomenology of emotions is importantly different from perceptual experiences (Salmela 2011; Dokic and Lemaire 2013; Deonna and Teroni 2012, 68–69; Müller 2019). See Roberts (2013, 71–72) and Poellner (2016) for potential responses. More recently, Naar (2022) argues that emotions are better understood on the model of action than on that of perception. For more detail on various debates about perceptualism, see especially Brady (2013) and Döring and Lutz (2015).

perceptualism says, or else turn on optional commitments that perceptualists can avoid on independent grounds. Having argued how perceptualists should answer the Attitudinalist Challenge, the paper closes by issuing a challenge of its own for versions of attitudinalism that share perceptualism's commitment to the view that emotions are evaluative experiences.

1 What Perceptualism Is

Versions of perceptualism have been defended by Cooper (1699), Meinong (1972), Roberts (2013), and Tappolet (2016), among others. The basic view is as follows:

> PERCEPTUALISM. Emotional experiences essentially involve non-doxastic, affective representations of value.

Sabine Döring offers an intuitive illustration with reference to the emotion of indignation:

> In experiencing indignation at the harsh punishment of the toddler, it seems to you that the punishment is in fact unjust: your occurrent emotional state puts forward your indignation's content as correct. This is in analogy to the content of a sense perception. In perceiving that the cat is on the mat, it seems to you that the cat is actually there. (Döring 2007, 377)

Several aspects of perceptualism require clarification. First, the theory speaks of "emotional experiences" because it is only meant as an analysis of occurrent, conscious emotions. For example, while it may be true that Cassandra loves Sasha even when Cassandra is sleeping, "love" here is meant dispositionally. Perceptualism is not about emotions in this sense.

Second, the phrase "essentially involve" is non-committal about a key question, namely whether there are any necessary components of emotion other than non-doxastic, affective representations of value. We can thus distinguish between the following positions:

> PARTHOOD PERCEPTUALISM. Emotional experiences essentially involve non-doxastic, affective representations of value as a proper part.

Identity Perceptualism.   Emotional experiences are nothing
more than non-doxastic, affective representations of value.

To illustrate these two positions, consider efforts to analyze emotions commonly begin by listing paradigmatic features of emotional experiences. These include evaluations, bodily feelings, action tendencies, and patterns of attention, among other things (Brady 2019, 10). For example, a hiker who fears a nearby bear can be expected to evaluate the bear as dangerous, experience sensations characteristic of fear, be motivated to avert the threat, and attend to whether the bear really is dangerous and what the escape options might be. Identity perceptualists maintain that emotions are in essence their evaluative dimension, which they take to be a non-doxastic representation of value. By contrast, parthood perceptualists see the evaluative dimension as insufficient on its own. Perhaps, for example, it must be paired with a tendency to act in accordance with that representation. In the case of fear, for example, this might be a tendency to act so as to avoid what is experienced as dangerous.

Perceptualism's advocates are almost always identity perceptualists.[3] This may come as a surprise, given that perceptualism's close cousin, judgmentalism, does divide into two distinct camps. That is, there are some who think that judgment fully captures the nature of emotion (e.g., Nussbaum 2004), and others who think it must be supplemented (e.g., Green 1992).[4] As it happens, many of the motivations for perceptualism, including that it can provide a plausible basis for value epistemology and that it can explain how emotions rationalize action, only require parthood perceptualism. Furthermore, perceptualists who are willing to take seriously parthood perceptualism have additional resources for addressing the Attitudinalist Challenge. For if emotions include more than evaluative representations, then they may be distinguished not only by their evaluative content, but also by other features (e.g.,

---

3 I'm not aware of any philosophers who explicitly defend parthood perceptualism (though some leave open the possibility, e.g., Cowan 2016, 61–62; Milona 2016; Mitchell 2017). We may find inspiration, however, in the work of some appraisal theorists in psychology. For example, Richard Lazarus says the following: "[E]motion is a superordinate concept that includes cognition, which is its cause in a part-whole sense. Cognitive activity, A, about the significance of the person's beneficial or harmful relationships with the environment, is combined in an emotion with physiological reactions and action-tendencies, B, to form a complex emotional configuration, AB" (Lazarus 1991, 353–354). According to Lazarus, the role of appraisal (cognition) in emotion is analogous to that of germs in the production of a disease, being both a cause and a part.

4 Although Green maintains that beliefs are essential, these beliefs aren't always evaluative (1992, 78). For discussion of different forms of judgmentalism, see Naar (2019).

action-tendencies).[5] However, because identity perceptualism is the dominant version of the theory, and because the Attitudinalist Challenge is most serious for this version, I focus in what follows on identity perceptualism.

Other key questions for perceptualists concern the relationship between an emotion's purported evaluative content and its phenomenology. Perceptualists typically view an emotion's representation of value as inseparable from its affective (felt) dimension. Here is Roberts:

> Affect is not something in addition to emotion [...] Just as in the visual experience of a house one is appeared to in the way characteristic of house-sightings, so in fear one is appeared to (in feeling) in the way characteristic of threat confrontations (the threat being directed at something one cares about). (Roberts 2013, 47–48)

Others make similar claims about the inseparability of emotional affect/feeling and the representational dimension of emotion (e.g., Döring 2007, 374; Tappolet 2016, 27–28; see also Ballard 2021b, 121).[6] According to this position, to describe *what it is like* to have an emotional experience requires reference to value (see Poellner 2016, 270). In experiencing, say, anger, we cannot describe its phenomenology without reference to the property of being wronged. I take perceptualists to be committed to this *inseparability* of emotional phenomenology and value. Such a position is compatible with different views about the relationship between how an emotion feels and what it represents. For example, on one possible view, the affective aspect of an emotion (or at least part of it) grounds the evaluative representation. This would accord with an increasingly popular approach to perceptual content which grounds such content in the phenomenal character of perceptual experience (see Kriegel

---

5 It is an open question for parthood perceptualists whether these additional features are representational. For example, suppose a parthood perceptualist invokes action-tendencies as the additional feature. On one conception, these action-tendencies are non-representational feelings of one's body's readiness to act (Deonna and Teroni 2012). By contrast, Mitchell (2021) proposes an intriguing "object-side" model of action-readiness (or action-tendency) phenomenology. This is an experience of an object (e.g., a charging bear or a beautiful painting) as calling for, or demanding, action. As Mitchell points out, object-based action-readiness is plausibly representational.

6 This doesn't mean that *all* of the feelings that we typically associate with emotions are inseparable from the representation of value. In particular, the bodily feelings that typically come along with emotions are naturally treated by perceptualists as representing bodily changes rather than value and thus as ultimately non-essential for emotion, at least according to identity perceptualism (cf. Nussbaum 2004, 328–329).

2013). But here I am non-committal about whether the intentionality or phe-
nomenology of emotions is more basic (if either is).[7]

Additional details about how perceptualists should, or at least reasonably
can, develop their view will emerge in the course of addressing the Attitu-
dinalist Challenge. In particular, I suggest that perceptualists take up more
specific views about the affective, non-doxastic representation of value and
how it relates to ordinary sensory experience.

2 The Attitudinal Alternative

The *attitudinal theory* is an important alternative to perceptualism. While
attitudinalists agree that emotions are evaluations, they deny that emotions
have evaluative content (e.g., Deonna and Teroni 2012, 2015; Müller 2017).
Emotions are taken to be evaluative at the level of *attitude*.

The basic idea can be illustrated by way of a comparison with belief and
truth. A belief that *P* has *P* as its content. But there's more to a belief than its
content. After all, one can also *suppose* that *P*. One major difference between
a belief and a supposition with the same content is that the former is in some
sense *truth-directed*. However, a belief that *P* doesn't *represent* that *P* is true,
for a belief that *P* has different content than a belief that *P is true* (see Kriegel
(2019b), 10; Ballard (2021a), 852–853). So truth somehow characterizes the
very attitude of belief. That is, a belief is a way of *taking-as-true* some content.
According to Deonna and Teroni, matters are similar with emotion, except
that values, rather than truth, characterize emotional attitudes. So instead of
saying, for instance, that fear represents the property of being dangerous and
anger represents the property of being offensive, the attitudinalist says that
the attitude of fear is a way of *taking-as-dangerous* its content and that the
attitude of anger is a way of *taking-as-offensive* its content.

But what is it to take-as-dangerous or take-as-offensive? On the most widely
discussed version of attitudinalism, we find another similarity with perceptu-
alism: emotional experiences are a way of experiencing value (Deonna and

---

7 Prinz (2007) draws on a Dretske-style indicator semantics in arguing that emotions have evalu-
ative content. According to him, emotions involve representations of value insofar as they are
perceptions of bodily changes and these bodily changes have the function of tracking correspond-
ing values. For the sake of simplicity, I set this version of perceptualism aside (cf. Cowan 2016,
78, n8).

Teroni 2012, 2015).[8] And as with perceptualism, when all goes well, these are experiences through which we come to apprehend objects as having certain values. Deonna and Teroni describe these experiences in terms of the form of readiness to act involved in each emotion (cf. Frijda 2007). Here are two helpful illustrations:

> Fear of a dog is an experience of the dog as dangerous insofar as it is an experience of one's body being prepared to forestall its impact (flight, preventive attack, immobility, etc.), an attitude it is correct to have if, and only if, the dog is dangerous. In the same way, anger at a person is an experience of offensiveness insofar as it consists in an experience of one's body being prepared to retaliate, an attitude that is correct if, and only if, the person is offensive. (Deonna and Teroni 2015, 303; see also Deonna and Teroni 2012, 81)

Since Deonna and Teroni's theory explains the sense in which emotions are evaluative experiences by appealing to such action-tendencies, I refer to this as *action-tendency attitudinalism*.[9] By maintaining that emotions are ways of experiencing value, one might suppose that action-tendency attitudinalists can thereby secure many of the advantages (or at least ambitions) of perceptualism in value epistemology and action-theory. I briefly address these matters in the penultimate section.

It is important to note that while Deonna and Teroni's action-tendency attitudinalism is often treated as the representative version of attitudinalism (e.g., Rossi and Tappolet 2019; Ballard 2021a), the theory can take different forms. Attitudinalism as such merely claims that emotions are evaluative at the level of attitude rather than content. Thus an attitudinalist might agree with Deonna and Teroni that emotions are evaluative experiences but resist the idea that this has to do with experiences of action-readiness (cf. Kriegel 2019b, 13). I consider below (section 6) why an attitudinalist might favor such an

---

8 Müller (2017) similarly describes Deonna and Teroni as maintaining with perceptualists that emotions *apprehend* value, or at least apparently apprehend value. The term "experience" here is intended to be non-factive, covering both genuine experiences of value and mere experiences as of value.

9 One may further qualify that Deonna and Teroni's theory as *bodily* action-tendency attitudinalism. For as noted above, one may also attempt to capture the phenomenology of preparedness to act in non-bodily, representational terms (see footnote 5 and Mitchell 2021). For ease of presentation, though, I don't add this qualification throughout.

alternative characterization of emotions as evaluative attitudes. Furthermore, it is also consistent with attitudinalism to maintain that emotions aren't ways of experiencing value at all. For example, Müller (2017) argues that emotions are responses to pre-emotional experiences of value rather than experiences of value themselves; and these responses are such as to be correct in the presence of the relevant value.[10] So attitudinalism is *highly* flexible. To keep things manageable, however, I limit my discussion to versions of attitudinalism that take emotions to be evaluative experiences and likewise focus the ensuing discussion primarily on action-tendency attitudinalism.

3 The First Attitudinalist Objection: Perceptualism as a Bad Start

The first Attitudinalist Objection is simple, at least in outline. It emerges from similarities between how we pretheoretically conceptualize different

---

10 Müller offers multiple arguments against the view that emotions apprehend value. For example, one key argument starts with the thought that we often ask people *why* they are angry, sad, etc. in order to probe their motivating reasons for being angry, sad, etc. But Müller maintains that it doesn't make sense to ask similar questions about why someone apprehends something, and this therefore indicates emotions aren't apprehensions (Müller 2017, 286; see also Dietz 2018; and Mulligan 2010, 485). From a perceptualist perspective, this argument is structurally similar to the familiar argument that emotions admit of justificatory reasons while perceptual experiences don't (Deonna and Teroni 2012; Brady 2013). In both cases, the perceptualist's most straightforward response is to resist the view that emotions admit of either kind of reason. So, for example, while someone might say in some instance that they are angry for no reason (Dietz 2018, 1689), the perceptualist may say that, strictly speaking, one is *always* angry for no reason. We just tend to say what isn't quite true. But conformity to all pretheoretical ways of talking isn't decisive, as others who press a challenge similar to Müller's observe (Dietz 2018, 1690). Perceptualists, furthermore, have resources to explain our tendency to talk about motivating (or justificatory) reasons for emotions. For example, emotions are highly sensitive to choices and attitudes (e.g., beliefs) that *do* admit of such reasons. And so we can be motivated to bring it about that we experience certain emotions, or we can be (ir)rational in bringing about certain emotions (cf. Milona 2016, 903; Tappolet 2016, 37–38). Thus while there may be *a* cost for perceptualism here, it arguably isn't severe. [See Milona, manuscript, for an extended, and less concessive, response to these worries about motivating and justificatory reasons.] Action-tendency attitudinalists could follow a similar path. A complication, however, is that advocates of this view have objected to perceptualism precisely on the grounds that it fails to accommodate justificatory reasons for emotion (Deonna and Teroni 2012). And it isn't clear that justificatory reasons for apprehensions make any more sense than justificatory reasons for perceptions. So as Müller points out, Deonna and Teroni's own proposal "can be attacked on the same grounds on which they attack the Perceptual View" (Müller 2017, 286).

emotions as compared to attitudes such as belief, desire, perception, etc. Here is how Deonna and Teroni put it:

> [R]egarding the different types of emotions as different attitudes and not as one and the same attitude—for example the attitude of judging or that of perceiving—towards different contents is the default position […] Isn't it natural to understand the contrast between, say, fear, anger and joy as one between different ways the mind is concerned with objects and events? Shouldn't this contrast be located at the same level as that between desiring, believing and conjecturing and be clearly distinguished from the contrast between believing a given proposition and believing another? (Deonna and Teroni 2015, 296)

The argument can be summed up as follows. The first premise is that when we talk about believing, desiring, perceiving, etc., we are talking about different attitudes. The second premise is that if the foregoing premise is true, then by analogy, when we talk about emotions, including fear, envy, and so on, it is natural to assume that we are also talking about different attitudes. But, the argument continues, perceptualism denies that emotions are distinct attitudes. For according to perceptualism, all emotions are constituted by the same affective attitude. Call this the *Perceptualism as a Bad Start Objection*. It is easy to see why Deonna and Teroni, building on this objection, maintain that attitudinalism, rather than perceptualism, should be our starting point for theorizing the sense in which emotions are evaluations.

Before considering how the perceptualist might reply, we should consider what it is for something to be an attitude. There are different ways in which one might define such a technical (or quasi-technical) term. But as the passage from Deonna and Teroni above illustrates, they intend for the purposes of this objection a sense of "attitude" inclusive of perceiving (Deonna and Teroni 2015, 296; see also Kriegel 2019a). This makes sense given the present dialectic. The objection isn't that perceptualists fail to treat emotions as attitudes; it's that they treat all of them as the *same* attitude, distinguished only by their contents.[11] Furthermore, defenders of perceptualism have recently been explicit

---

11 A referee rightly points out that other approaches to defining "attitude" may create trouble for perceptualism. But it's important to notice that these issues are distinct from the objection being considered here. For example, one might define "attitude" in terms of "taking a position" on something. Deonna and Teroni elsewhere gesture towards such a proposal in developing their version of attitudinalism (though not in pressing the Perceptualism as a Bad Start Objection).

that they don't mean to deny that emotions are attitudes (Rossi and Tappolet 2019, 553). I thus suspect that Deonna and Teroni have in mind a capacious view of attitudes whereby an attitude is "a way of having content" (Siegel 2021). A perceptualist would certainly grant that emotions are attitudes in this sense.

In addressing the Perceptualism as a Bad Start Objection, I focus in particular on *experiential* ways of having content. By this I mean to refer to ways of having content such that there is *something it is like* to represent in that way.[12] By focusing on experiential ways of having content, perceptualists can ensure that their response hews close to the surface of our emotional life and so doesn't lose sight of the intuition driving the objection. I therefore won't be concerned with sub-personal ways of representing, or with sub-personal processes that give rise to experiences with certain content (cf. Siegel 2021; Kriegel 2019a). To illustrate, suppose that a perceptualist attempts to address the challenge by appealing to distinct neural machinery underlying different emotions (see Tracy and Randles 2011). The various processes by which different emotions arise may lead a perceptualist to say that there are many different emotional attitudes insofar as they involve the functioning of distinct

---

For example, they say "we should conceive of emotions as distinctive types of bodily awareness, where the subject experiences her body holistically as taking an attitude towards a certain object [...]" (Deonna and Teroni 2012, 79). Insofar as it is just the body, and not the agent, that is experienced as taking a stance, this proposal may not conflict with the idea that perceptual experience qualifies as an attitude. But Deonna and Teroni seem to also have in mind that the agent moreover experiences *themselves* as taking a stance (2012, 79–81). Perceptual experiences don't seem to qualify as attitudes in *this* sense. But perceptualists would (or at least in my view, should) resist that emotions are this sort of attitude. Here I think that they are on solid footing phenomenologically: emotions (passions) seem to be passive in a way that is difficult to describe in terms of the (emoting) agent's taking a stand on the world (but see Müller 2019). Yet the idea that emotions are attitudes in this sense may persist in light of the fact that emotions seem to admit of reasons (motivating and normative). See footnote 10 and citations therein for details about how perceptualists might answer these concerns about emotions and reasons.

12 A perceptualist might argue that there is no distinctive attitudinal phenomenology, maintaining instead that the phenomenology of emotion is tied *entirely* to content. This would mirror a familiar approach to sensory experience (see Tye 1995). But on the basis of considerations outlined below (section 4), I think of the phenomenology of perceptual experience as corresponding to both attitude and content. Of course, such phenomenological considerations are contestable. But here it's worth noticing two additional points. First, we've already seen (section 2) a reason to think *belief* admits of an attitude/content distinction; and this gives defeasible reason to think other attitudes work similarly (see also section 5 below). Second, the present version of perceptualism shares attitudinalism's commitment to the thesis that an emotion's correctness conditions are a function of attitude and content, and so it helps to focus our attention on the real points of disagreement between the two approaches to theorizing emotion.

biological capacities. But it seems to me that Deonna and Teroni's objection doesn't hinge on the underlying neural architecture of emotion but is rather focused on the surface of how we pretheoretically talk and think about emotions.[13] By focusing on attitudes as experiential ways of having content, then, we mitigate the risk of missing the point.

How, then, should a perceptualist respond to the objection? The most straightforward reply is already suggested by the core of perceptualism, namely its analogy with sensory experiences. To see why, recall that the objection invites us to have the intuition that just as perception, belief, and desire are all distinct attitudes, so too are the various emotions, including joy, anger, sadness, etc. But there are alternative comparisons that, from a pretheoretical perspective, we might just as easily have made. More specifically, we might have compared emotional experiences and experiences in different sensory modalities, including visual, auditory, tactile, etc. experiences. Here again, the focus is on the sensory experiences themselves, rather than the underlying sub-personal processes.[14] And here too we can ask what makes an experience in one modality experientially, or phenomenologically, distinct from an experience in another modality. One salient difference, of course, concerns the contents of experiences in different modalities. For example, a visual experience has colors as part of its content while an auditory experience has sounds (even if some of the content of an auditory and visual experience overlap). Indeed, perhaps *all* of the experiential differences between visual, auditory, etc. experiences are a function of content (Speaks 2015, ch. 24–26; see also Chalmers 2004). But if it were reasonable to maintain that talk of visual, auditory, etc. experiences refers to a single experiential way of having content that is uniform across different sensory experiences, then presumably it is likewise reasonable, for all we've seen, for perceptualists to maintain that talk of anger, sadness, etc. refers to a single attitude that is uniform across different emotions. If this were correct, then the Perceptualism as a Bad Start Objection would fail to gain independent leverage insofar as it stacks the deck by inviting a tendentious comparison between emotions (emotional experiences) and perception, belief, desire, etc. rather than visual, auditory, tactile, etc. experiences.

---

13   Thanks to a referee for helpful feedback on this issue.
14   See Grice (1962) on different ways of talking about sensory modalities.

But is it plausible that different sensory modalities involve a single experiential way of having content?[15] One important argument for an affirmative answer builds on the phenomenon of *perceptual binding*.[16] To illustrate, suppose a person sees a basketball as orange and spherical. They don't just simultaneously see something orange and something spherical but rather experience a single entity as orange and spherical. This is *intramodal* perceptual binding. Such binding can also occur *intermodally*. For example, one may perceptually experience a brown dog as barking (Speaks 2015, 180). This isn't merely the co-occurrence of a visual experience as of a brown dog at a certain location and an auditory experience as of barking nearby. The brown and the barking are experienced as having a common source. But since the sound (barking) isn't seen and the color (brown) isn't heard, this experience seems to be intermodal in character. Following Speaks, let's call this intermodal experience a *C-representation* (2015, 183–184).

Consider now the question of whether in C-representing the dog as brown and barking one likewise C-represents the dog as brown and C-represents the dog as barking. There is pressure to say yes. To see this, consider how other attitudes work. For example, if one believes that the dog is brown and barking, then one believes that the dog is brown and believes that it is barking. Or returning to the example of intramodal binding, in seeing the basketball as orange and spherical, one sees the basketball as orange and sees it as spherical. Barring a persuasive argument to the contrary, we should likewise say that C-representations distribute over conjunction in just the same ways as believing and seeing. But now it looks like C-representations are, as Speaks puts it, "swallowing up the other species of perceptual representation" (2015, 184). Rather than insisting C-representations occur alongside visual, auditory, etc. experiences with the same content, Speaks suggests that there is a single experiential way of having content common to each.[17] In other words,

---

15 One might think that experiences in different sensory modalities *must* be different ways of representing. After all, a visual and auditory experience might be about the same thing even while their phenomenology differs (see Block 1996). This difference in phenomenology, one might think, must be explained by a difference in the way that visual and auditory experiences represent. This parallels one of the Attitudinalist Objections against perceptualism and is addressed below (see section 4).

16 The argumentation in the next two paragraphs follows Speaks (2015, 177–185). For related arguments, see Tye (2007) and Bourget (2017). For an overview of the phenomenon of perceptual binding, see O'Callaghan (2015).

17 According to Speaks, this way of having content isn't limited to the five senses (2015, 186–188). He thinks that it also applies to bodily sensations.

visual, auditory, etc. experiences aren't each distinctive attitudes in their own right; they rather qualify a singular perceptual attitude. While I cannot fully investigate the prospects for this view here (though section 4 addresses an important objection that parallels another of the Attitudinalist Objections), it offers an attractive framework in which to develop perceptualism.

Faced with the Perceptualism as a Bad Start Objection, then, perceptualists should say that just as different sensory experiences involve the same underlying attitude, so too do emotions. Moreover, on a straightforward version of perceptualism, the experiential way of having content implicated in emotions is the same as that involved in sensory perceptual experience. Such a view pairs naturally with the standard perceptualist idea that describing *what it is like* to have an emotional experience requires reference to value just as describing what it is like to have a sensory experience requires reference to what the sensory experience is about. I referred to this perceptualist idea above as the *inseparability* of emotional phenomenology and value.[18] According to the present proposal, this similarity between emotional and sensory experience is explained by the fact that the experiential way of having content is the same in each case.[19] Of course, this doesn't mean that there won't be differences.

---

18 This idea is commonly endorsed by perceptualists. Indeed, for some perceptualists (including myself), this thought is part of what makes perceptualism so attractive in the first place. (See Döring 2007, 374; Roberts 2013, 71–72; Tappolet 2016, 27–28; Milona 2016; Poellner 2016, 270.) But others may think that this proposal is phenomenologically implausible. For example, Demian Whiting maintains that emotional experiences/feelings "do *not* manifest phenomenally a representational character or content" (Whiting 2012, 97). According to him, while (say) nervousness involves a " 'restless' or 'nervous' sensation" and fear an "unpleasant edgy sensation", it is important to notice that "these feelings—the *only* feelings manifest in the emotions—do not have the representational properties that the perceptual value theorist is after" (Whiting 2012, 101). But while I can't respond to Whiting in full, it seems to me that he hasn't offered a compelling case. Talk of restlessness, edginess, etc. strikes me as referring to a combination of bodily and evaluative representations. On this proposal, the "edginess" in fear might be understood as the combination of an evaluative experience of danger with a bodily experience of readiness to flee (or fight) in light of that danger. One attraction of this approach is that it can explain (what seems to me possible) why fear sometimes lacks edginess. Imagine a person who, while afraid of losing their job, recognizes that there is nothing they can do right now and so lacks fear's bodily manifestations and thereby any "edgy" phenomenology.

19 Given this account of the phenomenology of emotions, one can further buttress the thesis that emotions and sensory experiences involve the same attitude by appealing to cases of intermodal binding similar to the ones Speaks invokes in his argument. For example, a person who fears a snarling dog, according to the perceptualist, experiences the snarling dog as dangerous. This isn't merely the copresence of a visual experience of a snarling dog and an affective experience of danger. Rather, the snarling dog is experienced as the *source* of danger (much as the brown dog is experienced as the source of the barking in the example above). But one doesn't affectively

For just as experiencing an odor is very different from experiencing a sound, so too is experiencing value very different from experiencing either. But the perceptualist position is that these are differences in content rather than differences in experiential ways of having content. In sections 4 and 5 below I'll have more to say on how perceptualists can theorize about these ways of having content.

Before moving on, it's worth noticing that perceptualism's fate isn't necessarily beholden to the view that emotions involve the same way of representing as ordinary sensory experience. Nevertheless, if a perceptualist doesn't follow this path, it raises concerns about whether they will ultimately have an adequate response to the Perceptualism as a Bad Start Objection. Such a perceptualist has two options. On the one hand, they may say that talk of different emotions refers to a single attitude of "emoting." But then perceptualists would face the burden of saying what such emoting consists in, including how it is distinct from the attitude implicated in ordinary sensory-perceptual experience (Deonna and Teroni 2012, 78). On the other hand, a perceptualist could observe that perceptualism is compatible with taking different types of emotions to be distinct attitudes. That is, a perceptualist may argue that fear is a *fearful* representation of something as dangerous, anger an *angry* representation of something as offensive, and so on. The difficulty here is that it isn't clear what an angry or fearful way of having content is. Analyzing them in terms of their corresponding values may seem objectionably redundant, given that those values are already in the content. And taking them to be primitive ways of representing strikes me as theoretically disappointing, best reserved as a last resort. Taking seriously the Perceptualism as a Bad Start Objection thus pressures perceptualists to maintain that emotions and sensory experiences involve the same experiential way of representing.

---

represent the snarling dog or visually experience the danger. So to avoid the problems of invoking an additional intermodal attitude (similar to Speaks's C-representation hypothesis), it is better to understand "affective" and "visual" to qualify the contents of a singular perceptual way of representing.

4 The Second and Third Attitudinalist Objections: Portable Contents and Fading Emotions

4.1 *Unpacking the Objections*

I turn now to the second dimension of the Attitudinalist Challenge, which consists of two related arguments. Answering these objections reveals hitherto underappreciated points of disagreement between attitudinalism and perceptualism. This will take some work to see, however, since Deonna and Teroni's arguments may initially appear question-begging.

To begin, Deonna and Teroni (2015, 297) observe that we often talk as if distinct emotions are about the same thing. For example, we might say that one person is angry about something that another finds amusing. But perceptualism denies this insofar as it ascribes different content to anger than it does to amusement. Put generally, the objection is as follows. The first premise is that different types of emotion can be about the same thing. The second premise is that if instances of different emotion types can be about the same thing, then emotions as such do not contribute anything to what is represented. But then this is a problem for perceptualism, since perceptualism says that each emotion is tied to a corresponding value that it represents. In other words, perceptualism is committed to the following claim that the attitudinalist rejects: the full content of one emotion type (anger) is *never* entirely portable to another emotion type (e.g., amusement). Call this the *Portable Contents Objection*.

Deonna and Teroni offer what they take to be a similar argument using an example involving a single emotion. Here is what they say:

> Maurice is not amused anymore by Barbara's excellent joke for he heard it a hundred times. This is because his attitude towards the joke has changed, not because of a change in the content of the joke. We expect Maurice to insist that the joke is very funny while stressing the fact that at that point he heard it too many times (Herzberg 2012, 81). We have no apparent reason to think that these everyday situations imply a difference in *what* the subject's mind is concerned with as opposed to the *way* his mind is concerned with it. (Deonna and Teroni 2015, 297)

This example involving a single emotion is meant to illustrate that emotions can come and go without changing what one represents. Maurice continues

to represent the joke as funny—presumably by way of a belief—even as his amusement fades. Although Deonna and Teroni group this objection with the Portable Contents Objection, it will, for reasons that become clear below, be worth keeping separate. I call this the *Fading Emotions Objection*.

These objections may appear question-begging. As Mauro Rossi and Christine Tappolet point out in their defense of perceptualism, we must not conflate what they call the *intentional object* of an emotion with its *entire content* (Rossi and Tappolet 2019, 552). The intentional object of, say, Maurice's amusement at Barbara's joke is the joke itself. But then the perceptualist adds to this a story about what amusement is, namely an experience of its object as amusing. So for Deonna and Teroni to insist that different emotions can have the same content is to beg the question. And Rossi and Tappolet could add that in cases where amusement fades (though they don't address cases of this sort directly), we must not simply assume that nothing changes about what the agent represents. The perceptualist will say that even if the agent continues to believe that the joke is funny once the amusement has faded, they no longer *emotionally experience* it as such. In other words, what they once represented in two ways, namely through judgment and emotion, they subsequently only represent in one.

It turns out, however, that the Portable Contents Objection (and similarly the Fading Emotions Objection) can be further developed in a way that isn't question-begging. One possibility, suggested by Rossi and Tappolet (2019), is that the objection may proceed from general commitments about the nature of *formal objects*, and the formal objects of emotions in particular (see also Deonna and Teroni 2012, 76). Formal objects are distinguished from *intentional objects*, or *particular objects* (see Kenny 1963; Teroni 2007, 396). In general, formal objects "are supposed to shed light on specific categories of mental states" (Teroni 2007, 396). For example, the intentional object of a belief that *P* is *P*, but the formal object, at least according to one common view, is truth. Whereas *P* can figure in the content of many different mental states (e.g., one can suppose that *P*), the formal object, truth, seems to tell us something important about the nature of belief itself. Similarly, according to a familiar story about emotions, the formal objects of emotions are the values corresponding to each emotion. Fear of a bear, say, has two objects: the intentional object is the bear and the formal object is danger. Such formal objects perform at least two main tasks (Rossi and Tappolet 2019, 549). First, they help to determine an emotion's correctness conditions. Fearing that *P* is correct just in case *P* is dangerous. Second, the formal object individuates the type of emotion in

question. For example, anger is distinct from fear because these emotions have distinct formal objects.[20] Rossi and Tappolet then point out that, according to Deonna and Teroni, the formal object of an attitude is never part of its content (Deonna and Teroni 2012, 76). This picture of formal objects, then, denies the perceptualist any gap between the intentional object of emotions and the "entire" content of emotions. So when we say, for example, that one person is angry about what another person finds amusing, perceptualism can't make sense of this. And so, the thought goes, attitudinalism is a better starting point for emotion theory.

Rossi and Tappolet offer a reply on behalf of perceptualism. Their reply begins by conceding that the formal object of *many* mental states resides outside those states' content. For example, a belief that *P* doesn't represent that *P* is true. It's rather that truth characterizes the correctness conditions for the attitude type rather than its content (Rossi and Tappolet 2019, 555). But according to them, the formal objects of *some* non-emotional attitudes *do* feature in those attitudes' content. Here they point to chromatic perceptual experiences. These include visual experiences of red, green, etc. Take a visual experience of an object as red. This experience has redness as part of its content. But if formal objects individuate attitudes and determine their correctness conditions, then redness is likewise the formal object. For as Rossi and Tappolet observe, "redness is that which, in conjunction with the intentional object of a perception of red, determines whether the perception is correct or not" (2019, 551). And "redness is the property that individuates the type of perception in question, namely, a perception of red" (Rossi and Tappolet 2019, 551).

As it stands, advocates of the Portable Contents and Fading Emotions Objections are unlikely to find Rossi and Tappolet's defense of perceptualism persuasive, and reasonably enough. This is because the notion of formal

---

20 According to some, formal objects play a third role, namely that of serving as a constraint on an emotion's intelligibility. As Müller puts it, "this intelligibility constraint specifies how the subject of an attitude must construe its intentional content in order for her to intelligibly hold that attitude" (Müller 2017, 287). This may not seem to be a problem for perceptualists, since they agree that experiencing an emotion of a given type requires a "construal" (perceptual experience) in terms of the formal object. But according to Müller, the best way to interpret this constraint requires us to invoke *pre-emotional* apprehensions of value. But then this suggests that emotions are responses to apprehensions of value (or experiences of value) rather than apprehensions of value themselves. If Müller is right, then this is a problem not only for perceptualism but also Deonna and Teroni's brand of perceptualism. For the purposes of this paper, I set aside these broader concerns about whether emotions are experiences of value at all.

objects has arguably been cheapened to the point that they are no longer
revelatory of the attitude or mental state in question (cf. Teroni 2007, 396;
Müller 2017, 284). To illustrate, suppose that chromatic perceptual experi-
ences, including "reddish" visual experiences, "bluish" visual experiences,
etc., mark distinctive attitudes with their own formal objects. One may worry
that, if this were the case, then there are as many distinctive attitudes and
formal objects as there are properties that can be perceived. This includes
not only colors such as red but specific shades of red, specific shapes, motion
properties, etc. And beyond perceptual experience, if groupings of similar
contents are viewed as sufficient grounds for invoking distinctive attitudes and
corresponding formal objects, it isn't clear why this line of response wouldn't
generate the result that, say, chromatic *beliefs* also have colors as their formal
objects (perhaps in addition to truth). So Deonna and Teroni can reasonably
deny that adding qualifications such as "chromatic" (or "shaped," etc.) to "per-
ceptual experience" and "belief" marks a new attitude with its own formal
object.

   As we'll see momentarily, Rossi and Tappolet's reply gets something impor-
tantly right. Perceptualists should take the relation between emotions and
values to be analogous to the relation between chromatic perceptual experi-
ences and colors. But perceptualists need to be cautious about the language
of formal objects, perhaps even setting it aside (at least initially) as something
that tends to obfuscate the most natural ways of framing perceptualism. The
perceptualist reply that I offer to the Portable Contents and Fading Emotions
Objections emerges by attending in the right way to the core comparison
between emotions and sensory experiences that motivates perceptualism in
the first place.

4.2 *Answering the Portable Contents and Fading Emotions Objections*

Perceptualists can still answer the Portable Contents and Fading Emotions
Objections, but doing so requires being careful about the contemporary dogma
that the formal objects of emotions are corresponding values. A bit of extra
terminology will help to clarify the dialectic. This is the language of *representa-
tional guises*, a notion with roots as far back as Aquinas (see Tenenbaum 2006).
The intuitive idea is that a representational guise is a way of representing that
"casts" content in a certain light. Here is how Kriegel describes such castings:

I propose that we capture this by saying that when a mental state represents *p* under the guise of the *F*, the state does not represent *p* as *F*, but rather *represents-as-F p*. Thus, a belief that *p* does not represent *p* as true, but *represents-as-true p*. That which it represents is simply *p*. Representing-as-true is a *way*, or *mode*, of representing the mode characteristic of belief. (Indeed, it would not be far-fetched to hold that believing *just is* representing-as-true.) What this means is that in representing *p* under the guise of the true, the belief that *p* represents *p* in a "truth-committal" manner. It takes a truth-y stance toward *p*. Similarly, a desire that *p* does not represent *p* as good, but represents-as-good *p*. (Kriegel 2019b, 10) [21]

These remarks indicate a close relationship between the role of representational guises and those often assigned to formal objects. Whereas Deonna and Teroni invoke formal objects to distinguish the attitude of belief from that of desire, Kriegel invokes representational guises to make this distinction. Indeed, Kriegel is explicit that (at least for some attitudes) he recommends conceiving of the property typically cited as the formal object as the representational guise (2019b, 16).

Perceptualists, however, should distinguish between representational guises and formal objects. For the sake of sticking with the custom in emotion theory, they can continue to treat an emotion's formal object as its corresponding value. But then what about the representational guise of emotions? The answer is almost irresistible. After all, the view is called *perceptualism*. As we have seen, the natural perceptualist response to the Perceptualism as a Bad Start Objection says that emotions involve the same experiential way of having content as paradigmatic perceptual experiences. The notion of a representational guise offers a more concrete understanding of this proposal. That is, perceptualism pairs naturally with the view that emotions have the same representational guise as ordinary perceptual experience. One natural candidate for the guise involved in perceptual experience is the following: *representing-as-present* (cf. Kriegel 2019a, 159–160).[22] The idea here is to capture an important feature of the phenomenology of perceiving, namely that

---

21 Kriegel further illustrates the proposal: "If you want to grasp the nature of the attitude of belief, say, think of truth-ascribing content and then rethink the"truthy" aspect of that content as pertaining rather to the psychological attitude taken toward that content" (Kriegel 2019b, 11).

22 Schafer (2013) says that both perceptual experience and belief represent their contents with a certain *force*, namely that of truth (see also Smithies 2018). Schafer uses "force" similarly to

in perceiving one has an impression of certain objects and properties *as being present*; and when a perceptual experience is veridical, one is acquainted with those very properties.[23] So, on this proposal, in perceiving the brown dog one stands in a relation to the dog such that one represents-as-present the brown dog. Similarly, fearing the dog might consist in representing-as-present the dangerous dog.

One worry about this proposal stems from the temporal orientation of some emotions.[24] To illustrate, it might seem as if sadness and fear can't represent-as-present since sadness is about the past and fear about the future. But on closer inspection, there's no immediate cause for concern here. For even if such emotions include in their cognitive bases thoughts directed to the past or future, it doesn't follow that the evaluative properties that they represent would not be present.[25] Consider a person who is sad about having been fired from work. This past event can explain things, most obviously certain absences, that matter *now* (e.g., an absence of fulfilling work). Furthermore, and in general, when a past event ceases to explain anything of negative value in the present (e.g., one finds a better job), one is typically no longer sad, or at least it seems fitting not to be; and so it strikes me as *prima facie* plausible that sadness represents-as-present some negative value (typically grounded in an apparent absence explained by a past event).[26] A similar point works for fear,

---

how I am using "representational guise". But notice that if perceptual experience and belief have the same representational guise, then the phenomenological difference between belief and perceptual experience will not be (even partly) a function of their guises. On one possible view, the phenomenological difference between perceptual experience and belief is primitive (Kriegel 2019a). And while such primitivism is compatible with perceptualism, the view I sketch here aims to avoid this.

23 The idea that emotions/perceptions involve acquaintance with objects and properties is proposed in Ballard (2021b, 121), who is, in turn drawing on Roberts and Wood (2007). Ballard's aim is to argue that such acquaintance is central to the epistemic significance of emotions. In contrast, my aim here is to suggest that this idea can be used to defend a view about the representational guise distinctive of emotions. It should also be noted, however, that this proposal leads likewise to a distinction in the *content* of an emotion and corresponding belief (e.g., fearing something versus believing that that thing is dangerous). The latter relates to a proposition. Thus I think it is somewhat misleading when Roberts (2013, 132) says that emotions can involve "a perceptual acquaintance with moral truths". Strictly speaking, the content of emotions isn't truth-evaluable, though they may be able to justify corresponding beliefs with true (or false) content.

24 Thanks to a referee for raising this issue.

25 I set aside the more familiar worry (independent of the specific proposal here) that emotions cannot be perceptual since they often include non-perceptual states (e.g., imaginings) in their cognitive base (Tappolet 2016, 24–31; Milona and Naar 2020).

26 See Farennikova (2013) on absence perception.

as already indicated by the brown dog example above. In particular, while fear can be driven by thoughts of a possible future outcome, it is the prospect of that outcome *now* that makes something dangerous. In general, then, sadness and fear aren't obviously exceptions to the proposal that emotions represent-as-present. Of course, whether certain emotions are temporally oriented in such a way that they can't be understood to represent-as-present value depends on a detailed study of particular emotions. And while I'm optimistic such explorations will vindicate the present proposal, this is beyond what I can hope to accomplish here.[27]

Whatever one thinks about this specific proposal about the guise involved in perceiving, however, the big picture perceptualist idea is just this: emotions have that very same representational guise as perceptual experience. So insofar as it seems as if formal objects are revelatory of the nature of *attitudes*, rather than the *content* of attitudes, this is because we are overlooking a key point: perceptualism naturally generates a key distinction between an emotion's representational guise and its formal object.[28] The former is common to all emotions while the latter is distinctive of the emotion type in question.

We're now positioned to see how the perceptualist ought to respond to the Portable Contents and Fading Emotions Objections. Let's start with the latter. In presenting that objection, recall that Deonna and Teroni describe Maurice's fading emotional response to Barbara's joke. Despite no longer being amused by the joke, he still *believes* that it's funny. They say, "The fact that an evaluative

---

27 The worry about temporal orientation isn't the only possible concern in the vicinity. For example, one may object that my proposal doesn't extend to emotions in response to fictions (see Teroni 2019). Here people seem to experience emotions (e.g., fear on behalf of a fictional character), even though the relevant value (e.g., danger) isn't present. My favored view is that such emotions systematically misrepresent value, but in a way that can nevertheless be fitting (at least in a sense) insofar as the emotion arises from well-functioning emotional dispositions [see Milona, manuscript, on this sort of fittingness]. Such systematic misrepresentation is explained by the way in which more primitive emotional capacities interact with sophisticated forms of human cognition. By contrast, a more concessive response would allow that there are distinct classes of emotional attitudes, only some of which are strictly speaking subject to a perceptual analysis (cf. Mitchell 2022).

28 Tappolet (2016, 15–16, n40) mentions in passing that perceptualism distinguishes an emotion's formal object (a value) from its constitutive aim (truth or correctness). Constitutive aims are not obviously the same as representational guises. Depending on one's view, the former might indicate a normative standard (cf. Wedgwood 2002) whereas the latter seem to indicate a descriptive or phenomenal feature; but Tappolet confirms (in conversation) that her footnote is meant to gesture at a broadly similar thought to the one developed here (albeit not in the course of addressing the Attitudinalist Challenge). See also footnote 23 above for why perceptualists should be hesitant about taking emotions to aim at truth.

property features in the content of a mental state is hardly sufficient to make it an emotion, let alone an emotion of a specific type" (Deonna and Teroni 2015, 297). But now consider an analogous argument centering on perceptual experience. In particular, take the following, clearly misguided, objection to the view that a perceptual experience as of a red car represents redness (which parodies Deonna and Teroni's statement of the Fading Emotions Objection; cf. Deonna and Teroni (2015), 297):

> Kunal sees Melinda's new red car in his driveway. While they are out riding bikes, he and Melinda chat about her new car. Despite no longer seeing the car, he continues to represent it as red. This indicates that Kunal's color perceptions don't tell us anything about the properties he represents the car as having.

But this objection doesn't work. This is because visual experiences involve a distinctively *perceptual* way of representing certain contents that is importantly different from the way contents are represented in belief. On one view, the difference between perceptual and cognitive ways of representing is primitive, at least on the phenomenal level we're concerned with here (see Kriegel (2019a)). But the notion of representational guises offers hope for (at least partially) analyzing this difference. For example, following Kriegel's suggestion above, and in accord with those who take the formal object of belief to be truth, we may say that believing that *P* is a matter of *representing-as-true P* (Kriegel 2019b, 10; see also Deonna and Teroni 2015, 308).[29] By contrast, perceptual experiences are plausibly oriented to objects and properties, which are more aptly described as present rather than true.

Turn now to the Portable Contents Objection. Recall that, according to this objection, everyday discourse about emotions suggests that different emotions can be *about* the same thing. For example, we might say that one person is angry about what was amusing to another. But if different types of emotions are about the same thing, then, contrary to perceptualism, emotions don't contribute anything to what is represented. To see why this objection shouldn't persuade us, turn once again to ordinary sensory experience. We might say that while Cassandra *heard* the ambulance approaching, Benny *saw*

---

29 Here I am assuming that beliefs, or at least some occurrent beliefs, have a phenomenal character. If they don't, then perceptualists have an easier response to the Fading Emotions Objection. In that case, they would be able to say that emotions are a phenomenal way of having content while beliefs aren't.

the ambulance approaching. The presence of this common content paired with the difference in the phenomenology of the two experiences, may tempt one to conclude that vision and audition are different experiential ways of having content. But this inference would be a mistake.[30] The reason is because Cassandra's auditory experience and Benny's visual experience only have overlapping content, not the same content. After all, Cassandra's experience included various sounds as part of its content while Benny's included various colors and shapes. And perceptual experiences with color content have a very different phenomenology from perceptual experiences with sound content. So when we transfer the reasoning behind the Portable Contents Objection to the perceptual case, the argument fails to show that experiences in different sensory modalities can share their entire content. The Portable Contents Objection, then, really only shows that emotions have overlapping contents, and perceptualists agree with that.

The perceptualist position being proposed here can be further illustrated by way of comparison with the attitude of *disbelieving*.[31] For example, one might say that Obama disbelieves what Trump believes. Here the content of the disbelief and the content of the belief are not exactly the same. This is because "disbelieves" refers to both an attitude as well as a content, perhaps among other things. In particular, it seems to be a shorthand way of referring to a belief that something is not the case (see Price 1989, 120–121). The perceptualist thinks that talk of emotions functions similarly. That is, talk of sadness, anger, joy, etc. refers both to an attitude as well as a content; and it's the content represented under a certain guise that makes a given emotion the emotion that it is.[32]

5 The Fourth Attitudinalist Objection: Standards of Correctness

If what I have argued so far is correct, then perceptualists can also answer the fourth and final Attitudinalist Objection, what I call the *Standards of*

---

30 See Speaks (2015, 178–179). Speaks is drawing on Tye (1995, 156–157).

31 The reasoning in this paragraph draws on Gregory (2021, 10–17). Gregory's aim is to defend the view that desire is a kind of belief. I adapt his reasoning here to support perceptualism.

32 Parallel arguments could be offered for other mental states, e.g., that of rejecting *P*. This likewise seems to refer to an attitude as well as part of its content (cf. Mulligan 2007, 218). Note that, while these proposals about disbelief and rejecting are in my view intuitive and useful for illustrating perceptualism, they aren't unrivaled. See Mulligan (2013) for a detailed discussion.

Correctness Objection. According to this objection, the attitudinal theory better explains the correctness conditions for emotions. By way of comparison, consider that a belief that *P* is correct just in case it is *true* that *P*. Similarly, a desire that *P* is correct just in case it is *desirable* that *P* (or, alternatively, *good* that *P*). The different correctness conditions for the belief and desire are, according to many, explained by the nature of the respective attitudes rather than their contents. After all, as Deonna and Teroni point out:

> [F]ew philosophers go along with Davidson in insisting that believing requires representing a proposition *as true*, or that desiring requires representing a proposition [...] *as desirable* (Deonna and Teroni 2015, 298) [33]

Given a rejection of the Davidsonian approach, they then draw the connection to emotions:

> This encourages the thought that a distinction between the respective contributions of content and attitude to the correctness conditions akin to the one sketched above for belief and desire also holds true for the emotions. To the question: "Why is fear or anger correct if the object or situation to which these emotions are directed is dangerous or offensive?", the straightforward answer is "Because one has the attitude of fear or anger towards it" and not "Because it is represented as being dangerous or offensive." (Deonna and Teroni 2015, 299)

The first point to notice is that Deonna and Teroni seem mistaken in an assumption about perceptualism. They take it as a data point that fear is a correct response to what is dangerous for the trivial reason that one has the attitude of fear toward it (Deonna and Teroni 2015, 299). They also suggest that perceptualists are barred from saying as much. But perceptualists *can* say this. Of course, they also happen to think that what fear consists in is a perceptual way of representing its object as dangerous, in a manner similar to how a visual experience of redness involves a perceptual way of representing its object as red. It is this feature of fear that helps us to understand more deeply *why* fear is a correct response to what is dangerous.

---

33 The Davidsonian approach, at least with respect to desire, is more popular than this quote indicates. For recent defenses of the view that desires involve representations of the good, see Oddie (2005), Schroeder (2007), and Boswell (2018). See Milona and Schroeder (2019) for additional citations and discussion.

The core of Deonna and Teroni's objection, however, is that attitudinalism does a better job of respecting the contributions of both attitude and content to the correctness of an emotion. For example, a belief that *P* is correct just in case it is true that *P*. The content identifies a certain proposition while the attitude (belief) requires that the proposition be true. Similarly, a desire might be thought correct just in case its content is good; and so on for other attitudes. If this is how it works for other attitudes, shouldn't it be the same for emotions? Fearing that *P* is correct if and only if *P* is the case and *P* is dangerous; anger that *P* is correct if and only if *P* is the case and *P* is offensive; and so on for other emotions.

But if what I argued in the previous section is on track, then perceptualists needn't deny that attitude and content both contribute to the correctness conditions of emotions. Perceptualists should say that emotions share their representational guise with ordinary sensory experiences, and this guise contributes to the correctness conditions of different emotions. This is not to my knowledge a point that perceptualists have emphasized.[34] But it's hard to overstate how natural it is for a *perceptualist* about emotions to say this in response to the Standards of Correctness Objection. Incidentally, this is also what perceptualists about *desire* should say. That is, philosophers who maintain that desires are a perceptual representation of some normative property or relation can say that desires represent-as-present their contents (e.g., Oddie 2005). This proposal on behalf of perceptualism about desire, as with emotion, concerns the *total* content of the desire. Perceptualists about emotion/desire think that talk of emotion/desire refers both to an attitude and its proprietary content, each of which make contributions to the correctness conditions of the attitude. And as we saw in the last section, there is nothing obviously *ad hoc* about taking talk of emotions, or desires for that matter, to refer both to attitudes and contents.

Over the course of the last two sections, I have argued that perceptualists should draw a perhaps surprising distinction between an emotion's representational guise—treating it as identical to perceptual experience—and its formal object—taking it to be a value proprietary to the type of emotion in question. I close this section by raising a question about whether attitudinalists may have reason to adopt their own distinction between representational

---

34 But cf. Tappolet's (2016, 15–16, n40) brief remark on the constitutive aim of emotions as well as footnote 28 above.

guises and formal objects.[35] Whether they do may depend on whether they agree with perceptualists about a key dimension of how perceptualists characterize emotional phenomenology. On the view sketched here, perceptualists maintain that emotions share a representational guise with sensory experience, namely that certain content is represented-as-present. This is a way of unpacking Döring's thought that emotions and perceptions put forth certain contents as *actually there* (Döring 2007, 377). Emotional experience is thus unlike (voluntary) imaginative experiences, or suppositions, which do not put forth their contents in this way, and therefore imagination and supposition do not have correctness conditions mirroring that of perceptual experience and emotion. A question thus arises for attitudinalists about whether they would agree with those perceptualists who take emotions to put forth their contents as present. And if so, then there is reason for the attitudinalists to complexify what they take the formal objects of emotions to be, or alternatively to draw their own distinction between representational guises and formal objects.[36] The aim here is not to present an objection to attitudinalism but rather to raise a question that helps us to better frame the possible points of (dis)agreement between perceptualism and various versions of attitudinalism.

6 The Choice between Perceptualism and Attitudinalism

This paper has taken for granted the popular position that emotions are evaluative experiences. The aim has been to show that the interlocking objections comprising the Attitudinalist Challenge do not establish attitudinalism as a better starting point for this position. In this final section, I explain why we might ultimately favor perceptualism over Deonna and Teroni's version of attitudinalism (i.e. action-tendency attitudinalism).

As we've seen, action-tendency attitudinalists maintain that emotional attitudes consist in feelings of readiness to act, and these feelings explain why emotions count as *evaluative* experiences. Here is how Deonna and Teroni describe their position:

---

35 Since writing this paper, I came across an argument in Gregory (2021, 14, n10) that makes similar points to the ones in this paragraph, though in the context of the literature on desire rather than emotion.

36 The question raised here is principally for those attitudinalists who maintain that emotions are evaluative experiences. But as noted in section 2, attitudinalism comes in different forms. Some attitudinalists deny that emotions are experiences of value (e.g., Müller 2017). Attitudinalists of this form may argue that perceptualists have confused the phenomenology of emotion with evaluative feelings that precede emotions.

> Fear of a dog is an experience of the dog as dangerous insofar as
> it is an experience of one's body being prepared to forestall its
> impact (flight, preventive attack, immobility, etc.), an attitude it
> is correct to have if, and only if, the dog is dangerous. (Deonna
> and Teroni 2015, 303; see also Deonna and Teroni 2012, 81)

Deonna and Teroni also maintain that there is a *non-contingent* connection
between the experiential dimension of an emotion and its correctness conditions:

> The body is felt in the form of a gestalt of bodily sensations, which
> consists in being ready to respond in a given way to the object. If
> experiencing such an attitude is all there is to experiencing some-
> thing in evaluative terms, then of course the relation between
> the attitude and the fact that the evaluative property enters into
> the correctness conditions of the mental state is anything but
> contingent. (Deonna and Teroni 2012, 87)

The bodily sensations in fear, for example, are such that they necessarily
count as experiences of their object as dangerous; and this is why fear has
the correctness conditions that it does. To motivate this thought, they point
out that it isn't intelligible that amusement could be a way of making danger
manifest. Given the nature of fear, it seems as if that is the only emotion that
could be an experience of danger (Deonna and Teroni 2012, 86).

A major challenge for action-tendency attitudinalism is to demystify how
emotional experiences count as *evaluative* experiences. Such evaluative expe-
riences aren't simply a matter of covariation:

> [T]he connection between the emotional experience and the eval-
> uative property cannot be modeled on that between smoke and
> fire, namely as one of natural co-variation. Experiencing the eval-
> uative property of an object is not taking the way one's body feels
> as an indication, a sign, or a symptom of the fact that this object
> has this property. (Deonna and Teroni 2012, 87).

Deonna and Teroni argue that a covariational conception of the link between
emotion and value fails to capture the thought that emotional experiences
involve a presentation or manifestation of value.[37] My concern, however, is

---

37 Deonna and Teroni remark, "[W]e cannot conceive of the connection between, for instance,
the phenomenology of fear and danger as arbitrary. Intuitively, no other emotional experience

that action-tendency attitudinalism may ultimately turn out to be, in an important sense, a version of the covariation model (perhaps a kind of necessary covariation), ultimately failing to deliver anything like a presentation of value.

Notice first what the action-tendency attitudinalist *isn't* saying. First, and most obviously, they aren't saying what perceptualists say. A perceptualist, as we've seen, says that danger features in experiences of fear similar to how empirical properties feature in sensory experience (e.g., Roberts (2013, 72–73); Tappolet (2016, 26–28); *inter alia*). Such a view thus well-suited to make sense of the idea that values are manifest in emotional experiences. But Deonna and Teroni deny that emotions make value manifest *in this way* (2012, 68–69).

There is another important view in the vicinity of perceptualism that likewise isn't the action-tendency attitudinalist's. This view can be understood as adapting the proposal sketched above about the representational guise of perceptual experience. According to that proposal, a full description of a perceptual experience requires reference to an attitudinal phenomenology of *representing-as-present* (a being-present-y mode of representation; cf. Kriegel (2019b, 10)). Building on this thought, an attitudinalist might then take emotions to have evaluative representational guises in the manner that perceptual experiences have a representing-as-present guise. Fear, for instance, might be thought to have an attitudinal phenomenology that must be described as representing-as-dangerous.[38] But action-tendency attitudinalists don't have in mind representational guises *of this sort*, either (see Kriegel 2019b, 13).[39] Instead, the action-tendency attitudinalist maintains that the phenomenology of emotional attitudes is properly described in terms of one's body being acti-

---

than that of fear is a suitable candidate for presenting the world in terms of a danger" (Deonna and Teroni 2012, 86). Of course, Deonna and Teroni deny that emotional phenomenology is *exclusively* a matter of value becoming manifest (2015, 308).

38 Kriegel describes such a view with respect to moods. He says the following about the mood of euphoria in particular: "As before, expressions such as 'represents-as-wonderful' function as winks of sorts, with the wink's message being: To grasp the nature of euphoria's distinctive character, think of a wonderfulness-ascribing content and then rethink its 'wonderfulness' dimension as pertaining actually to the subject's attitude toward the content" (Kriegel 2019b, 12). The suggestion here is that an attitudinalist might extend Kriegel's view of moods to emotions.

39 I believe that this is the position attitudinalists *should* adopt, at least insofar as they want to take seriously the view that emotions make value manifest. Such a view also provides a tempting response to Dokic and Lemaire's (2015) argument that attitudinalism collapses into perceptualism (or at least a view that faces as many problems as perceptualism) insofar as it claims that emotions make us aware of value. Unfortunately, however, I haven't space to develop this view and canvass its advantages and disadvantages with respect to perceptualism.

vated in a particular way rather than in evaluative terms. Deonna and Teroni point to the following passage from Nico Frijda to unpack their view:

> In self-focus, analytic attention reduces felt bodily engagement to just that. Felt impulse to shrink back from a threat is transformed into felt muscle tension, just as the feeling of pointing can be transformed into feeling one's finger stretched. (Frijda 2005, 382; quoted in Deonna and Teroni 2015, 308, n19)

Contrast this with the view of perceptual experience offered in section 4: whereas attending to a perceptual experience, according to that proposal, involves attending to the property of *being present* as a dimension of attitudinal phenomenology, the action-tendency attitudinalist doesn't think that attending to emotional experience involves attending to value as a dimension of attitudinal phenomenology.

So how exactly does the action-tendency attitudinalist understand emotions as evaluative experiences? As we've seen, Deonna and Teroni say that emotions are "a gestalt of bodily sensations, which consists in being ready to respond in a given way to the object" (2012, 87). For example, a person who fears a snarling dog may have an experience of their body shrinking away from the snarling dog. But it's not clear that this makes sense of emotions as *evaluative* experiences, or as manifesting value. Even if we add that the action-tendencies associated with different emotions are (necessarily) correct responses to the relevant value, it wouldn't thereby follow that emotions are evaluative experiences. But consider the following: might it be that emotional experiences are evaluative but don't seem evaluative when we *attend* to them?[40] We can see the difficulty with this proposal by returning to Frijda's example of pointing quoted above (2005, 382). Following Frijda, Deonna and Teroni appear to think that in attending to what it feels like to point, the experience seems to just be that of one's finger being stretched. But notice that attending to the entirety of the experience isn't describable simply in terms of the experience of a stretching finger. And even if we attend to the experience in abstraction from what is being pointed to, we aren't left with merely an experience of a stretching finger. This is because a crucial part of the experience of pointing is an experience of indicating, and we *can* attend to this dimension—either in isolation or in conjunction with an object. So if the pointing case provides a model for emotions, then, contrary to what Deonna

---

40 Thanks to a referee for pushing me to consider this possibility.

and Teroni suggest, a description of what we're attending to in emotional experience—even in isolation from the emotion's object—should require reference to an experience of value. But if the action-tendency attitudinalist says this, then they have drifted in the direction of the sort of perceptualist-adjacent phenomenology they want to resist, namely one that retains a representational mode phenomenology even in higher-order attention on the experience itself. So unless the action-tendency attitudinalist can somehow make sense of emotional experiences as evaluative experiences that don't seem evaluative when we attend to them, there is pressure to give up the view that emotions are evaluative experiences.

But how much does it matter whether action-tendency attitudinalism can make sense of emotions as evaluative experiences? The answer depends on what one hopes to accomplish with a theory of emotions. For example, one may be tempted by the view that evaluative knowledge is ultimately rooted in evaluative experiences. Or, more modestly, one may think that evaluative experiences are *an* important route to evaluative knowledge. And mental states like emotions provide a tempting non-mysterious source for what such value experiences might be (Roberts (2013); Tappolet (2016); Milona (2016); *inter alia*).[41] Furthermore, perceptualists are often attracted to the idea that emotions are able to rationalize action and maintain, moreover, that perceptualism can explain how this is possible. We might appeal to fear, for instance, to explain a person's fleeing a bear. If fear is an experience of its object as dangerous, then this renders the action intelligible (Döring 2007). Yet, again, if emotions aren't evaluative experiences, if they are mere felt tendencies to act, then it is not clear that they can rationalize action (as opposed to merely cause it).

7 Conclusion

This paper has explored the Attitudinalist Challenge to perceptualism. The objections comprising the challenge are meant to illustrate that much of our

---

41 Of course, one might assign other (more modest) roles for emotions in value epistemology that don't require emotions to be evaluative experiences. For example, emotions might tend to fix our attention on objects of potential significance, helping us to notice things we otherwise might have missed (see Brady 2013). Furthermore, it's not clear to what extent the roles that Deonna and Teroni assign to emotions require their thesis that emotions are evaluative experiences (2012, 118–125; see also Müller 2017, 304–305). Indeed, as a referee points out to me, some opponents of perceptualism might think that perceptualist's epistemological ambitions lead them to implausible accounts of the nature of emotions.

pretheoretical discourse about emotions conflicts with the perceptualist theses that emotions have, and are individuated by, evaluative content. However, the Attitudinalist Challenge is unpersuasive. Still, adequately addressing the objections requires perceptualists to present their view with greater clarity. In particular, the version of perceptualism presented here draws a crucial and perhaps surprising distinction between an emotion's representational guise, which is uniform across emotions and other perceptual experiences, and its formal object, which is specific to that emotion type. This version of perceptualism emerged in large part by comparing emotions and sensory perceptual experiences, and to this extent marks a natural development of the theory.*

Michael Milona
0000-0003-1474-6367
Toronto Metropolitan University
Milonamc@gmail.com

References

Ballard, Brian Scott. 2021a. "Content and the Fittingness of Emotion." *The Philosophical Quarterly* 71(4): 845–863, doi:10.1093/pq/pqaa074 .

—. 2021b. "The Epistemic Significance of Emotional Experience ." *Emotion Review* 13(2): 113–124, doi:10.1177/1754073920957082.

Block, Ned. 1996. "Mental Paint and Mental Latex." in *Philosophical Issues 7: Perception*, edited by Enrique Villanueva, pp. 19–49. Atascadero, California: Ridgeview Publishing Co., doi:10.2307/1522889.

Boswell, Paul. 2018. "Intelligibility and the Guise of the Good." *Journal of Ethics and Social Philosophy* 13(1), doi:10.26556/jesp.v13i1.225.

Bourget, David. 2017. "Representationalism and Sensory Modalities: An Argument for Intermodal Representationalism." *American Philosophical Quarterly* 54(3): 251–268, doi:10.2307/44982142.

Brady, Michael Sean. 2013. *Emotional Insight. The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199685523.001.0001.

—. 2019. *Emotion: The Basics*. London: Routledge. Copyright year is given as 2019.

Chalmers, David J. 2004. "The Representational Character of Experience." in *The Future for Philosophy*, edited by Brian Leiter, pp. 153–181. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780195311105.003.0011 .

Cooper, Anthony Ashley (third Earl of Shaftesbury). 1699. *An Inquiry Concerning Virtue, in two Discourses*. London: A. Bell, E. Castle, S. Buckley.

Cowan, Robert. 2016. "Epistemic Perceptualism and Neo-Sentimentalist Objections." *Canadian Journal of Philosophy* 46(1): 59–81, doi:10.1080/00455091.2015.1123037.

D'Arms, Justin and Jacobson, Daniel. 2003. "The Significance of Recalcitrant Emotion (or, Anti-Quasijudgmentalism)." in *Philosophy and the Emotions*, edited by Anthony Hatzimoysis, pp. 127–146. Royal Institute of Philosophy Supplement n. 52. Cambridge: Cambridge University Press, doi:10.1017/S1358246100007931.

Deonna, Julien Amos and Teroni, Fabrice. 2008. *Qu'est-ce qu'une émotion?* Chemins Philosophiques. Paris: Librairie philosophique Jean Vrin.

—. 2012. *The Emotions. A Philosophical Introduction*. London: Routledge. Substantially enlarged edition and translation of Deonna and Teroni (2008).

—. 2014. "In What Sense Are Emotions Evaluations?" in *Emotion & Value*, edited by Sabine Roeser and Cain Todd, pp. 15–31. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199686094.001.0001.

—. 2015. "Emotions as Attitudes." *Dialectica* 69(3): 293–311. Special issue "Beyond Perceptualism," edited by Sabine A. Döring and Anika Lutz, doi:10.1111/1746-8361.12116.

Dietz, Christina Hope. 2018. "Reasons and Factive Emotions." *Philosophical Studies* 175(7): 1681–1691, doi:10.1007/s11098-017-0929-y.

Dokic, Jérôme and Lemaire, Stéphane. 2013. "Are Emotions Perceptions of Value?" *Canadian Journal of Philosophy* 43(2): 227–247, doi:10.1080/00455091.2013.826057.

—. 2015. "Are Emotions Evaluative Modes?" *Dialectica* 69(3): 271–292. Special issue "Beyond Perceptualism," edited by Sabine A. Döring and Anika Lutz, doi:10.1111/1746-8361.12117 .

Döring, Sabine A. 2007. "Seeing What to Do: Affective Perception and Rational Motivation." *Dialectica* 61(3): 363–394, doi:10.1111/j.1746-8361.2007.01105.x.

Döring, Sabine A. and Lutz, Anika. 2015. "Beyond Perceptualism: Introduction." *Dialectica* 69(3): 259–270. Special issue "Beyond Perceptualism," edited by Sabine A. Döring and Anika Lutz, doi:10.1111/1746-8361.12106.

Farennikova, Anna. 2013. "Seeing Absence." *Philosophical Studies* 166(3): 429–454, doi:10.1007/s11098-012-0045-y.

Frijda, Nico H. 2005. "Emotion Experience." *Consciousness and Emotion* 19(4): 473–497, doi:10.1080/02699930441000346.

—. 2007. *The Laws of Emotion*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

GREEN, O. H. 1992. *The Emotions: A Philosophical Theory*. Philosophical Studies Series n. 53. Dordrecht: D. Reidel Publishing Co., doi:10.1007/978-94-011-2552-9.

GREGORY, Alex. 2021. *Desire as Belief: A Study of Desire, Motivation, and Rationality*. Oxford: Oxford University Press, doi:10.1093/oso/9780198848172.001.0001.

GRICE, H. Paul. 1962. "Some Remarks about the Senses." in *Analytical Philosophy, First Series*, edited by Ronald J. BUTLER, pp. 133–153. New York: Barnes; Noble. Reprinted in Grice (1989, 248–268).

—. 1989. *Studies in the Way of Words*. Cambridge, Massachusetts: Harvard University Press.

GRZANKOWSKI, Alex. 2017. "The Real Trouble with Recalcitrant Emotions." *Erkenntnis* 82(3): 641–651, doi:10.1007/s10670-016-9836-4.

KENNY, Anthony John Patrick. 1963. *Action, Emotion and the Will*. London: Routledge & Kegan Paul. Second edition: Kenny (2003).

—. 2003. *Action, Emotion and the Will.* 2nd ed. London: Routledge. First edition: Kenny (1963), doi:10.4324/9780203711460.

KRIEGEL, Uriah. 2013. "The Phenomenal Intentionality Research Program." in *Phenomenal Intentionality*, edited by Uriah KRIEGEL, pp. 1–26. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199764297.003.0001.

—. 2019a. "The Intentional Structure of Moods." *Philosophers' Imprint* 19(49).

—. 2019b. "The Perception/Cognition Divide: One More Time, With Feeling." in *The Philosophy of Perception. Proceedings of the 40th International Wittgenstein Symposium*, edited by Christoph LIMBECK-LILIENAU and Friedrich STADLER, pp. 149–170. Publications of the Austrian Ludwig Wittgenstein Society (new series) n. 26. Berlin: Walter De Gruyter, doi:10.1515/9783110657920-009.

LAZARUS, Richard S. 1991. "Cognition and Motivation in Emotion." *American Psychologist* 46(4): 352–367, doi:10.1037//0003-066x.46.4.352.

MEINONG, Alexius. 1917. "Über emotionale Präsentation." *Sitzungsberichte der philosophisch-historischen Klasse der Kaiserlichen Akademie der Wissenschaften in Wien* 183(2): 3–181. Reprinted in Meinong (1968, 3:285–465).

—. 1968. *Abhandlungen zur Werttheorie*. vol. 3. Alexius Meinong Gesamtausgabe. Graz: Akademische Druck- und Verlagsanstalt. Bearbeitet von R. Kindinger.

—. 1972. *On Emotional Presentation*. Evanston, Illinois: Northwestern University Press. Translation of Meinong (1917) andIntroduction by Marie-Luise Kalsi, foreword by John Drummond, doi:10.21985/N26T5J.

MILONA, Michael. 2016. "Taking the Perceptual Analogy Seriously ." *Ethical Theory and Moral Practice* 19(4): 897–915, doi:10.1007/s10677-016-9716-7.

MILONA, Michael and NAAR, Hichem. 2020. "Sentimental Perceptualism and the Challenge from Cognitive Bases." *Philosophical Studies* 177(10): 3071–3096, doi:10.1007/s11098-019-01360-7.

Milona, Michael and Schroeder, Mark. 2019. "Desiring under the Proper Guise." in *Oxford Studies in Metaethics*, volume XIV, edited by Russ Shafer-Landau, pp. 121–143. Oxford: Oxford University Press, doi:10.1093/oso/9780198841449.003.0006.

Mitchell, Jonathan. 2017. "The Epistemology of Emotional Experience." *Dialectica* 71(1): 57–84, doi:10.1111/1746-8361.12171 .

—. 2021. "The Bodily-Attitudinal Theory of Emotion." *Philosophical Studies* 178(8): 2635–2663, doi:10.1007/s11098-020-01567-z.

—. 2022. "Two Irreducible Classes of Emotional Experiences: Affective Imaginings and Affective Perceptions." *European Journal of Philosophy* 30(1): 307–325, doi:10.1111/ejop.12648.

Müller, Jean Moritz. 2017. "How (Not) to Think of Emotions as Evaluative Attitudes." *Dialectica* 71(2): 281–308, doi:10.1111/1746-8361.12192.

—. 2019. *The World-Directedness of Emotional Feeling. On Affect and Intentionality*. London: Palgrave Macmillan, doi:10.1007/978-3-030-23820-9.

Mulligan, Kevin. 2007. "Intentionality, Knowledge and Formal Objects." *Disputatio* 2(23): 205–228, doi:10.2478/disp-2007-0010.

—. 2010. "Emotions and Values." in *The Oxford Handbook of Philosophy of Emotion*, edited by Peter Goldie, pp. 475–500. Oxford Handbooks. New York: Oxford University Press, doi:10.1093/oxfordhb/9780199235018.003.0022.

—. 2013. "Acceptance, Acknowledgment, Affirmation, Assertion, Belief, Certainty, Conviction, Denial, Judgment, Refusal and Rejection." in *Judgement and Truth in Early Analytic Philosophy and Phenomenology*, edited by Mark Textor, pp. 97–136. History of Analytic Philosophy. London: Palgrave Macmillan, doi:10.1057/9781137286338_6.

Naar, Hichem. 2019. "Emotion: Animal and Reflective." *The Southern Journal of Philosophy* 57(4): 561–588, doi:10.1111/sjp.12359.

—. 2020. "The Real Issue with Recalcitrant Emotions: Reply to Grzankowski (2017)." *Erkenntnis* 85(5): 1035–1040, doi:10.1007/s10670-018-0063-z.

—. 2022. "Emotion: More like Action than Perception." *Erkenntnis* 87(6): 2715–2744, doi:10.1007/s10670-020-00324-2.

Nussbaum, Martha Craven. 2004. "Emotions as Judgments of Value and Importance." in *Thinking about Feeling. Contemporary Philosophers on Emotion*, edited by Robert C. Solomon, pp. 183–199. Series in Affective Science. Oxford: Oxford University Press, doi:10.1093/oso/9780195153170.001.0001.

O'Callaghan, Casey. 2015. "The Multisensory Character of Perception." *The Journal of Philosophy* 112(10): 551–569. Reprinted in O'Callaghan (2017, 171–188), doi:10.5840/jphil20151121035.

—. 2017. *Beyond Vision: Philosophical Essays*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198782964.001.0001.

ODDIE, Graham. 2005. *Value, Reality and Desire*. Oxford: Oxford University Press, doi:10.1093/0199273413.001.0001.

POELLNER, Peter. 2016. "Phenomenology and the Perceptual Model of Emotion." *Proceedings of the Aristotelian Society* 116(3): 261–288, doi:10.1093/arisoc/aow015.

PRICE, Huw. 1989. "Defending Desire-as-Belief." *Mind* 98(389): 119–127, doi:10.1093/mind/XCVIII.389.119.

PRINZ, Jesse J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199571543.001.0001.

ROBERTS, Robert C. 2013. *Emotions in the Moral Life*. Cambridge: Cambridge University Press, doi:10.1017/CBO9781139061469.

ROBERTS, Robert C. and WOOD, W. Jay. 2007. *Intellectual Virtues. An Essay in Regulative Epistemology*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199283675.001.0001.

ROSSI, Mauro and TAPPOLET, Christine. 2019. "What Kind of Evaluative States are Emotions? The Attitudinal Theory vs. the Perceptual Theory of Emotions [critical notice of Deonna and Teroni (2012), Deonna and Teroni (2014), Deonna and Teroni (2015)]." *Canadian Journal of Philosophy* 49(4): 544–563, doi:10.1080/00455091.2018.1472516.

SALMELA, Mikko. 2011. "Can Emotion be Modelled on Perception?" *Dialectica* 65(1): 1–29, doi:10.1111/j.1746-8361.2011.01259.x.

SCHAFER, Karl. 2013. "Perception and the Rational Force of Desire." *The Journal of Philosophy* 110(5): 258–281, doi:10.5840/jphil2013110528.

SCHROEDER, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199299508.001.0001.

SIEGEL, Susanna. 2021. "The Contents of Perception." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language; Information. Revision , August 30, 2021, of the version of 18 March 2005, https://plato.stanford.edu/entries/perception-contents/.

SMITHIES, Declan. 2018. "Reasons and Perception." in *The Oxford Handbook of Reasons and Normativity*, edited by Daniel STAR, pp. 631–661. Oxford Handbooks. New York: Oxford University Press, doi:10.1093/oxfordhb/9780199657889.013.28.

SOLOMON, Robert C. 1976. *The Passions. The Myth and Nature of Human Emotion*. Notre Dame, Indiana: Doubleday.

SPEAKS, Jeff. 2015. *The Phenomenal and the Representational*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198732556.001.0001.

STOCKDALE, Katie. 2017. "Losing Hope: Injustice and Moral Bitterness ." *Hypatia: A Journal of Feminist Philosophy* 32(2): 363–379, doi:10.1111/hypa.12314.

TAPPOLET, Christine. 2016. *Emotions, Values, and Agency*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199696512.001.0001.

Tenenbaum, Sergio. 2006. *Appearances of the Good. An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press, doi:10.1017/CBO9780511498855.

Teroni, Fabrice. 2007. "Emotions and Formal Objects." *Dialectica* 61(3): 395–415, doi:10.1111/j.1746-8361.2007.01108.x.

—. 2019. "Emotion, Fiction and Rationality ." *The British Journal of Aesthetics* 59(2): 113–128, doi:10.1093/aesthj/ayz015.

Tracy, Jessica L. and Randles, Daniel. 2011. "Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt." *Emotion Review* 3(4): 397–405, doi:10.1177/1754073911410747.

Tye, Michael. 1995. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, Massachusetts: The MIT Press.

—. 2007. "The Problem of Common Sensibles ." *Erkenntnis* 66(1–2): 287–303, doi:10.1007/s10670-006-9035-9.

Wedgwood, Ralph. 2002. "The Aim of Belief." in *Philosophical Perspectives 16: Language and Mind*, edited by James E. Tomberlin, pp. 267–297. Oxford: Basil Blackwell Publishers, doi:10.1111/1468-0068.36.s16.10.

Whiting, Demian. 2012. "Are Emotions Perceptual Experiences of Value?" *Ratio* 15(1): 93–107, doi:10.1111/j.1467-9329.2011.00518.x.

Abstracting and Indexing Services

The journal is indexed by the Arts and Humanities Citation Index, Current
Contents, Current Mathematical Publications, Dietrich's Index
Philosophicus, IBZ — Internationale Bibliographie der Geistes- und
Sozialwissenschaftlichen Zeitschriftenliteratur, Internationale Bibliographie
der Rezensionen Geistes- und Sozialwissenschaftlicher Literatur, Linguistics
and Language Behavior Abstracts, Mathematical Reviews, MathSciNet,
Periodicals Contents Index, Philosopher's Index, Repertoire Bibliographique
de la Philosophie, Russian Academy of Sciences Bibliographies.

Contents