

dialectica

International Journal of Philosophy

Contents

ADAM PATTERSON, <i>Avoid Avoiding the Wishful Thinking Problem</i>	257
RAN LANZET, <i>Whittle vs. Cantor on the Size of Infinite Sets</i>	271
EMILIE PAGANO, <i>Functionalism, Pluralities, and Groups</i>	283
SERENA GALLI, <i>Causal Inference from Big Data?: A Reply to Pietsch (2021)</i>	319
CIRO DE FLORIO & ALDO FRIGERIO, <i>Retro-Closure Principle and Omniscience</i>	335
AGNÈS BAEHNI, <i>Review of Schofield (2021)</i>	363
STEVE HUMBERT-DROZ, <i>Review of Lepine (2023)</i>	369
JONATHAN SIMON, <i>Review of Landgrebe and Smith (2022)</i>	377

dialectica

International Journal of Philosophy

Official Organ of the European Society of Analytic Philosophy

founded in 1947 by Gaston Bachelard, Paul Bernays and Ferdinand Gonseth

Editorial Board

Jérôme Dokic, EHESS, Paris, France

Pascal Engel, EHESS, Paris, France

Manuel García-Carpintero, Universitat de Barcelona, Spain

Diego Marconi, Università di Torino, Italy

Carlos Moya, Universitat de València, Spain

Martine Nida-Rümelin, Université de Fribourg, Switzerland

François Recanati, Collège de France, Paris

Marco Santambrogio, Università degli Studi di Parma, Italy

Peter Simons, Trinity College Dublin, Ireland

Gianfranco Soldati, Université de Fribourg, Switzerland

Marcel Weber, Université de Genève, Switzerland

Editors

Fabrice Correia, University of Geneva

Philipp Blum, University of Lucerne

Marco R. Schori, University of Bern (managing editor)

Review Editors

Stephan Leuenberger and Philipp Blum

Editorial Committee

Sara Amighetti, Joshua Babic, Davood Bahjat, Philipp Blum (né Keller), Claudio Calosi, Alessandro Cecconi, Zoé Christoff, Fabrice Correia, Matthias Egg, Andrea Giananti, Martin Glazier, Aleks Knoks, Arturs Logins, Jörg Löschke, Giovanni Merlo, Robert Michels, Ryan Miller, Michael Müller, Paolo Natali, Donnchadh O'Conaill, Simone Olivadoti, Edgar Phillips, Stephanie Rennick, Sebastian Schmidt, Marco R. Schori, David Schroeren, Mike Stuart, Daniel Vanello.

Consulting Board

Johannes Brandl (Salzburg), João Branquinho (Lisboa), Elke Brendel (Bonn), Ingar Brinck (Lunds), Eros Corazza (Ikerbasque and Carleton), Josep Corbi (València), Michael Esfeld (Lausanne), Dagfinn Føllesdal (Stanford and Oslo), Frank Jackson (Australian National University, Canberra), Max Kistler (Paris I), Max Kölbel (Wien), Jan Lacki (Genève), Karel Lambert (Irvine), Paolo Leonardi (Bologna), Fraser Macbride (Manchester), Josep Macià (Barcelona), Genoveva Martí (Barcelona), Élisabeth Pacherie (Institut Jean Nicod, Paris), David Piñeda (Girona), Wlodek Rabinowicz (Lund), Barry Smith (Buffalo), Christine Tappolet (Montréal), Neil Tennant (Ohio State), Mark Textor (King's College London), Achille Varzi (Columbia University), Alberto Voltolini (Torino), Timothy Williamson (Oxford).

September 2023

Contents

ADAM PATTERSON, <i>Avoid Avoiding the Wishful Thinking Problem</i>	257
RAN LANZET, <i>Whittle vs. Cantor on the Size of Infinite Sets</i>	271
EMILIE PAGANO, <i>Functionalism, Pluralities, and Groups</i>	283
SERENA GALLI, <i>Causal Inference from Big Data?: A Reply to Pietsch (2021)</i> . .	319
CIRO DE FLORIO & ALDO FRIGERIO, <i>Retro-Closure Principle and Omniscience</i>	335
AGNÈS BAEHNI, <i>Review of Schofield (2021)</i>	363
STEVE HUMBERT-DROZ, <i>Review of Lepine (2023)</i>	369
JONATHAN SIMON, <i>Review of Landgrebe and Smith (2022)</i>	377

PROOF

1 Avoid Avoiding the Wishful Thinking 2 Problem

ADAM PATTERSON

3 The wishful thinking problem purported to be a new problem for pure
4 non-cognitivist expressivist views in metaethics in addition to the similar,
5 yet distinct, Frege-Geach problem. After a smattering of initial responses,
6 discussion of the problem has faded. One might think this is because the
7 responses were fatal, and the problem is not really a problem. I do not
8 think so. I aim to re-start discussion of the wishful thinking problem. I
9 do so by recasting it in terms of the distinction between propositional
10 and doxastic justification. Doing so is instructive, for it shows some of
11 the initial, prominent responses to the problem fail. The problem is thus
12 not as dead as one might otherwise think.

13 Consider two things. First, consider the distinction between *propositional* and
14 *doxastic* justification for some subject, S's, belief that *p*. On one hand, S is
15 *propositionally* justified in believing that *p* when S has sufficient reasons, R,
16 to believe that *p*. S's belief that *p* is *justifiable*, in other words. On the other
17 hand, S's belief that *p* is *doxastically* justified when S believes that *p*, *given*
18 (or *on the basis of*) R (Silva and Oliveira 2024). That is, S's belief that *p* is
19 *justified* (Korcz 2000).¹ The crucial difference is the basing relation: there is
20 a difference between *having* (available, or at hand) R to believe that *p* and
21 actually believing *p based on* R (Alston 1985). Second, consider the younger
22 cousin of the popular Frege-Geach problem for non-cognitivist, expressivist
23 meta-ethical views: the oft-neglected wishful thinking problem (Dorr 2002).
24 While the latter is about *validity*, the former is about *justification*. How are
25 these two things related?

26 The above distinction plays an important role in both characterizing the
27 problem and evaluating its proposed solutions. Yet few (if any) explicitly
28 acknowledge this. I remedy that here. Doing so is instructive, for as we will see,

1 Henceforth, "S *justifiedly* believes that *p*" (and similar expressions) means the belief is doxastically justified, and "S *justifiably* believes that *p*" means the belief is "propositionally justified."

understanding both the problem and its solutions with the above distinction in mind reveals the ways that the responses fail. This is an interesting result, for the problem is thus not as dead as it seems.

In section 1, I recast the wishful thinking problem in terms of two kinds of justification. In section 2, I do the same for several prominent responses. I also argue that they fail. In section 3, I conclude.

3.1 The Wishful Thinking Problem: Recast

Consider the following moral-descriptive² *modus ponens* (Dorr 2002). Call it the *Liar Argument* (Long 2016).³

LIAR ARGUMENT

(P₁) If lying is wrong, the souls of liars will be punished in the afterlife.

(P₂) Lying is wrong.

(C) So, the souls of liars will be punished in the afterlife.

As we can see, moral-descriptive *modi ponentes* have as their major premise a conditional claim. That conditional's antecedent is a moral claim, whereas the conclusion is a non-moral (descriptive) claim.

Now consider Edgar. Edgar is reasoning himself through the **LIAR ARGUMENT**. The states of affairs as he does can be represented as follows (Dorr 2002, 98):

T₁ Edgar's belief that (P₁) and ?(C)⁴ are both doxastically justified. He believes \neg (P₂).

At T₁, it seems *irrational* for Edgar to believe (C) for two reasons. First, it is incoherent to believe that (C) given that he already justifiably believes (P₁)—on the basis of reliable testimony—and also believes that \neg (P₂). Second, any belief that (C) at this time lacks *propositional* justification. Edgar, in fact, right now has *good reason* to be ambivalent about (C) and is ambivalent precisely because of those reasons (Dorr 2002, 98).

² The "moral-descriptive" label is from Schroeder (2011).

³ Mabrito (2013) calls this "the damnation argument."

⁴ The "?C" denotes ambivalence about C. I borrow this from Guan (2014).

56 Now suppose that Edgar then reads some moral philosophy. As a result, he
 57 reconsiders his moral beliefs and thereby comes to immediately, justifiably
 58 believe (P₂). Thus:

59 T₂ Edgar has doxastic justification for (P₁), ?(C), and (P₂).

60 He now does as attitude-coherence demands: He revises. Given that Edgar has
 61 doxastic justification for (P₁) and (P₂), he jettisons his previous ambivalence
 62 about (C). So, he comes to believe (C) on the basis of (P₁) and (P₂). Hence:

63 T₃ Edgar has doxastic justification for (C).

64 According to Dorr (2002), pure non-cognitivist expressivism⁵ struggles to
 65 explain cases like this. Why is that?

66 At T₁, it is irrational for Edgar to believe (C). This seems plausible. After all,
 67 Edgar has no justification to believe that (C). Moreover, he cannot justifiably
 68 believe that (C) on the basis of what he believes at T₁ on pains of incoherence.
 69 So, any belief that (C) of Edgar's is neither propositionally nor doxastically
 70 justified.

71 Now, from T₁–T₂, Edgar's mental states changed. In particular, it seems
 72 like he underwent a change in beliefs; he gained a new one. He *came to believe*
 73 *that* (P₂) for the first time at T₂, and justifiably so. Now, a diachronic change
 74 in mental states in general is compatible with non-cognitive expressivism.
 75 The trouble is that on non-cognitive expressivism, Edgar did not gain a belief.
 76 Rather, his *non-cognitive states changed*. That is, Edgar only gained a new
 77 *non-cognitive* state. This is because (P₂), remember, is a moral claim. And that
 78 the state of accepting a moral claim is a *non-cognitive state* is part and parcel
 79 of non-cognitivist expressivism.

80 This is bad for non-cognitive expressivism. It means that on that view, Edgar
 81 *still* cannot justifiably believe that (C) at T₂. That is, at T₂, Edgar's belief that
 82 (C) still lacks doxastic justification for him—just as did for him at T₁. This is
 83 because a new non-cognitive state cannot be that on the basis of which Edgar
 84 justifiedly believes that (C). So, for the non-cognitive expressivist, nothing
 85 changed from T₁–T₂ that explains why it seems intuitively rational for Edgar
 86 to believe (C) *on the basis of* (P₁) and (P₂), or why Edgar seems rational to
 87 justifiedly believe that (C) on the basis of only accepting the premises.

5 I henceforth drop the "pure."

Said differently, it is intuitively plausible that the following two states can obtain over time in the Edgar case:⁶

- (i) At T_1 , it is *irrational* for Edgar to believe (C). For he justifiably believes (P₁), ?(C), and believes \neg (P₂). So, the belief that (C) is not propositionally justified and cannot be doxastically justified (C) given what he believes;

and

- (ii) It is *rational* for Edgar at T_3 to believe (C). He has doxastic justification for both (P₁) and (P₂), and he accepts (C) on their basis. So, he has doxastic justification for (C).

Yet on non-cognitivist expressivism, both cannot obtain. On that view, (ii) is not possible. This is because Edgar's belief that (C) at T_3 is neither justified nor justifiable. Remember: this was also true at T_1 . The only diachronic change in Edgar's mental states was his attaining a new *non-cognitive state* at T_2 . But non-cognitive states cannot be justifiers, things on which it is rational to base our beliefs; they are just the wrong kinds of things.⁷

Call this the *wishful thinking problem* for pure, non-cognitivist expressivism.⁸ It concerns justification. More precisely, given the *doxastic/propositional justification* distinction, one can see that it concerns whether Edgar's coming to believe (C) over time is rational because (and insofar as) said belief

6 This is how Mabrito (2013) perspicuously frames the problem.

7 Dorr (2002, 99) seems to implicitly rely on this to explain why "only a change in one's cognitive states, or in one's evidence, can make the difference between a case in which it would be irrational to believe something and one in which it would be rational to do so." But why think that non-cognitive states cannot be justifiers? Because revising your views about the world is rational when the change coheres with your *belief set*. It is, he says, "irrational to modify your views about the world so that they cohere with your desires and feelings" (Dorr 2002, 99). But again, one might ask: Why? Pryor's (2005) discussion of Davidson might be relevant. According to Pryor (2005), for Davidson, if some state, x , is a justifier, then x has propositional content, content that is expressible with "that-clauses" and which functions to assertively represent the world as being-such-and-such-a-way. This is because only by standing in *logical* relation to a belief can some state doxastically justify a belief. And since the state of accepting (P₂) is, by dint of non-cognitive expressivism, a *non-cognitive state*, and these do not function to assertively indicate anything about the world, it cannot stand in a logical relation to Edgar's belief that (C). Hence, Edgar's coming to believe (P₂), on pure non-cognitivist expressivism, cannot doxastically justify (C).

8 On non-cognitive expressivism, moral evaluations seem like danglers qua justifiers.

108 becomes *doxastically justified* after having been not previously even proposi-
 109 tionally justified, i.e., *justifiable*.

110 With this in mind, let us now re-evaluate some proposed solutions. As we
 111 will see, I find them all wanting, given this understanding of the problem.

112 Re-Evaluating Some Proposed Solutions

2131 The Decalogue Proposal

114 Consider the following argument (Lenman 2003). Call it the *decalogue argu-*
 115 *ment* (Schroeder 2011). Suppose that S is reasoning through it over time, and
 116 we can represent how his beliefs seemingly rationally change over time like
 117 this:

- 118 T₁ (P₃) S never contravenes the Decalogue.
 119 (P₄) All and only contraventions of the Decalogue are wrong.
 120 (P₅) S never does anything wrong.

121 Currently, for S to believe the descriptive claim that “S never looks at a woman
 122 with lustful intent” would be neither *justified* nor *justifiable*. It is irrational
 123 for S to believe that at T₁. Now, suppose that on the basis of (P₃) and (P₄), S
 124 comes to *justifiably believe* the following:

- 125 T₂ (P₆) If looking at a woman with lustful intent is wrong, then S never
 126 looks at a woman with lustful intent.

127 Thus, at T₂, the belief that (P₆) is doxastically justified for S; it is justifiedly
 128 believed on the basis of a pair of claims, (P₃) + (P₄). However, S’s belief in
 129 the descriptive claim “S never looks at a woman with lustful intent” is still
 130 neither justified nor justifiable at T₂ and is irrational to believe.

131 Moving on, S subsequently comes to believe two *more* things:

- 132 T₃ (P₇) Looking at a woman with lustful intent contravenes the Decalogue.
 133 (P₈) Looking at a woman with lustful intent is wrong.

134 More precisely, here S gains at least one new belief: (P₈). S believes it, given (or
 135 on the basis of) his belief in (P₄) + (P₇). So, the belief that (P₈) is doxastically
 136 justified for S. Also, at T₃, notice that it is *justifiable* for S to believe that “S

137 never looks at a woman with lustful intent.” This is because (P₃) + (P₇)⁹
 138 is sufficient reason to believe it. Finally, after all this, S comes to justifiedly,
 139 rationally believe that (C₁) on the basis of (P₆) + (P₈).

140 T₄ (C₁) S never looks at a woman with lustful intent.

141 What is the point of all of this? In general, Lenman’s (2003) goal seems to be
 142 to show that it can be rational to infer the conclusion of a moral-descriptive
 143 *modus ponens* without wishful thinking. This is achieved by showing that S is
 144 guaranteed to have evidence for (C₁) that can justify S’s coming to believe (C₁)
 145 without wishful thinking; S’s justification for believing (C₁) is guaranteed to
 146 be overdetermined, in other words.¹⁰ For in the very act of accepting a moral
 147 claim like (P₈), one is guaranteed to accept beliefs that support it and also
 148 support (C₁) independently of (P₆) and (P₈).

149 I have said nothing yet about propositional/doxastic justification. But now
 150 I ask: how can one understand the *decalogue argument* and this proposed
 151 solution in general, given the propositional/doxastic justification distinction?
 152 Asked differently: how can this solution be recast with the distinction—with it,
 153 how can one explain the way in which this proposal vindicates non-cognitive
 154 expressivism from the wishful thinking problem?

155 To see, recall (C₁). Also, recall that the solution works, in part, by guaran-
 156 teeing overdetermined justification for S’s belief that (C₁) by the very process
 157 of coming to believe it in the first place. Further, recall: this overdetermined
 158 justification is secured by finding two pairs of claims. One is (P₆) + (P₈). The
 159 other is (P₃) + (P₇). Both sets justify S’s coming to believe (C₁). Hence, with
 160 the relevant distinction between kinds of justification in mind, the proposal
 161 works by establishing one of two things. S’s belief is guaranteed to have, for S,

162 (iii) overdetermined doxastic justification;

163 (In which case, the *decalogue case* shows that S’s belief that (C₁) is always
 164 partly based on (P₃) + (P₇). So, said belief is doxastically justified, i.e., based
 165 on two sets of claims, one set of which lacks a moral claim. And thus, basing
 166 the belief on (C₁) is always, in part, not wishful thinking on S’s part.) or,

9 And remember: both (P₃) and (P₇) were each part of the arguments for claims that make S’s belief that (C₁) doxastically justified, namely (P₆) and (P₈).

10 I borrow characterizing this proposal in terms of guarantees from an anonymous referee.

167 (iv) overdetermined propositional justification.

168 (In which case, two sets of claims always support (C₁), and thus S's coming to
 169 believe (C₁) is rational. This is because (and insofar as) while S's believe that
 170 (C₁) is *justifiedly inferred* on the basis of a non-cognitive claim—(P₈)—the
 171 inference is guaranteed to be *justifiable* given (P₃) + (P₇.) This puts the
 172 non-cognitivist in a dilemma.

173 If the proposal establishes (iii), then it relies on the following assumption:
 174 Namely, for all of S's doxastically-justified moral beliefs, whenever a moral
 175 belief is doxastically justified for S on the basis of R, and that moral belief
 176 entails some non-moral claim, then R doxastically justifies S's belief in the non-
 177 moral claim. This is dubious, though. S can be unaware of what *propositionally*
 178 *justifies* the non-moral belief, which is the fact that it is entailed by the moral
 179 belief. This means that S need not necessarily form the non-moral belief
 180 *on the basis of* the moral belief, in which case that non-moral belief is not
 181 doxastically justified.

182 Moreover, if the proposal establishes (iii), then S is still basing their be-
 183 lief that (C₁) *in part* on the basis of (P₈). There is still wishful thinking
 184 present; there is just less of it on this proposal since it is also based on (P₃)
 185 and (P₇)—neither of which are non-cognitive claims given non-cognitive ex-
 186 pressivism. The proposal would be better to establish the following: in coming
 187 to believe (C₁) at T₄, S's belief is *only* based on that other pair of claims—the
 188 pair that lacks a moral claim, namely, (P₃) + (P₇)—and thereby makes the
 189 justified belief in (C₁) not a case of wishful thinking.

190 If the proposal establishes (iv), then it does not help with the wishful
 191 thinking problem. It needs to be shown that S is rational because (and insofar
 192 as) S went from, at one time, having no justification for believing (C₁) to
 193 having doxastic justification for it (and crucially without believing (C₁) on the
 194 basis of a non-cognitive attitude). But if the proposal establishes (iv), this still
 195 happens: S still comes to accept (C₁) on the basis of (P₈). So, S's belief is based
 196 on, problematically, a non-cognitive attitude. That S's belief is guaranteed
 197 some bonus propositional justification may lessen the sting of a charge of
 198 irrationality. But it stings nonetheless.

199 So, if the proposal showed that S's belief in (C₁) is rational insofar as the
 200 belief is guaranteed to be always *justified* or *justifiable* for S without wishful
 201 thinking, then it either relies on a false assumption or fails to address the
 202 problem. Either way, the wishful thinking problem remains.

2a.2 *The Modified Proposal*

204 Another way to deal with the *wishful thinking problem* is to attempt to argue
 205 that, in reasoning through the **LIAR ARGUMENT**, S is guaranteed to be *proposi-*
 206 *tionally justified* in believing (C). So, for example, whenever S argues through
 207 the **LIAR ARGUMENT**, S has available to them at any time via introspection
 208 the following modified, companion argument (Enoch 2003):

209 MODIFIED ARGUMENT¹¹

- 210 (P₉) If I accept that lying is wrong, the souls of liars will be punished in the
 211 afterlife.
 212 (P₁₀) I accept that lying is wrong.
 213 (C) So, the souls of liars will be punished in the afterlife.

214 This **MODIFIED ARGUMENT** is always available through introspection. The
 215 idea is thus that the **LIAR ARGUMENT** will never lead Edgar to irrationally
 216 believe (C) in the sense that S believes it without that which sufficiently
 217 justifies it.

218 This line is also problematic. It fails to guarantee that S's belief that (C) is
 219 justifiable for S. Why is that? The **MODIFIED ARGUMENT** is "available" to S
 220 in a weak sense. It is guaranteed to be possible that S can come to possess a
 221 **MODIFIED ARGUMENT**. But this does not entail that S actually *has available*
 222 (or is in the possession of) the relevant argument. The mere *presence* of the
 223 **MODIFIED ARGUMENT** only propositionally justifies Edgar's acceptance of
 224 (C) only if Edgar actually *has accepted* (P₉) and (P₁₀).

225 Suppose that I am wrong. Suppose that, on this strategy, any time S accepts
 226 (P₁) and (P₂), S will always possess (in some suitably strong sense) propo-
 227 sitionally justification for (C) since there will always be other things that
 228 propositionally justify (C).

229 This proposal does not address the problem, either, for familiar reasons. In
 230 particular, the problem is that when S forms the belief that (C) *on the basis of*
 231 both (P₁) and (P₂),

- 232 (P₁) If lying is wrong, the souls of liars will be punished in the afterlife;

11 Here, I continue the sequential numbering of premises from the decalogue argument. This does not mean that the **MODIFIED ARGUMENT** is a part of (or some extension of) the *decalogue argument*. I use this numbering convention to avoid referring to one premise in a distinct argument with the same expression.

233 (P2) Lying is wrong;

234 then S's subsequent belief that (C) is doxastically justified. And that means that
 235 S's belief that (C) is doxastically justified whether S is also justified in having
 236 some *other, auxiliary belief(s)*. Said differently: what generates the problem in
 237 the original case is that it seems rational to believe that (C) because the belief
 238 is doxastically justified for S regardless of whether S has some other, auxiliary
 239 claims available that are themselves doxastically justified.

243 *The Entailment Proposal*

241 Another way to think about the *wishful thinking problem* is that there is a
 242 condition that needs to be met for the states (i) and (ii) (section 1) to simul-
 243 taneously obtain, and the pure non-cognitivist expressivist cannot satisfy it.
 244 As we saw, the constraint seems to be something like this: S's belief that *p*
 245 goes from not doxastically justified to being just that only if S acquires a new
 246 *cognitive state*. Call this *Dorr's constraint* (Mabrito 2013, 1072). Given this, the
 247 *decalogue problem* and *modified proposals* can be thought of as attempts to
 248 show that the non-cognitivist expressivist can meet *Dorr's constraint*.

249 This is not the only way to vindicate non-cognitive expressivism, though.
 250 One can also attempt to argue that S can meet the constraint and still be
 251 rational (Mabrito 2013). How might one show this? One idea is to say that
 252 while the Edgar case violates *Dorr's constraint*, it is compatible with another,
 253 independently motivated constraint (Mabrito 2013). The obedience of *this*
 254 constraint vindicates the intuitive rationality of Edgar coming to believe (C)
 255 in the LIAR ARGUMENT. Here is the argument again:

256 LIAR ARGUMENT

257 (P1) If lying is wrong, the souls of liars will be punished in the afterlife.

258 (P2) Lying is wrong.

259 (C) So, the souls of liars will be punished in the afterlife.

260 This other constraint is called the *entailment constraint* (EC). With respect to
 261 the *wishful thinking argument*:

262 ENTAILMENT CONSTRAINT (EC). S moves from T_1 (during which
 263 S lacks justification for believing that *p*) to T_2 (during which S has
 264 justification for believing that *p*) only if S comes to accept claims

that entail p or acquires evidence that supports p (Mabrigo 2013, 1074).

Edgar's case seems compatible with this because the *wishful thinking problem* assumes that the *Frege-Geach problem* is solved: it concedes that moral-descriptive *modi ponentes* are valid. So, while Edgar initially lacks justification for believing (C), he later comes to accept two claims that entail (C)—securing the obedience of the **ENTAILMENT CONSTRAINT**—and is thereby rational for believing (C) on the basis of them.

Again, though, we should think of how this solution goes in terms of the doxastic/propositional justification distinction. To illustrate, consider the **ENTAILMENT CONSTRAINT** itself. As a necessary condition, with respect to changes in what kinds of justification is the condition plausible? And obeying which of these various formulations also helps with the *wishful thinking problem*?

Suppose that the relevant change concerns a belief's status with respect to being propositionally justified:

ENTAILMENT CONSTRAINT 2 (EC₂). S moves from T_1 (during which S lacks propositional justification for believing that p) to T_2 (during which S has propositional justification for believing that p) only if S comes to accept claims that entail p or acquires evidence that supports p .

(EC₂) is plausible but irrelevant. No one denies that Edgar's belief that (C) of the **LIAR ARGUMENT** is *justifiable*—i.e., denies that there are reasons to accept it. The worry is that Edgar's belief seems *justifiedly inferred* from his belief in the premises. So, a case that obeys *this* constraint is explanatorily moot with respect to the relevant intuitive rationality that needs preservation on the *wishful thinking problem*.

Suppose that the salient change is from a belief being not *justifiable* to actually being *justified*. Thus:

ENTAILMENT CONSTRAINT 3 (EC₃). S moves from T_1 (during which S lacks propositional justification for believing that p) to T_2 (during which S has doxastic justification for believing that p) only if S comes to accept claims that entail p or acquires evidence that supports p .

298 (EC₃) is false. One can both believe that p for no reason and then only later
 299 base one's belief that p on very good reasons, and yet neither come to believe
 300 that which entails p nor acquire evidence for p . This happens in cases where
 301 one dogmatically believes that p but only later comes to accept p on the basis
 302 of good reasons that one already had. This is because one can fail to recognize
 303 good reasons for beliefs when they have them. Hence, one can fail to base
 304 that which they already believe on the basis of those good reasons.

2024 The Hopeful Proposal Proposal

306 Another proposed solution to show that wishful thinking is sometimes rational
 307 is accepting that the premises of arguments akin to moral-descriptive *modi*
 308 *ponentes* do make the belief in the conclusion justifiable (Long 2016). As an
 309 example, consider the following argument that S is reasoning through. Call it
 310 the *hopeful proposal argument*.

311 HOPEFUL PROPOSAL ARGUMENT

- 312 (P11) If I hope that my proposal will be accepted, then my proposal will be
 313 accepted.
 314 (P12) I hope that my proposal will be accepted.
 315 (C2) So, my proposal will be accepted.

316 This argument is *wishful thinking* as it is normally understood outside the
 317 seminar room: forming *beliefs* about how the world *is*, given one's wants,
 318 desires, hopes, dreams, etc. (which tell us how the world is *not*). Now, how
 319 does one argue that it is rational to infer (C2)? The idea is that "accepting all
 320 of the [...] premises is often a reason to accept its conclusion, since paradigm
 321 cases of wishful thinking are often valid" (Long 2016, 3).

322 Once again, the propositional/doxastic justification distinction renders this
 323 proposal ambiguous. So, for example, S's belief that (C2) is rational because
 324 (and to the extent that) either it is *doxastically* justified or *propositionally*
 325 justified (given the fact that *modus ponens* is valid).¹²

326 No one would deny that S coming to believe (C2) is justifiable, given the
 327 validity of *modus ponens*. The issue, recall, is whether S is justifiably rational
 328 in going on to believe (C2) *on the basis of* the premises—as it intuitively seems.

12 This assumes a solution to the Frege-Geach problem.

329 So, if this proposal is to work, the very fact of the argument's validity must be
 330 a part of that set of things on the basis of which S believes that (C₂).

331 Here is how this works with respect to the LIAR ARGUMENT. S rationally,
 332 justifiedly infers (C),

333 (C) So, the souls of liars will be punished in the afterlife;

334 when the belief that (C) is based on the following:

335 (P₁) If lying is wrong, the souls of liars will be punished in the afterlife;

336 (P₂) Lying is wrong; and

337 (P₁₃) (P₁) and (P₂) entail that the souls of liars will be punished in the afterlife.

338 The trouble is that this proposal misses the mark. The mere availability of
 339 (P₁₃) does not mean that S is doxastically justified in believing (C). S can fail
 340 to base their belief in (C) on (P₁₃). S would need to be shown to, in every case,
 341 in fact base their belief in (C) on the basis of (P₁) + (P₂) + (P₁₃).

342 Moreover, the *wishful thinking* problem remains. The non-cognitivist expres-
 343 sivist still cannot make sense of how S's belief that (C)—even in that case—is
 344 doxastically justified. For (P₂), on their view, is still a non-cognitive state. And
 345 S cannot, it seems, justifiedly believe (C) *on the basis of* (P₁) + (P₂) + (P₁₃) for
 346 that very reason: (P₂) is the wrong kind of thing for one to justifiedly believe
 347 something else on the basis of. It would help if (P₃) was itself sufficient for S to
 348 justifiedly believe (C) on the basis of. We saw attempts at this. But the problem
 349 with those attempts remains. It would have to be shown that in every case, S
 350 in fact believes the conclusion on the basis of the stuff that is both guaranteed
 351 to be present; would, in fact, make the belief doxastically justified if the belief
 352 was based on them; and are the right kinds of things to base beliefs on.

353 Conclusion

354 The wishful thinking problem seemed dead, the recipient of several fatal blows.
 355 I hope to have shown that the distinction between propositional/doxastic
 356 justification helps clarify the nature of the problem, the nature of proposed
 357 solutions, and why those solutions are dubious. Perhaps, then, it is premature
 358 to ignore the wishful thinking problem.*

* I received many great comments on this paper. I'd like to thank the audiences at the Philosophy in Progress conference at the University of Nottingham and the Syracuse Working Papers Workshop

Adam Patterson

 0000-0003-0351-3402

Syracuse University

apatters@sy.edu

References

- ALSTON, William P. 1985. "Concepts of Epistemic Justification." *The Monist* 68(1): 57–89, doi:[10.5840/monist198568116](https://doi.org/10.5840/monist198568116).
- DORR, Cian. 2002. "Non-Cognitivism and Wishful Thinking." *Noûs* 36(1): 97–103, doi:[10.1111/1468-0068.00362](https://doi.org/10.1111/1468-0068.00362).
- ENOCH, David. 2003. "How Noncognitivists Can Avoid Wishful Thinking." *The Southern Journal of Philosophy* 41(4): 527–545, doi:[10.1111/j.2041-6962.2003.tb00965.x](https://doi.org/10.1111/j.2041-6962.2003.tb00965.x).
- GUAN, Chengying. 2014. "The Wishful Thinking Problem for Non-Cognitivism: Does It Really Make Sense?" *Kriterion – Journal of Philosophy* 28(1): 30–44, doi:[10.1515/krt-2014-012804](https://doi.org/10.1515/krt-2014-012804).
- KORCZ, Keith Allen. 2000. "The Causal-Doxastic Theory of the Basing Relation." *Canadian Journal of Philosophy* 30(4): 525–550, doi:[10.1080/00455091.2000.10717542](https://doi.org/10.1080/00455091.2000.10717542).
- LENMAN, James. 2003. "Noncognitivism and Wishfulness." *Ethical Theory and Moral Practice* 6(3): 265–274, doi:[10.1023/a:1026037820411](https://doi.org/10.1023/a:1026037820411).
- LONG, Joseph. 2016. "Non-Cognitivism and the Problem of Moral-Based Epistemic Reasons: A Sympathetic Reply to Cian Dorr (2002)." *Journal of Ethics and Social Philosophy* 10(3), doi:[10.26556/jesp.v10i3.186](https://doi.org/10.26556/jesp.v10i3.186).
- MABRITO, Robert. 2013. "Are Expressivists Guilty of Wishful Thinking?" *Philosophical Studies* 165(3): 1069–1081, doi:[10.1007/s11098-012-0003-8](https://doi.org/10.1007/s11098-012-0003-8).
- PRYOR, James. 2005. "There is Immediate Justification." in *Contemporary Debates in Epistemology*, edited by Ernest Sosa and Matthias Steup, 1st ed., pp. 181–201. Contemporary Debates in Philosophy n. 3. Boston, Massachusetts: Blackwell Publishers. Second edition: Steup, Turri and Sosa (2014, 202–221).
- SCHROEDER, Mark. 2011. "How Not to Avoid Wishful Thinking." in *New Waves in Metaethics*, edited by Michael Sean Brady, pp. 126–140. New Waves in Philosophy. Basingstoke, Hampshire: Palgrave Macmillan, doi:[10.1057/9780230294899_7](https://doi.org/10.1057/9780230294899_7).
- SILVA, Paul, Jr. and OLIVEIRA, Luis R. G. 2024. "Propositional Justification and Doxastic Justification." in *The Routledge Handbook of the Philosophy of Evidence*, edited

for helpful comments and discussion. In particular, I'd also like to thank the following individuals for their useful feedback: Neil Sinclair, Preston Werner, Kris McDaniel, Hille Paakkunainen, Nate Sharadin, and Evelyn Hudson. Finally, I would also like to thank anonymous referees from the following journals for their thorough comments: *Dialectica*, *Analysis*, *Thought*, and the *Journal of Ethics & Social Philosophy*.

- 392 by Maria LASONEN-AARNIO and Clayton LITTLEJOHN, pp. 395–407. Routledge
393 Handbooks. New York: Routledge, doi:[10.4324/9781315672687-36](https://doi.org/10.4324/9781315672687-36).
- 394 SOSA, Ernest and STEUP, Matthias, eds. 2005. *Contemporary Debates in Epistemology*.
395 1st ed. Contemporary Debates in Philosophy n. 3. Boston, Massachusetts: Blackwell
396 Publishers. Second edition: Steup, Turri and Sosa (2014).
- 397 STEUP, Matthias, TURRI, John and SOSA, Ernest, eds. 2014. *Contemporary Debates in*
398 *Epistemology*. 2nd ed. Contemporary Debates in Philosophy n. 3. Malden, Mas-
399 sachusetts: Wiley-Blackwell. First edition: Sosa and Steup (2005), doi:[10.1002/97](https://doi.org/10.1002/9781394260744)
400 [81394260744](https://doi.org/10.1002/9781394260744).

PROOF

Whittle vs. Cantor on the Size of Infinite Sets

RAN LANZET

I examine several arguments by Bruno Whittle against the Cantorian conception of the size of infinite sets. I find that none of them succeed.

Sets A and B are the same size— A has exactly as many elements as B does—iff there is a bijection between them; A is at least as large as B —has at least as many elements as B does—iff there is an injection from B to A . This conception of set size has become standard following Cantor. What Cantor's theorem means, on this conception, is that the powerset of every set A is larger than A (see, e.g., Enderton 1977, 132; Hrbacek and Jech 1999, 65; Smullyan and Fitting 1996, 7–8). Thus interpreted, the theorem entails that there are different sizes of infinity (assuming, of course, that there is at least one infinite set and that that set has a powerset).

Whittle (2015a, 2015b, 2018) objects to this standard conception. He holds that we do not in fact have good reason for believing either of the following two principles, and that we are thus not in a position to know either of them or to know that there are infinite sets of different sizes:

SIZE \rightarrow FUNCTION

For any sets A and B , if A is the same size as B , then there is a bijection from A to B .

SIZE* \rightarrow FUNCTION

There is some “size-like” property—a property *similar* to size—size* such that, for any sets A and B , if A is the same size* as B , then there is a bijection from A to B .

Whittle puts forward a series of arguments that aim to establish this. Five of his arguments purport to refute the following widely accepted theses:

427 THESIS 1. There being a bijection between A and B is what it is for
428 A and B to be the same size.^{1,2}

429 THESIS 2. The notion of cardinality defined in set theory is at least
430 “size-like”: it is at least similar to our ordinary notion of size (of sets)
431 and perhaps a natural generalization of the notion of finite size.

432 THESIS 3. Cantor’s theorem establishes that there are infinite sets
433 of different sizes.³

434 Here, I examine Whittle’s arguments against theses 1 and 2; I argue that
435 none of them succeed as refutations of either of those theses. The arguments,
436 appearing in their fully developed form in Whittle (2018), are:

- 437 (i) an argument against THESIS 1 based on its interpretation as stating that
438 “ c is the same size as d ” and “There is a bijection from c to d ” express
439 the same structured proposition;
- 440 (ii) an argument against the same thesis, based on what would be true in a
441 mathematically-impossible situation in which there are no functions
442 from certain sets;
- 443 (iii) a Benacerraf-style challenge for accounts of size in terms of functions—
444 accounts that may be offered in support of THESIS 1;
- 445 (iv) an objection to THESIS 2 based on (ii).

446 Whittle also argues directly against THESIS 3 in his (2015a) and (2015b). I
447 examine and reject his argument against this thesis in (manuscript).

-
- 1 Whittle targets this thesis and THESIS 2 below since they can be offered as grounds for $\text{SIZE} \rightarrow \text{FUNCTION}$ and $\text{SIZE}^* \rightarrow \text{FUNCTION}$.
 - 2 Some may consider THESIS 1 too strong and prefer a more tolerant approach, on which the identification of size with cardinality-as-defined-in-set-theory is just one of several legitimate options. However, if this identification is indeed untenable, as some of Whittle’s arguments purport to establish, then it is unclear how one could legitimately adopt it—even tolerantly. I thank an anonymous referee for drawing my attention to this point.
 - 3 Whittle also attacks the following theses, which, I believe, are not as popular as the ones I mentioned in the main text: We can infer that $\text{SIZE} \rightarrow \text{FUNCTION}$ holds for sets in general from the fact that it holds for finite sets; we can infer $\text{SIZE} \rightarrow \text{FUNCTION}$ by Inference to the Best Explanation, since size differences are the only explanation of the absence of a bijection between two given sets; we are entitled to consider $\text{SIZE} \rightarrow \text{FUNCTION}$ as a basic mathematical truth; we can argue for $\text{SIZE} \rightarrow \text{FUNCTION}$ inductively, based on the consequences that it allows us to derive. I will not discuss Whittle’s arguments against those additional theses here.

448 If I am right in rejecting the arguments against theses 1–3, then Whittle
 449 fails to establish his claim that we are not in a position to know that there
 450 are different sizes of infinity. For these three theses, if correct, provide routes
 451 to such knowledge (this is, indeed, why Whittle attacks them); and this is so
 452 whether or not some additional routes are successfully blocked by those of
 453 Whittle’s arguments that I will not consider here.⁴

454 Although my treatment of Whittle’s arguments can be considered a defense
 455 of the Cantorian view of size, I will not offer positive arguments for this view
 456 or argue that other views, incompatible with it, are incorrect or unjustified;
 457 my point is only that Whittle fails to establish this for Cantor’s view.

458 Sections 1–4 below are each dedicated to the examination of one of the
 459 arguments (i)–(iv). I briefly conclude in section 5.

460 1 Structured Propositions

461 Consider the following two sentences:

462 SIZE
 463 *c* is the same size as *d*.

464 FUNCTION
 465 There is a bijection from *c* to *d*.

466 According to Whittle, the most straightforward interpretation of **THESES 1** is
 467 this:

468 (*) **SIZE** and **FUNCTION** express the same proposition.⁵

469 Whittle thinks of propositions here as structured, Russellian propositions.
 470 The propositions that **SIZE** and **FUNCTION** seem to express are, according to
 471 him,

472 (p_S) $\langle \exists!P[\text{Size}(P) \wedge c \text{ has } P] \wedge \exists!Q[\text{Size}(Q) \wedge d \text{ has } Q] \wedge$
 473 $\exists R[\text{Size}(R) \wedge c \text{ has } R \wedge d \text{ has } R] \rangle$

474 and

4 See footnote 3.

5 An alternative interpretation of **THESES 1** would be that **SIZE** and **FUNCTION** describe the same feature of reality. Whittle’s argument against **THESES 1** on *this* reading is the one considered in the next section.

475 $(p_F) \langle \exists f \text{ } f \text{ is a bijection from } c \text{ to } d \rangle$,

476 respectively.

477 Whittle now suggests that each of the steps in the following argument
478 against (*) is at least very plausible, or prima facie correct (Whittle 2018,
479 855–856):

- 480 (1) (p_S) and (p_F) are about very different sorts of things: (p_S) is about a
481 certain sort of property—a size—that c and d have in common; (p_F) is
482 about a function.
- 483 (2) If (*) is true, then either both SIZE and FUNCTION express (p_S) , or both
484 of them express (p_F) .
- 485 (3) Most sentences about functions express propositions that are genuinely
486 about functions.
- 487 (4) There is no plausible account on which the following sentence expresses
488 a proposition about properties:

489 CONSTANT

490 There is a constant function from c to d .

- 491 (5) The sentences FUNCTION and CONSTANT seem to be very similar.
- 492 (6) From (5): FUNCTION and CONSTANT should express similar
493 propositions—not ones about completely different sorts of things.
- 494 (7) From (1), (4), and (6); and also from (1) and (3): FUNCTION cannot
495 express (p_S) .
- 496 (8) SIZE has the same general form as the following:

497 COLOR

498 e is the same color as g .

499 HEIGHT

500 e is the same height as g .

501 SEX

502 e is the same sex as g .

- 503 (9) From (8): SIZE and sentences like COLOR, HEIGHT, and SEX should
504 express similar propositions—propositions of the same general form.

- 505 (10) The propositions expressed by sentences such as **COLOR**, **HEIGHT**, and
 506 **SEX** are not about functions.
- 507 (11) From (1), (9), and (10): **SIZE** cannot express (p_F).
- 508 (12) From (7) and (11): It is not the case that both **SIZE** and **FUNCTION**
 509 express (p_S), and it is not the case that both of them express (p_F).
- 510 (13) From (2) and (12): (*) is false.

511 Whittle's argument here relies on the following view of structured proposi-
 512 tions:

513 **PROP**

514 For each declarative sentence such as **SIZE** and **FUNCTION**, there is
 515 one unique structured proposition that the sentence can be taken to
 516 express; the structure of that proposition, as well as what it is about,
 517 is given by the most straightforward formalization of the sentence
 518 in higher-order predicate logic.

519 Without **PROP**, we seem to have no reason to accept the inference from (5) to
 520 (6) or the one from (8) to (9). As I will now argue, however, it is illegitimate
 521 to rely on **PROP** in this context. This is so, since **PROP** is incompatible with
 522 Whittle's reading of "what it is"—a reading on which

- 523 • There being a bijection between A and B is what it is for A and B to be
 524 the same size

525 is correctly interpreted as

- 526 • **SIZE** and **FUNCTION** express the same structured proposition.

527 **PROP**, together with this reading, yields absurd consequences, as illustrated
 528 by the following example.

529 Being a man who has never been married is just what it is to be a bachelor.
 530 But **PROP**, together with the aforementioned reading of "what it is," entails that
 531 this is not so. To see this, note first that the most straightforward formalizations
 532 in predicate logic of the following differ in structure:

- 533 (i) x is a bachelor.
- 534 (ii) x is a man who has never been married.

535 (i) and (ii) are most straightforwardly formalized as, respectively:

- 536 (iii) Bx
 537 (iv) $M^1x \wedge \neg \exists y(M^2xy)$.

538 This, by **PROP**, means that (i) and (ii) express different propositions. Hence,
 539 on Whittle's interpretation of "what it is," being a never-married man is *not*
 540 what it is to be a bachelor.

541 One can, of course, hold that, in formalizing (i), we should take into account
 542 the analysis of "bachelor" as "never-married man" and correspondingly allow
 543 the formalization of both (i) and (ii) as (iv). This is, I think, a reasonable
 544 approach, but note that it deviates from **PROP**. Moreover, if a similar deviation
 545 is allowed in the case of **SIZE** and **FUNCTION**, then Whittle's argument does
 546 not go through, as there is no longer reason to accept, e.g., the inference from
 547 (8) to (9).

548 Whittle's approach entails, then, that being a never-married-man is *not*
 549 what it is to be a bachelor. Similarly, it entails that:

- 550 • There being an isometry that transforms a figure c in the Euclidean
- 551 plane into a figure d is *not* what it is for c and d to be congruent.
- 552 • There being a formal proof of a first-order sentence ϕ from a first-order
- 553 theory T is *not* what it is for ϕ to be a theorem of T .

554 Additional, similar examples are not hard to produce.

555 I conclude that Whittle's argument from structured propositions fails by
 556 relying on two incompatible principles: **PROP** and a reading of "what it is," on
 557 which **THEESIS 1** is correctly interpreted as (*). If **PROP** is given up, then, as
 558 noted above, Whittle's argument seems not to go through; if the aforemen-
 559 tioned reading of "what it is" is given up, then the argument does go through,
 560 but its conclusion, (13), no longer means that **THEESIS 1** is false.

562 Sets in an Impossible Situation

562 **THEESIS 1**, recall, was this: There being a bijection between A and B is what
 563 it is for A and B to be the same size. This thesis can be supported by what
 564 Whittle calls a *functional account of size*: an account of size properties of sets
 565 in terms of functions. According to Whittle, however, no such account can be
 566 correct (Whittle 2018, 860–861).

567 Whittle specifically considers a functional account on which complex prop-
 568 erties are "Russellian propositions with gaps" and sizes are properties of the
 569 form

570 ⟨There is a bijection from $_$ to κ ⟩,

571 where κ is a von Neumann cardinal number (i.e., a von Neumann ordinal
 572 equipollent with none of its elements). This, of course, is not the only way to
 573 characterize size properties in terms of functions. A more natural functional
 574 account—and one that avoids commitment to a specific metaphysical account
 575 of properties—is, perhaps, this: A size-property is the property of belonging
 576 to a given bijection-type⁶ (cf. Whittle 2018, fn.15). Whittle explains, however,
 577 that his argument can be adapted so as to refute *any* functional account of
 578 size, and, in fact, any account of size in terms of things other than those that
 579 constitute the set, such as its elements or those elements' parts or elements.

580 Whittle's argument involves the sets $A = \{0, 1\}$ and $B = \{2, 3\}$ in a (mathe-
 581 matically impossible) situation S that is “exactly like the actual world, except
 582 that there are no functions from either set” (Whittle 2018, 860). The argument
 583 is this:

- 584 (1) In S , there are no bijections from A or from B .
 585 (2) In S , A and B have the same size.
 586 (3) From (1): In S , neither A nor B has any property of the form
 587 ⟨There is a bijection from $_$ to κ ⟩.
 588 (4) From (3): If the functional account under consideration is correct, then,
 589 in S , neither A nor B has a size.
 590 (5) From (4): If the functional account is correct, then, in S , A and B do *not*
 591 have the same size.
 592 (6) From (2) and (5): The functional account is incorrect.

593 There is a good reason to suspect this argument, if not to reject it outright:
 594 Arguments very similar to it, and ones that are not any less appealing, lead to
 595 absurd conclusions. Consider, for example, the following account of what a
 596 *circle* is:

597 CIRCLE

598 To be a circle is to be a set of all the points in the Euclidean plane
 599 that are at the same given distance r from a given point o .

600 I submit that this is a *correct* account of what a circle is. But an argument
 601 similar to Whittle's leads to the conclusion that this is not so: Let C be a circle,

6 A bijection type here is a class X for which the following condition holds: For some set a , X is the class of all possible sets x such that there is a bijection between x and a .

602 and consider an impossible situation S' that is exactly like the actual world
 603 except that there is no point at equal distances from all the elements (points)
 604 of C . (Reasoning about S' seems to make as much sense as reasoning about S
 605 does.) Then:

- 606 (1) In S' , no point is at equal distances from all the elements of C .
- 607 (2) In S' , C is a circle. (This seems as plausible as Whittle's premise that A
 608 and B have the same size in S .)
- 609 (3) From (1): If **CIRCLE** is correct, then, in S' , C is *not* a circle.
- 610 (4) From (2) and (3): **CIRCLE** is incorrect.

611 (Note that I am not claiming that this is a very convincing argument—just
 612 that it is similar to, and not less appealing than, Whittle's argument.)

613 This “bad company” indicates that something is wrong with Whittle's argu-
 614 ment. But what? There are, I believe, two major problems with the argument.
 615 First, it is not at all clear that we can make sufficient sense of mathematically
 616 impossible situations like S to determine the truth value of statements like
 617 Whittle's premise (2) (“In S , A and B have the same size”). Whittle claims that
 618 we *are* capable of judging what would be true in S , but it is far from clear
 619 that he is correct about that. Whittle also claims that an argument similar to
 620 his can be given not in terms of what is true under an impossible hypothesis
 621 but, instead, in terms of what is an immediate consequence of the hypothesis.
 622 If this is indeed possible, then the resulting (modified) argument would not
 623 be any more problematic than a standard *reductio*. It is unclear, however,
 624 how the imagined modification of the argument is supposed to proceed, and
 625 Whittle gives no indication of that. At least on the face of it, his argument is
 626 very much *unlike* a *reductio*: It is more similar, it seems, to a demonstration of
 627 non-entailment using a counterexample; for, rather than deriving an absurdity
 628 from an impossible hypothesis, he seems to be relying on judgments made
 629 under such a hypothesis in order to reject a universal statement (to the effect
 630 that certain properties always coincide).

631 Second, assuming (for the sake of argument) that we can make sufficient
 632 sense of mathematically impossible situations like S , it is unclear why we
 633 should think that A and B have the same size in S . Perhaps Whittle holds that
 634 this is so since (allegedly) A and B are still of size 2 in S . But it is unclear why
 635 we should think that this is so. Especially if, in S , *there is no way of counting*
 636 *the elements of A or those of B* , and there is indeed no way of counting those
 637 elements in S , at least if the following jointly hold: (a) a way of counting

638 the elements of, e.g., A is a way of correlating them 1-1 with the elements of
 639 {"one", "two"}; (b) a way of thus correlating the elements of these two sets is
 640 a *bijection* between them; (c) in S , there are no bijections between A and any
 641 other set.

642 Perhaps Whittle holds that, e.g.,

643 SIZE2

644 The size of A is 2 in *every* situation, mathematically possible or not,
 645 in which A exists.⁷

646 But this, it seems, cannot serve as a ground for Whittle's premise (2), since an
 647 argument similar to his leads to the conclusion that SIZE2 is *false*: Consider
 648 a situation S'' that is exactly like the actual world, except that there are no
 649 numbers other than 0 and 1. (Reasoning about S'' seems to make as much
 650 sense as reasoning about S does.) Then:

- 651 (1) It is not the case that, in S'' , something equals 2.
 652 (2) If SIZE2 is true, then, in S'' , the size of A equals 2.
 653 (3) From (2): If SIZE2 is true, then, in S'' , something equals 2.
 654 (4) From (1) and (3): SIZE2 is false.

655 I am not claiming that (1)–(4) is a convincing argument or that reasoning in
 656 this way about S'' makes any sense; I only claim that this is so *by Whittle's stan-*
 657 *dards*. If this claim is correct, then Whittle cannot rely on SIZE2; consequently,
 658 premise (2) of his argument remains unfounded.

659 I conclude that Whittle's argument from S fails as a refutation of functional
 660 accounts of size and that it is therefore ineffective against THESIS 1.

663 Benacerraf's Problem

662 Whittle (2018, 862–863) argues that the specific functional account that fea-
 663 tured in his argument from an impossible situation faces a version of "Be-
 664 nacerraf's problem."⁸ That functional account, recall, identified sizes with
 665 properties of the form

7 He may think that this is so, since the following (presumably) holds in any situation in which A exists: $\exists u \exists v (u \neq v \wedge u \in A \wedge v \in A \wedge \forall w [w \in A \rightarrow (w = u \vee w = v)])$.

8 The original problem, presented in Benacerraf (1965), arises for theories that take the natural numbers to be a particular collection of sets.

666 ⟨There is a bijection from $_$ to κ ⟩,

667 where κ is a von Neumann cardinal number. The problem is this: By replacing
668 the cardinal numbers with other, equipollent sets, we can obtain different,
669 incompatible functional accounts of size. Since, moreover, there is no reason
670 to favor any one of those competing accounts over the others, each of them
671 seems arbitrary and therefore, according to Whittle, incorrect.

672 This problem does not seem to afflict *all* functional accounts of size. Con-
673 sider, for instance, the account I mentioned in the previous section, on which
674 a size property is the property of belonging to a given bijection type. This ac-
675 count does not commit to a particular metaphysical theory of what properties
676 are, and the Benacerraf problem, it seems, does not arise for it. Whittle would
677 seem to agree, but he does not consider this to be a problem for his attack
678 on [THESIS 1](#), since he takes the functional accounts immune to Benacerraf’s
679 problem to be refuted by the argument discussed in the previous section (see
680 his [2018, fn.15](#)). If my arguments in the previous section are correct, how-
681 ever, then there is a genuine problem here for Whittle; for the argument he is
682 relying on fails.

684 **4 Are Cardinalities Size-Like?**

684 [THESIS 2](#) was this:

685 The notion of cardinality defined in set theory is at least “size-like”: it
686 is at least similar to our ordinary notion of size (of sets) and perhaps
687 a natural generalization of the notion of finite size.

688 Whittle takes cardinalities to be properties of the form

689 ⟨There is a bijection from $_$ to κ ⟩,

690 where κ is a cardinal number. He makes two related points against [THESIS 2](#)
691 ([Whittle 2018, 864](#)):

- 692 1. The notion of cardinality is not a generalization of the notion of finite
693 size, since the collection of all cardinalities does not contain the finite
694 sizes. This is allegedly established by the argument from [S](#) (discussed in
695 section [3](#) above), as that argument is supposed to show that finite sizes,
696 unlike finite cardinalities, cannot be understood in terms of bijections.

- 697 2. Cardinalities are “just a completely different sort of property from sizes”
 698 (Whittle 2018, 864). For, sizes have nothing in particular to do with func-
 699 tions; they are, rather, *intrinsic*: they can be accounted for only in terms
 700 of the things that constitute the set, such as its elements, their parts,
 701 or their elements. This is supposed to be established by an argument
 702 similar to the argument from *S*.

703 There are, however, two serious problems with this line of argument. First, if
 704 my criticism of Whittle’s argument from an impossible situation (see section 3
 705 above) is correct, then that argument fails to establish that finite sizes are
 706 distinct from finite cardinalities, and it is unclear how a similar argument
 707 could establish the supposed intrinsic nature of size properties.

708 Second, even if the argument from *S* did establish that sizes are intrinsic
 709 and that finite sizes are thus distinct from finite cardinalities, it would still *not*
 710 follow that sizes are not *size-like*—i.e., not similar to sizes—in any important
 711 or interesting ways, or that it is not obvious that they are. For, whether or not
 712 sizes are intrinsic, there are several well-known points of similarity between
 713 cardinalities and finite sizes that are, arguably, both interesting and important.
 714 These include the following:

- 715 (i) Finite cardinalities are at least *co-extensive* with finite sizes (Whittle
 716 concedes this).
 717 (ii) If a subset *B* of *A* is smaller* than *A* (in the sense of cardinalities), then
 718 it is a *proper* subset of *A*; more generally, the pigeonhole principle holds:
 719 If a set *B* is smaller* than *A*, then there is no injection from *A* to *B*.
 720 (iii) The following version of Hume’s principle (which Whittle accepts for
 721 finite sizes) holds: Sets *A* and *B* have the same size* iff there is a bijection
 722 between them.
 723 (iv) Assuming the axiom of choice, sizes* (i.e., cardinalities) are well-
 724 ordered by the relevant smaller-than* relation.


725 Given these problems, I submit, Whittle’s points against [THESIS 2](#) do not
 726 suffice as a refutation of it.

725 5 Conclusion

728 Whittle puts forward four arguments against theses 1 and 2; these, as I ex-
 729 plained in the introduction, constitute an essential component of his objection

730 to Cantor's conception of infinite size. As I hope to have established, however,
731 none of Whittle's arguments against those two theses succeed.*

Ran Lanzet

 0000-0003-2475-562X

University of Haifa

lanzetr@gmail.com

736 References

- 737 BENACERRAF, Paul. 1965. "What Numbers Could Not Be." *The Philosophical Review*
738 74(1): 47–73, doi:10.2307/2183530.
- 739 ENDERTON, Herbert B. 1977. *Elements of Set Theory*. New York: Academic Press.
- 740 HRBACEK, Karel and JECH, Thomas J. 1978. *Introduction to Set Theory*. 1st ed. Pure
741 and Applied Mathematics n. 45. New York: Marcel Dekker.
- 742 —. 1999. *Introduction to Set Theory*. 3rd ed. Pure and Applied Mathematics n. 220.
743 New York: Marcel Dekker. First edition: Hrbacek and Jech (1978).
- 744 MCGEE, Vann. 2015. "Whittle's Assault on Cantor's Paradise [on Whittle (2015a)]." in
745 *Oxford Studies in Metaphysics*, volume IX, edited by Karen BENNETT and Dean W.
746 ZIMMERMAN, pp. 20–32. New York: Oxford University Press, doi:10.1093/acprof:
747 oso/9780198729242.003.0002.
- 748 SMULLYAN, Raymond M. and FITTING, Melvin Chris. 1996. *Set Theory and the Continuum Problem*.
749 Oxford Logic Guides n. 34. Oxford: Oxford University Press. second
750 edition: Smullyan and Fitting (2010).
- 751 —. 2010. *Set Theory and the Continuum Problem*. 2nd ed. Mineola, New York: Dover
752 Publications. First edition: Smullyan and Fitting (1996).
- 753 WHITTLE, Bruno. 2015a. "On Infinite Size." in *Oxford Studies in Metaphysics*, volume
754 IX, edited by Karen BENNETT and Dean W. ZIMMERMAN, pp. 3–19. New York:
755 Oxford University Press, doi:10.1093/acprof:oso/9780198729242.003.0001.
- 756 —. 2015b. "Reply to McGee (2015)." in *Oxford Studies in Metaphysics*, volume IX, edited
757 by Karen BENNETT and Dean W. ZIMMERMAN, pp. 33–41. New York: Oxford
758 University Press, doi:10.1093/acprof:oso/9780198729242.003.0003.
- 759 —. 2018. "Size and Function." *Erkenntnis* 83(4): 853–873, doi:10.1007/s10670-017-
760 9917-z.

* For helpful comments and discussions I thank Balthasar Grabmayr, David Kashtan, Aviv Keren, Eli Pitcovski, Gil Sagi, and four anonymous referees. The research leading to this paper was supported by post-doc fellowships at the University of Haifa and at the Hebrew University of Jerusalem.

Functionalism, Pluralities, and Groups

EMILIE PAGANO

762 It's widely accepted that *pluralism* about groups—the view that groups
 763 are pluralities—is incompatible with the following: one group can have
 764 different individuals as members at both different times and in differ-
 765 ent worlds (Difference), and more than one group can have the same
 766 individuals as members at both the same times and in the same worlds
 767 (Sameness). As a result, it's widely accepted that pluralism is false. In
 768 this paper, I argue that these “arguments from Difference and Sameness”
 769 are unsound. First, I articulate a functionalist account of what it is to be a
 770 group that's neutral with respect to pluralism and its primary opponent,
 771 *monism*. According to the version of functionalist pluralism I propose,
 772 groups are pluralities of functional roles. Second, I argue that because
 773 different individuals can play a role at both different times and in differ-
 774 ent worlds, and because the same individuals can play different roles at
 775 both the same times and in the same worlds, functionalist pluralism is
 776 invulnerable to the arguments from Difference and Sameness. Lastly, I
 777 raise a challenge for functionalist monism: whereas functionalism seems
 778 to favor “external” individuation conditions, monism seems to favor “in-
 779 ternal” individuation conditions, and it's up to the functionalist monist to
 780 square them. In the process, I hope to have shown that functionalism—
 781 whether pluralistic or monistic—is worthy of our attention.

782 Groups are everywhere. We rely on them when we marry, matriculate, and
 783 mortgage, when we pray, purchase, and patronize, and when we lend, loot,
 784 and lecture. They systematically guide our interactions. They have members,
 785 they do things, and they come and go. They matter. As a result, the question
 786 arises: What are they?

787 As with “What is X?” questions generally, this one's ambiguous. On the
 788 one hand, to ask what a group is is to ask what it is *to be* a group. That's
 789 a question about what defines the kind GROUP. It's asking: What is it for a

particular collection of individuals to be a group rather than a mere collection of individuals? Call it *the definitional question*.¹

On the other hand, to ask what a group is is to ask what kinds of *things* groups are. That isn't a question about what defines the things we call groups qua groups. It's asking: What, say, instantiates GROUP? Groups are the kinds of things that fit into churches, corridors, and courtrooms. And it might be—indeed, it's quite plausible—that GROUP doesn't fit into churches, corridors, or courtrooms. As a result, we'll want to know what does. Call this *the ontological question*.^{2,3}

Interestingly, philosophers have nearly universally privileged the ontological question. Predictably, there's significant disagreement among them. Nonetheless, the vast majority of philosophers accept that groups are “one,” that they're fundamentally singular things (e.g., [Effingham 2010](#); [Ritchie 2013, 2015, 2020](#); [Hawley 2017](#); and [Fine 2020](#)).⁴ I call their view *monism*. Nonetheless, a minority of philosophers accept that groups are “many,” that they're fundamentally plural “things” (i.e., pluralities) ([Uzquiano 2018](#); [Faller 2021](#);

1 I have “real” rather than linguistic definition in mind. See [Dorr \(2016\)](#), [Correia \(2017\)](#), and [Rosen et al. \(2018\)](#) for discussion.

2 Should one be disinclined to accept this distinction—in particular, because one is inclined to accept that GROUP is defined by the kinds of things groups are, or conversely—here's an argument:

- (i) Groups exist.
- (ii) It's possible that nothing defines GROUP.
- (iii) If it's possible that nothing defines group, GROUP isn't defined by the kinds of things groups are; that is, the kinds of things that exist.
Therefore, GROUP isn't defined by the kinds of things groups are (i.e., the kinds of things that exist).

Of course, one might deny (ii). But the point is: whether it's possible that GROUP is indefinable is an open question, and, so, one can't deny (ii) without argument.

- 3 Though the distinction between the definitional and ontological questions is familiar, my conception resembles a distinction metaethicists draw between value and its “bearers.” Again, we might ask what it is to *be* good, what it is that defines GOOD. That's a definitional question. However, we might ask what kinds of things are good, what, in other words, “bears goodness.” That's an ontological question. (See [Lemos 2005](#) for an excellent discussion.) Clearly, it's possible that GOOD is indefinable; a whole slew of philosophers accept it. Nonetheless, things are good, things bear value. As a result, GOOD isn't defined by its bearers; that is, by the kinds of things that are good.
- 4 It's important to note that [Epstein \(2015, epstein_b:2019a\)](#) and [Thomasson \(2019\)](#) accept “hybridism” about groups, according to which some groups are one and some groups are many. In what follows, I set hybridism aside, in particular, because problems for monism *are* problems for hybridism. If groups aren't one, they're not one *and* many. Rather, they're many.

806 Horden and López de Sa 2021; and Wilhelm 2022). They accept that a group
807 is in some sense its members, that they *are* “it.”⁵ I call their view *pluralism*.

808 In general, monists accept the following argument against pluralism:

- 809 (1) A group can have different members at both different times and in
810 different worlds. (Difference)
- 811 (2) If groups are pluralities, groups can’t have different members at either
812 different times or in different worlds.
813 *Therefore*, groups aren’t pluralities.
- 814 (3) According to pluralism, groups are pluralities.
815 *Therefore*, pluralism is false.

816 Moreover, they accept:

- 817 (4) Different groups can have the same members at both the same times
818 and in the same worlds. (Sameness)
- 819 (5) If groups are pluralities, different groups can’t have the same members
820 at either the same times or in the same worlds.
821 *Therefore*, groups aren’t pluralities.
- 822 (6) According to pluralism, groups are pluralities.
823 *Therefore*, pluralism is false.

824 I call these “the arguments from Difference and Sameness.” In this paper, I
825 argue that they’re unsound. Both (2) and (5) are false.

826 Admittedly, my argument takes some—ultimately necessary—twists and
827 turns. Here’s how it’ll go. In section 1, I articulate a functionalist account of
828 groups as an answer to the definitional question. In section 2, I argue that
829 functionalism is neutral with respect to both monism and pluralism. In sec-
830 tion 3, I argue that by obscuring the definitional question, the arguments from
831 Difference and Sameness assume a particularly naïve version of pluralism,
832 and show that an attractive version of functionalist pluralism is invulnerable
833 to them. Lastly, in section 4, I raise a challenge for functionalist monism:

5 Of course, that’s an odd way of speaking. However, I have roughly what Baxter (1988) has in mind when he suggests that “a” group—a *loosely* singular entity—is identical to its members—a *strictly* plural entity (i.e., a plurality). In other words, though we might *speak* of groups as if they’re strictly one—we tend to say that *it* nominated me rather than that *they’ve* nominated me, for instance—that we do follows from, say, our willingness to *count* the department as something over and above its members; that is, from strictly *practical* considerations. And that’s quite compatible with its being strictly many.

834 whereas functionalism seems to favor “external” individuation conditions,
 835 monism seems to favor “internal” individuation conditions, and it’s up to
 836 the functionalist monist to square them. I don’t claim that my challenge is
 837 dispositive, however. Rather, I claim that it exposes an important source of
 838 disagreement that’s worth pursuing.

839 Although the implications for monism and pluralism are clear, one of my
 840 aims is to generate interest in functionalism, whether monistic or pluralistic.
 841 Although functionalism about social goings-on—specifically, about artifacts—
 842 has pedigree, its application to groups hasn’t been explored.⁶ There are details
 843 to sort out, of course. And though I’ll make suggestions as I go, ultimately, I
 844 hope to have provided a framework for sorting them out that’s worthy of our
 845 attention.

846 **1 The Definitional Question: Functionalism**

847 As an answer to the definitional question, I propose a functionalist account of
 848 groups. Functionalism about social goings-on is an established view; in partic-
 849 ular, functionalism about artifacts (Searle 1995, 2010; Baker 2007; Thomasson
 850 2019; Eynine 2016; Guala 2016; Koslicki 2018). However, it hasn’t been pur-
 851 sued as an account of what groups are. But as Haslanger (1995) suggests, we
 852 might think of groups as special kinds of artifacts, as products of some of
 853 the things we do, whether intentionally or unintentionally. We might think
 854 of them as things we in some sense “use” to do them. And this suggests we
 855 might expand functionalism to account for them. In this section, I give it a try.

856 First, a note. Generally, philosophers think of artifactual functions teleo-
 857 logically, as things that serve purposes artificers impose on their products.
 858 However, I take my cue from functionalism about mental states, according to
 859 which functions aren’t teleological but, e.g., causal. Ultimately, that’s a choice
 860 point. One might translate the account I propose in terms of teleological func-
 861 tions (see Thomasson 2019 for rumblings). As a result, it’s easily assimilated
 862 into the wider literature.

6 Sometimes Ritchie (2013)’s account seems like a version of functionalism about groups, as when she claims that “[a group] structure is realized when each of its functionally defined nodes or places are occupied” (Ritchie 2013, 257). Nonetheless, she doesn’t pursue functionalism about groups per se.

1631 *Functionalism*

864 In general, functionalism about x is the view that x is a functional kind. A kind,
 865 K , is a functional kind when something is an K because of its extrinsic rather
 866 than its intrinsic properties; specifically, what it does—the way it functions—
 867 within a particular system in which it's embedded. Ordinarily, functionalism
 868 is associated with a particular account of mental states.⁷ Functionalists about
 869 mental states accept that kinds of mental states (e.g., PAIN) are defined by
 870 functions (e.g., to avoid physical harm) that are realized⁸ by whatever plays
 871 the relevant roles (e.g., the “pain role”) within a particular cognitive system.
 872 Generally, they accept that PAIN's realizations are pains.

873 Functionalism's claim to fame is the ease with which it accommodates multiple
 874 realizability. Multiple realizability is a feature of mental states whereby
 875 a single mental state might be grounded in multiple non-mental (e.g., physical)
 876 states. In creatures with cognitive systems like ours, PAIN is realized
 877 by c-fibers that cause the relevant kinds of responses; flinches, winces, and
 878 wails, for instance. In extraterrestrials with cognitive systems unlike ours,
 879 however, PAIN might be realized by gunky she-fibers that cause relevantly
 880 similar responses; shlinches, shinces, and shails, for instance. But both are
 881 pains because pains *are* what function that way.

882 Although there are many details about which functionalists disagree, I'll
 883 assume that PAIN is a higher-order property of the form *an-input-linking-*
 884 *an-output*, where the relevant kind of linking is causal.⁹ A particular pain
 885 might be a-pinch-causing-a-flinch within a particular cognitive system or a
 886 shinch-causing-a-slinch in theirs. But that's merely for convenience. For the
 887 important thing is this: to be a pain is to do what pains do. As a result—and
 888 with relevant disagreements aside—I propose

889 GROUP FUNCTIONALISM (FUNCTIONALISM). For xx to be a group,
 890 is for

7 Of course, there are many varieties of functionalism. However, the variety of functionalism I've described is “commonsense,” represented in, e.g., Block (1978), Fodor (1968), Lewis (1980), and Putnam (1975).

8 Realization is a kind of non-causal determination that's figured in influential arguments for functionalism. See Bickle (2020) for an excellent survey.

9 Regrettably, I can't discuss the many—and radically complex—differences in functionalist accounts of mental properties here. But, ultimately, the relevant differences won't jeopardize the view I propose. In fact, they're bound to enrich it.

- 891 i. *xx* to be an instance of a group kind, *K*, and for
 892 ii. *xx* to serve a function that defines *K*,
 893 iii. within a particular social system

894 where *xx* is either a singular or a non-singular plurality. Here, I argue that
 895 **FUNCTIONALISM** provides an attractive answer to the definitional question
 896 because it does what an account of groups should do; it satisfies several
 897 desiderata.

898 First, a clarification. One might worry that in appealing to group kinds,
 899 **FUNCTIONALISM** is circular. Group kinds are *group* kinds, and one might
 900 reasonably insist that **GROUP** can't be defined by them.¹⁰ (See Bealer 1997 for
 901 the corresponding objection about mental states.) Here, then, is a conception
 902 of group kinds I'll assume throughout:

903 **GROUP KIND FUNCTIONALISM.** *K* is a group kind iff *K* is

- 904 iv. a type of collection of individuals, *C*, such that
 905 v. the individuals comprising *C* realize roles, *rr*, that
 906 vi. give them reasons to act as members of *C*

907 where, again, *rr* is either a singular or non-singular plurality. Clearly, **FUNC-**
 908 **TIONALISM** plus **GROUP KIND FUNCTIONALISM** isn't circular. **GROUP KIND**
 909 **FUNCTIONALISM** says nothing about **GROUP**. Rather, it says that a particu-
 910 lar collection of individuals—Jonathan, Jennifer, and Julia, say—is, e.g., a
 911 department of philosophy because they realize roles that push them around
 912 in particular ways—in ways characteristic of departments of philosophy. In
 913 other words, their being a group isn't what makes them a department. Rather,
 914 their being a department is what makes them a group. Of course, **GROUP**
 915 **KIND FUNCTIONALISM** assumes we understand what roles are. But because
 916 everyone owes us an account of them, that's okay. As a result, I'll carry on as
 917 planned.¹¹

10 For instance, one might insist that particular group kinds are defined by **GROUP** as particular pains might be defined by **PAIN**. But that's in tension with multiple realizability. For if this or that mental state is essentially a pain, anything that's in that mental state has to be in pain. But that's precisely what the functionalist denies. Because mental states are multiply realizable, something might be in that mental state and not be in pain because of the nature of its cognitive system more generally. That's why pains are defined by their extrinsic rather than their intrinsic features.

11 Alternatively, one might use "group kind" ostensibly as a placeholder for whatever things are relevantly like departments, courts, and baseball teams. Again, I take my cue from functionalism about mental states, and it's plausible that we can say that **MENTAL STATE** is defined by "mental

1.2 *Desiderata*

919 **FUNCTIONALISM** satisfies several desiderata. In particular, it accounts for

- 920 (a) the distinction between groups and mere collections of individuals;
- 921 (b) the fact that groups and the individuals that are their members might
- 922 malfunction—alternatively, might err¹²—as the kinds of things they
- 923 are; and
- 924 (c) the fact that groups are embedded in wider social systems.

925 That it does is reason to take it seriously.

926 As for (a): again, **FUNCTIONALISM** is the view that groups are what they are
 927 because they do what they do.¹³ A paradigmatic group like the Department
 928 of Philosophy is the group it is because the individuals that are its members
 929 in some way “do philosophy” within a particular college, within a particular
 930 university, and as part of a particular department of education.¹⁴ Whatever
 931 their doing what they do amounts to, the Department of Philosophy is the
 932 particular group it is because the individuals that are its members do what
 933 they do. It’s a department of philosophy, an instance of DEPARTMENT OF
 934 PHILOSOPHY, and, so, it’s a group.

state kinds” without circularity. We know which responses are pain responses, and so we can appeal to them, whatever they are, in defining PAIN. (See Shoemaker 2003; McCullagh 2000; Tooley 2001 for this kind of response.) Similarly, we know which collections of individuals are, e.g., baseball teams, and so we can appeal to them, whatever they are, in defining BASEBALL TEAM, too.

- 12 The difference between malfunctioning and erring corresponds to the difference between teleological and nonteleological (e.g., causal) functions. Again, I’ll set that aside and mark it as a choice point.
- 13 We might worry about understanding groups in terms of their “doing things.” For there’s a question about how to understand the relation between groups and what they do, such that groups don’t disappear when their members *don’t* do it. For instance, it’s an open question whether departments of philosophy exist when their members don’t do philosophy because, e.g., they’re on strike, funding is suspended, it’s temporarily out of members. But this isn’t a special problem for **FUNCTIONALISM**. Whether groups are defined in terms of what they do, everyone is on the hook for providing an account of how they persist through changes like these, if they do.
- 14 There’s a question about how to define the relevant group kinds when some of the department’s members—its secretaries, say—*don’t* do philosophy in any straightforward sense. There’s a relatively straightforward solution to this problem, however. There’s a slew of ways to play the roles that define particular group kinds. For instance, although secretaries might not give seminars, what they do *enables* professors to give seminars. As a result, defining group kinds in terms of relatively course-grained functions like this one isn’t in itself problematic.

935 And this feature of **FUNCTIONALISM** underwrites a plausible explanation
 936 of the distinction between groups and mere collections of individuals. It's
 937 uncontroversial that groups, in some sense, consist of collections of individu-
 938 als. The Department of Philosophy, in some sense, consists of the individuals
 939 that are its members, of Jonathan, Jennifer, and Julia. Similarly, the Supreme
 940 Court, in some sense, consists of the individuals that are its members, of
 941 Roberts and his colleagues, and the Boston Red Sox of the individuals that
 942 are its members, of Martinez and his teammates.

943 But the Department of Philosophy, the Supreme Court, and the Red Sox are
 944 relevantly unlike collections of individuals like those wearing white t-shirts,
 945 those driving rental cars, and those that prefer chocolate to vanilla ice cream.
 946 Intuitively, whereas groups like the Department of Philosophy, the Supreme
 947 Court, and the Red Sox are such that the individuals that are their members do
 948 what they do *because they're members*, there's nothing the collection of individ-
 949 uals wearing white t-shirts do because they're "members." Of course, there's
 950 something they do: they wear white shirts. Unlike the groups in question,
 951 however, they're its members *because they're wearing white shirts*. In other
 952 words, whereas the individuals that are a group's members play particular
 953 kinds of roles—and so can act in their capacity as members—mere collec-
 954 tions of individuals don't and can't. And, again, that's what **FUNCTIONALISM**
 955 implies: because groups are individuated by what they do, the Department of
 956 Philosophy is a group because the individuals that are its members play the
 957 roles that define it, and not conversely.

958 As for (b): **FUNCTIONALISM** explains how and why groups and the individ-
 959 uals that are their members might malfunction as the kinds of things they
 960 are. Groups are like thermostats. Thermostats are what measure temperature.
 961 There's something they do and with respect to which they might fail. If they
 962 do, they're bad thermostats. Similarly, departments of philosophy are what
 963 do philosophy in the relevant ways. Like thermostats, there's something they
 964 do with respect to which they might fail. If they do, they're bad departments
 965 of philosophy.¹⁵

966 As for (c): **FUNCTIONALISM** explains how and why particular groups are
 967 embedded in wider social systems. Again, particular pains are what protect the

15 When we say the Department of Philosophy does philosophy badly, then we mean that it's failed as a department of philosophy. Though it's failing to be, e.g., collegial is grounds for criticism, this isn't a failure qua department of philosophy. On the contrary, its failing to do philosophy collegially is a kind of normative failure. But its normative failures don't entail that it fails to be a department of philosophy altogether.

968 pained, are causing's to avoid hot stoves, stubbed toes, and unfriendly blows.
969 This entails that they function within wider cognitive systems that enable the
970 relevant kinds of behaviors; in other words, that there are pained things—
971 organisms or otherwise—to protect. They're parts of cognitive systems, are
972 what function to avoid the physical harms that might befall the things they
973 cause to behave in the relevant ways. In a word: no things, no cognitive
974 systems, no pains.

975 Analogously, particular departments of philosophy are what do philosophy
976 in the relevant ways, what in some sense account for the fact that the individu-
977 als that are their members give particular kinds of seminars, invite particular
978 kinds of speakers, and host particular kinds of events. Again, this entails that
979 particular departments of philosophy function within wider social systems
980 that enable these kinds of behaviors. Again, the Department of Philosophy
981 is the group it is because its members do philosophy as part of a particular
982 college that's part of a particular university that's part of a particular depart-
983 ment of education, and these ultimately underwrite, say, its seminar offerings.
984 These groups require that the Department follow a particular curriculum and,
985 so, ultimately constrain how the Department of Philosophy does philosophy.

986 Relatedly, **FUNCTIONALISM** explains how and why the social systems in
987 which particular groups are embedded are *structured*. What social structures
988 are is an important and underexplored issue in contemporary philosophy. But
989 a few things are clear.

990 Minimally, structures are arrangements. They're complexes of relations.
991 They consist of "positions" at the intersections of these relations, and things
992 "occupy" them. Naturally, then, group structures are social arrangements;
993 complexes of social relations that consist of intersecting positions things
994 occupy. Baseball teams are structured, for instance. Every baseball team has a
995 pitcher and a catcher. In other words, a baseball team's structure partly consists
996 of particular asymmetric, irreflexive, and non-transitive relations—*pitches to*
997 *and returns to*, say—whose relata intersect in a particular way: pitchers pitch to
998 catchers and catchers return to pitchers. And that the relevant elements—*this*
999 *and that* individual—occupy the relevant positions—that they're arranged in
1000 the relevant way—partly grounds the fact that they're part of a baseball team
1001 rather than a mere heap of individuals.

1002 More than this, though, that a particular group functions in the particular
1003 ways it does is explained by the general arrangements of its elements. For
1004 instance, the Red Sox and the individuals that are their members play roles

1005 that are determined by the wider social system in which they're embedded.¹⁶
 1006 Like the Department of Philosophy and the Supreme Court, in other words,
 1007 the Red Sox are part of other groups; for instance, the American League East
 1008 (ALE) and Major League Baseball (MLB).¹⁷ And this wider social system
 1009 partly determines in which relations the individuals that are members of the
 1010 Red Sox stand both to one another and to other groups. For instance, we can't
 1011 explain the fact that the Red Sox won the World Series in 2018 by appealing to
 1012 how well they played. Rather, we must appeal to the relational fact that they
 1013 played *better* than the Los Angeles Dodgers—themselves part of the National
 1014 League West (NLW) and, so, the MLB—and to the rules that legitimated their
 1015 win. In other words, we must appeal to the social system in which both the
 1016 Red Sox and the Los Angeles Dodgers are embedded—the MLB—to explain
 1017 important facts about them. As a result, it's not merely that particular groups
 1018 are embedded in wider social systems, the social systems in which they're
 1019 embedded structure them. And, again, **FUNCTIONALISM** bears this out. (See
 1020 [Haslanger 2000](#) for a similar thought.)

1021 As a result, **FUNCTIONALISM** satisfies desiderata (a)–(c) and thereby pro-
 1022 vides a powerful framework for thinking both about what it is to be a group
 1023 per se and what it is to be a group of a particular kind. But because it's a sketch
 1024 of an account, there are details about which we might meaningfully disagree.
 1025 I'll consider a particularly important detail about which we might disagree in
 1026 section 3. But there's more.

1123 *Open Questions*

1028 Here's a brief survey. We might disagree about what a social system is and
 1029 about what it is for a complex of relations to constitute a structure. Both
 1030 Ritchie (2013, 2015, 2020) and Haslanger (2016) understand both in terms of
 1031 Shapiro (1997)'s influential conception of mathematical structure according
 1032 to which

16 Importantly, **FUNCTIONALISM** implies that, at any time and in any world, there's a social system that is structured by social relations that *aren't* embedded in any larger social structure. Maximal social systems—like “maximal” cognitive systems—can't be realized more than once at any time or in any world aside from trivial permutations of the same roles within a single group. As a result, every non-maximal social system there is part of a larger social system whose elements are ultimately structured by the most basic social relations there are.

17 I'm assuming that the MLB is itself a group with other groups as “parts.” Following Ruben (1985), I take it that being a member of a group and being part of a group are different relations. In what follows, however, I overlook this distinction.

1033 a system [is] a collection of [entities] with certain relations. [...]
1034 A structure is the abstract form of a system, highlighting the
1035 interrelationships among the [entities], and ignoring any features
1036 of them that do not affect how they relate to other [entities] in
1037 the system. (Shapiro 1997, 73–74)

1038 However, we needn't accept Shapiro's conception of structure to accommodate
1039 this feature of **FUNCTIONALISM**. What's important is that we accept that
1040 the social systems in which groups are embedded are holistic. And, again,
1041 **FUNCTIONALISM** bears this out.

1042 Moreover, we might disagree about whether particular individuals are
1043 unified in coming to be elements in a functionalist structure or whether
1044 it's sufficient that the relevant network of social relations is interdependent.
1045 (Indeed, I discuss its importance in section 4.) Relatedly, we might disagree
1046 about what the relevant functions are and what it is to realize them. For
1047 instance, there might be "basic" kinds of groups that correspond to basic
1048 kinds of functions—for instance, to competition—and non-basic kinds of
1049 groups that correspond to ways individuals might compete—for instance, to
1050 playing baseball.

1051 Lastly, we might disagree about which types of collections of individuals
1052 are groups—in particular, whether genders and races are.¹⁸ Functionalist
1053 accounts of genders and races are well-represented. For instance, MacKinnon
1054 (1993)'s remarkably influential account of gender—according to which for
1055 one to be a woman is for one to be sexually subordinated to men and to be
1056 a man to sexually subordinate women—is recognizably functionalist. (See
1057 Jenkins 2017 for a compelling case for this functionalist interpretation of
1058 MacKinnon; see also Young 1990; Witt 2011; and Haslanger 2012 for additional
1059 evidence.) Similarly, Charles Mills (1997)'s account of race—according to
1060 which to be, e.g., black is to be positioned within a social system (i.e., a "vertical
1061 race system") such that one is treated as a "sub-person"—is recognizably
1062 functionalist, too. (See Griffith 2020 for discussion.) But the important point is
1063 this: **FUNCTIONALISM** paves the way for a unification of otherwise disparate
1064 literatures about the metaphysics of groups generally.

18 Another thing we might disagree about is whether genders and races are groups at all. Ritchie (2020) claims that they're "group-like kinds," though one wonders whether they might rather be "kind-like groups." Nonetheless, I won't defend the view that they are. Ultimately, it doesn't matter for the question whether monism or pluralism is correct, though I think it's worthy of our attention.

1065 Although FUNCTIONALISM is ultimately neutral with respect to these dis-
 1066 agreements, it both clarifies what's at issue and provides a framework for
 1067 thinking about how they might be resolved. (My many footnotes attest to it!)
 1068 To the extent that they're meaningful disagreements, then we've reason to
 1069 take it seriously.¹⁹

1072 The Ontological Question: Monism & Pluralism

1071 Though FUNCTIONALISM allows for significant disagreements among func-
 1072 tionalists, there's an important detail that's central to what I've called the
 1073 arguments from Difference and Sameness, namely, whether groups are one
 1074 or many. In this section, I argue that FUNCTIONALISM is compatible both
 1075 with what I've called monistic and pluralistic answers to the ontological ques-
 1076 tion. Ultimately, in section 3, I show that the arguments from Difference and
 1077 Sameness are unsound because of it.

1078 Again, monism is the view that groups are singular things. Some monists
 1079 accept that groups are sets (Sider 2001; Effingham 2010), some that they're
 1080 "realizations of structure" (Ritchie 2013, 2015, 2020), some that they're fusions
 1081 (Hawley 2017), and some that they're "embodiments" of structure (Fine 2020).
 1082 Sets, realizations of structure, fusions, and embodiments of structure are
 1083 "something over and above" the collections of individuals out of which they're
 1084 made up, and, in each case, that something is one.²⁰

1085 And though pluralism is widely rejected, some have come to its defense. In
 1086 particular, some accept that groups are pluralities of individuals that embody
 1087 "plural conditions" (Uzquiano 2018), and others that they're "instantaneous
 1088 pluralities" (Wilhelm 2020). In each case, however, groups are many in the
 1089 sense that they're fundamentally plural "things" (i.e., pluralities), "the very
 1090 kind of ['object'] of which many is to be asserted," as Russell (1903) suggests.

1091 But the important point is this: one can be a functionalist, whether one
 1092 is a monist or a pluralist. In this section, I focus on Kit Fine and Gabriel
 1093 Uzquiano's monistic and pluralistic accounts of groups, respectively. Because
 1094 each appeals to Fine (1999)'s "theory of embodiment," focusing on their tidies
 1095 things up. Though other monists and pluralists will answer the ontological

19 Griffith (2020) takes FUNCTIONALISM, as I understand it, seriously.

20 It might be that a plurality is something over and above *another* plurality, as Uzquiano (2018) accepts with respect to groups and as Dasgupta (2014) suggests more generally with respect to the relation (i.e., grounding) between groups and their members. As a result, this is an unargued for assumption on the part of monists, though it's perfectly commonplace.

1096 question in meaningfully different ways, what I say in restricting myself to
 1097 embodiments is ultimately compatible with them.

1098 First, the framework. According to Fine (2020), groups are embodiments.
 1099 Embodiments are quite like Aristotelean compounds.²¹ Generally, Aristote-
 1100 lean compounds consist not only of “matter”—say, of a plurality of flowers—
 1101 but of “form”—say, of a “being bunched” arrangement—where form is un-
 1102 derstood to structure matter, to turn a plurality of flowers into a bouquet.
 1103 Similarly, embodiments consist not only of their parts but of “principles of
 1104 embodiment” (henceforth: principle/s) that specify the relevant arrangements
 1105 that structure their parts. The bunch is a plurality of flowers that embody the
 1106 arrangement the relevant principle specifies; the bunch is the flowers-qua-
 1107 bunch: a bouquet.²²

1108 Fine distinguishes two kinds of embodiments, what he calls “rigid” and
 1109 “variable” embodiments. On the one hand, rigid embodiments have their parts
 1110 necessarily. For instance, the bouquet is a rigid embodiment because exactly
 1111 the relevant flowers embody exactly the relevant bunching arrangement at
 1112 all times and in all worlds. The bouquet is the bouquet it is because it has
 1113 exactly those flowers arranged in exactly that way; replace one and you’ll have
 1114 another bouquet.

1115 On the other hand, variable embodiments have their parts contingently.
 1116 For instance, my bike is a variable embodiment because it has different “man-
 1117 ifestations” that embody a particular arrangement at particular times and in
 1118 particular worlds. Clearly, *my* bike has different bits of rubber, plastic, and
 1119 metal as parts, and these are arranged ‘bicycley’ at different times and in
 1120 different worlds. In other words, my bike isn’t identical to its manifestations,
 1121 to the objects consisting of bits of rubber, plastic, and metal arranged ‘bicycley’
 1122 at particular times and in particular words. Rather, it’s constituted by them
 1123 then and there. As a result, it persists when I replace a tire. (We might disagree
 1124 about examples, of course, but the distinction is clear enough.)

1125 Ultimately, groups can be either rigid or variable embodiments, though
 1126 Fine (2020) focuses on those that are variable embodiments. The Department

21 See Koons (2014) and Marmodoro (2013) for misgivings.

22 Importantly, embodiment isn’t mere instantiation. Should the flowers merely instantiate the property of being bunched, it follows that the flowers are merely contingently bunched. When one tosses the bunch and the flowers scatter, the bunch—the *flowers*—survive. If the bunch of flowers embody being bunched, however—if they’re the flowers-qua-bunch, the bouquet—they’re necessarily bunched, and, so, the bouquet doesn’t survive the scattering. But this feature of Fine’s account won’t matter for what I go on to say.

1127 is the group it is because it has manifestations at particular times and in
 1128 particular worlds consisting of the individuals that are its members and the
 1129 arrangements they embody at those times and in those worlds. What this
 1130 arrangement is is specified by the relevant principle; something like ‘being
 1131 arranged “departmentally.”’ And in embodying that arrangement, the Depart-
 1132 ment isn’t identical to its manifestations but constituted by them at particular
 1133 times and in particular worlds.²³

1134 But Fine’s theory is remarkably flexible. He places no restrictions on the
 1135 content of principles of embodiment generally. It’s up to say what they are.
 1136 Uzquiano (2018) exploits this flexibility and argues that groups are structured
 1137 by principles that encode *plural* rather than singular conditions, as Fine
 1138 assumes. In particular, he argues that their principles are relevantly like
 1139 “being scattered,” where to be scattered implies that what’s scattered *isn’t* one
 1140 but many.

1141 Moreover, he argues that because his account of groups can accommodate
 1142 what might be groups that others can’t—in particular, supposed groups like
 1143 queues and multitudes that are significantly less structured than, e.g., depart-
 1144 ments of philosophy—we should prefer it to Fine’s. As he puts it, “neither
 1145 queues nor multitudes appear to demand much of their individual members.”
 1146 They must “[embody] a certain spatial arrangement but they do not seem to
 1147 require a shared intentionality or agency from their members” (Uzquiano
 1148 2018, 423). In other words, though individuals that are “members” of queues
 1149 embody minimal arrangements such that they’re queues, this doesn’t entail
 1150 that they’re one.²⁴

1151 However, neither Fine nor Uzquiano say what *kinds* of principles define
 1152 groups rather than other variable embodiments. In other words, neither gives
 1153 us an account of what it is to be a group—what defines GROUP—such that
 1154 we can distinguish embodiments that are groups from those that aren’t. Ulti-
 1155 mately, that’s a desirable feature of the theory of embodiment. It was developed
 1156 as an alternative to traditional accounts of mereological composition. It tells
 1157 us what kinds of *things* groups are, not what it is to be a group *per se*. As a

23 Fine claims that this account can accommodate the view that groups can have different structures at different times and in different worlds. It’s not clear that he’s right, however. Whether he is depends on the content of the relevant principles of embodiment.

24 One wonders whether this is compatible with pluralism. It’s tempting to say that embodiment itself is unifying and, so, that even queues become one in embodying the relevant spatial arrangement, plural conditions be damned.

1158 result, it doesn't presuppose an answer to the definitional question. And that's
1159 good.

1160 But now we can see how both Fine and Uzquiano might answer it. For
1161 we can trace the differences between bikes and departments of philosophy
1162 to the principles they embody. Whereas my bike's parts might embody one
1163 functional arrangement—a principle that specifies what the relevant bits
1164 of rubber and metal do (e.g., enable riding)—a department of philosophy
1165 will embody another functional arrangement—concerning what *it* does (i.e.,
1166 philosophy). As a result, Fineans can appeal to **FUNCTIONALISM** to explain
1167 what it is to be a group per se. In particular, it can explain which embodiments
1168 are groups and which aren't by appealing to the distinctive kinds of principles
1169 the individuals that are their members embody. Because Fine and Uzquiano
1170 can agree that groups are functional, then **FUNCTIONALISM** is compatible
1171 with both monism and pluralism.

1173 **The Arguments from Difference & Sameness**

1173 Now we're ready for the biggest bit: the arguments from Difference and Same-
1174 ness. (Here²⁵ they are for easy access.) Again, they're supposed to be problems
1175 for a pluralistic account of the kinds of things groups are and *not* for a func-
1176 tionalist account of **GROUP**. They're supposed to imperil the pluralist's answer
1177 to the ontological question. In this section, I argue that an attractive version of
1178 functionalist pluralism is invulnerable to them. In particular, I argue that the

25 For reference:

- (1) A group can have different members both at different times and in different worlds.
- (2) If groups are pluralities, a group can't have different members either at different times or in different worlds.
Therefore, groups aren't pluralities.
- (3) According to pluralism, groups are pluralities.
Therefore, pluralism is false.

And:

- (4) Different groups can have the same members both at the same times and in the same worlds.
- (5) If groups are pluralities, different groups can't have the members either at either the same times or in the same worlds.
Therefore, groups aren't pluralities.
- (6) According to pluralism, groups are pluralities.
Therefore, pluralism is false.

1179 arguments from Difference and Sameness obscure a distinction between rela-
 1180 tions I call “being grouped” and “being a member.” Once we distinguish them,
 1181 **FUNCTIONALISM** comes to the rescue, and pluralism is back in business.

3₃₂₁ *The Arguments*

1183 Again, that paradigmatic groups can have different members both at different
 1184 times and in different worlds is widely accepted as a constraint on accounts
 1185 of what they are. I call it:

1186 THE DIFFERENCE CONSTRAINT (DIFFERENCE). A group can have
 1187 different members both at different times and in different worlds.

1188 It’s plausible that departments of philosophy can, e.g., hire professors. In par-
 1189 ticular, it’s plausible that the department once had members it no longer does
 1190 and that it might have had members it never did. Nonetheless, it’s precisely
 1191 the department it either was or might have been. If so, **DIFFERENCE** is true.

1192 Moreover, that different groups can have the same members both at the
 1193 same times and in the same worlds is widely accepted as a constraint on
 1194 accounts of what they are. I call it:

1195 THE SAMENESS CONSTRAINT (SAMENESS). Different groups can
 1196 have the same members both at the same times and in the same
 1197 worlds.

1198 The department and its bowlers, the HaeXeities, might have exactly the same
 1199 members. Nonetheless, it’s plausible that the department isn’t identical to the
 1200 HaeXeities. For instance, whereas members of the department are expected
 1201 to behave decorously in departmental dealings—and, so, might be sanctioned
 1202 for misbehaving—the HaeXeities aren’t; anything goes on the lanes. If so,
 1203 **SAMENESS** is true.

1204 But because it’s widely accepted that pluralities have their “members” es-
 1205 sentially, it’s widely accepted that pluralities can’t have different members
 1206 at either different times or in different worlds. If the plurality of individuals
 1207 that are members of the department here and now—Jonathan, Jennifer, and
 1208 Julia—consists of them essentially, it consists of them necessarily and, so, at
 1209 every time and in every world.

1210 As a result, both **DIFFERENCE** and **SAMENESS** presuppose a principle of
 1211 extensionality for pluralities:

1212 EXTENSIONALITY. One plurality, xx , is identical to another plurality,
 1213 yy , if and only if for all z , z is one of the xx 's if and only if z is one
 1214 of the yy 's.

1215 Assuming **EXTENSIONALITY**, pluralism entails that for groups to be different
 1216 *is* for them to have different members, and for them to be the same *is* for them
 1217 to have the same members. Because Jonathan, Jennifer, and Julia couldn't
 1218 be Jonathan, Jennifer, and Julia, *and Jim*, pluralism entails that each is a
 1219 different plurality and, so, that the department can't have different members
 1220 at different times and in different worlds. Hence the argument from **DIFFER-**
 1221 **ERENCE**. Similarly, if groups are pluralities, the department and the HaeXeities
 1222 aren't different groups because they have exactly the same members: Jonathan,
 1223 Jennifer, and Julia. Hence the argument from **SAMENESS**.

1224 The arguments from **DIFFERENCE** and **SAMENESS** have become *the* argu-
 1225 ments against pluralism in the literature, and, so, they're rarely resisted.²⁶
 1226 Nonetheless, the arguments are misleading. In keeping with **EXTENSION-**
 1227 **ALITY**, it's important to emphasize that (2) and (5) are true if and only if
 1228 pluralism is the view that

- 1229 (i) to be a group, G , is to be identical to a plurality, ab ; and
 1230 (ii) to be a member of G is to be a "member" of ab (i.e., to be either a or b).

1231 Again, **EXTENSIONALITY** entails that the department *is* its members, that they
 1232 *are* the department, because the relevant individuals are "members" of the
 1233 plurality with which it's identical. Similarly, the HaeXeities *is* its members,
 1234 they *are* the HaeXeities, because the relevant individuals are "members" of
 1235 the plurality with which it's identical.

1236 However, there's no one thing it is to be either one or many, and, so, there
 1237 are different versions of pluralism to which the arguments are inattentive.²⁷

26 In particular, though Horden and López de Sa (2021) are pluralists, they reject **SAMENESS**, which neither Uzquiano (2018) nor I are willing to do.

27 The monist will likely insist that there's no conception of plurality on which pluralities don't have their members essentially. And, if they're right, **EXTENSIONALITY** entails that pluralism is false. I both agree and disagree. On the one hand, I agree that the pluralist ought to accept **EXTENSIONALITY**. However, I disagree that accepting **EXTENSIONALITY** entails that there's no conception of pluralities on which they have their members essentially. That's the crux of my argument.

1238 As a result, neither (i) nor (ii) is entailed by pluralism per se. Rather, they
 1239 constitute a—particularly naïve—version of pluralism that both monists
 1240 and pluralists are right to resist. Because both (i) and (ii) are required to
 1241 substantiate the arguments from DIFFERENCE and SAMENESS, then they
 1242 misrepresent pluralism.

1243 The basic idea is this. The arguments from DIFFERENCE and SAMENESS
 1244 assume that EXTENSIONALITY entails that pluralism per se is false. However,
 1245 there's an intuitive version of *functionalist* pluralism that's compatible with
 1246 EXTENSIONALITY. If that's right, (2) and (5) are false, and the arguments from
 1247 DIFFERENCE and SAMENESS are unsound. That's what I argue in this section.

3.4.2 *Being Grouped vs. Being a Member*

1249 I begin with a distinction:

1250 BEING GROUPED. The relation between a group, G, and the plurality
 1251 of its members.

1252 And

1253 BEING A MEMBER. The relation between an individual that is a
 1254 member of a group, G, and G.

1255 Minimally, BEING GROUPED is a multigrade relation. Although it's a relation
 1256 between G and its members, its members' slot doesn't have a definite degree:
 1257 at some times and in some worlds, some number of members stand in this
 1258 relation, and at other times and in other worlds, another number of members
 1259 do. However, BEING A MEMBER is a unigrade relation; its members' slot does
 1260 have a definite degree. In particular, it's a binary relation in which a group
 1261 stands to a particular individual.²⁸

1262 That's sufficient to distinguish BEING GROUPED and BEING A MEMBER.
 1263 They're different relations because they have different properties. But it's
 1264 especially important to distinguish them because they imply the distinction
 1265 between (i) and (ii). For the claim that G is identical to *ab* is a claim about
 1266 the relation between a group and its members. Something makes a particular
 1267 collection of individuals a group rather than a mere collection of individuals.

²⁸ Whether these relations are symmetric, asymmetric, or, importantly, antisymmetric is something I won't—but that should be—discuss/ed in detail.

1268 Jonathan, Jennifer, and Julia are the department, in other words, because
 1269 they're related to it in a particular way; they do this and not that. Again, if (i)
 1270 is true, this relation is identity (i.e., to *ab*).

1271 Moreover, the claim that to be a member of *ab* is to be a "member" of
 1272 *ab* is a claim about the relation between the individuals that are a group's
 1273 members and the group of which they're members. Again, something makes
 1274 a particular individual a member of a particular group. Julia is a member of
 1275 the department because *she's* related to it in a particular way. And, again, if
 1276 (ii) is true, this relation is identity (i.e., to either *a* or *b*).

3.3 Functionalism Pluralism: Roles

1278 Importantly, to distinguish BEING GROUPED from BEING A MEMBER is to
 1279 recognize both that they needn't be identity and that neither (i) entails (ii)
 1280 nor that (ii) entails (i). Here's a version of pluralism that does the work:

1281 ROLES. For *G* to be a group is for

- 1282 (A) *G* to be a plurality of functional roles, *rr*, that are instances²⁹ of a kind,
- 1283 *K*, for
- 1284 (B) *K* to be defined by *rr* at particular times and in particular worlds,
- 1285 (C) within a particular social system.

1286 Simply: the collection of individuals we call the Department of Philosophy
 1287 is a group because the kind of which it's an instance—DEPARTMENT OF
 1288 PHILOSOPHY—is defined by a set of functional roles the individuals that are
 1289 its members realize.³⁰

29 Note that I'm appealing to a distinction between multiple realizability and multiple instantiability here. Again, I've suggested that functional properties like being a pain or being a baseball team are *second-order* properties that are defined in terms of particular first-order properties that are instantiated by particular individuals at different times and in different worlds. But because the Red Sox can have different instances of the same set of first-order realizers of the kind BASEBALL TEAM, it's multiple instantiability and *not* multiple realizability that explains how groups survive changes in membership. Nonetheless, it's plausible that multiple realizability explains how different baseball teams at the same times and in the same worlds each realize the functional property in different ways.

30 One might insist that it's an instance of finer-grained kind; say, being a department. But that's risky. In particular, there's a risk that in appealing to fine-grained social kinds, we can't accommodate the view that groups can have *different* structures at different times and in different worlds. For it's arbitrary to stop appealing to finer- and finer-grained social kinds in distinguishing these groups. If the Department of Philosophy is structured in the particular ways it is because it's a

1290 Importantly, in defining groups in terms of functional roles, **ROLES** is a
 1291 structuralist account of groups. As I suggested in section 1, arrangements
 1292 are essentially relational. There are no roofs without frames, and there are
 1293 no frames without foundations. Analogously, there are no pitchers without
 1294 catchers, no catchers without pitchers. The position **PITCHER** is defined by
 1295 the pitches to relation and thereby **CATCHER**, and the position **CATCHER** is
 1296 defined by the returns to relation and thereby **PITCHER**. And that's what
 1297 **ROLES** implies. Every group is defined by a plurality of functional roles, each
 1298 of which is played by particular individuals that are embedded within wider
 1299 social systems. **ROLES** is holistic, too.

1300 A quick clarification. It's plausible that a *version* of the ontological question
 1301 arises for **ROLES**. Philosophers interested in groups have said remarkably
 1302 little about the metaphysics of roles and, so, about what it is to play one.³¹
 1303 Here, then, is another detail about which functionalists might disagree. I'll
 1304 refer to role-types and role-tokens to simplify matters, but I intend to remain
 1305 ecumenical with respect to their metaphysics. We can reasonably expect any
 1306 account of roles to satisfy the corresponding versions of **DIFFERENCE** and
 1307 **SAMENESS**.

304 *Responding to the Arguments from Difference and Sameness*

1309 **ROLES** reveals that (i) doesn't entail (ii). The first of these claims—that G is
 1310 identical to *ab*—is neutral both with respect to what *a* and *b* are and with
 1311 respect to what it is to be a member of G. In particular, it tells us that *a* and *b*
 1312 aren't *individuals* but the roles they play.³² And, in that case, (ii) doesn't entail
 1313 (i) either.

department of philosophy, it's as structured in the particular ways it is because it's a department
 of philosophy at the University of Chicago, because it's *the* Department of Philosophy at the
 University of Chicago, because it's *the* Department of Philosophy at the University of Chicago *at*
 a particular time and in a particular world, and on and on. And with each successive refinement,
 we'll significantly limit the possibility that it'll survive changes in structure. (See footnote 36.)

31 For instance, a functionalist Finean might accept that *roles* are variable embodiments. That
 is, whereas Fine accepts that groups are variable embodiments, one might accept that groups
 are pluralities of variable embodiments. In other words, whereas Fine claims that groups are
 pluralities of individuals that collectively embody a formal principle, it might be that groups are
 constituted by pluralities of roles that individuals individually embody. And this underwrites an
 account of roles with which I'm ultimately sympathetic. I won't assume it here, however.

32 The words we use to refer to both groups and their members can—or, as Horden and López de
 Sa (2021) argue, do—mislead us. For what we might call “member terms,” such as “department
 chair,” figure in claims with both *de re* and *de dicto* interpretations.

1314 For if G is identical to a plurality of functional role-types, rr , that define the
 1315 kind in question—and *not* to a plurality of individuals, ab , with which they're
 1316 easily confused—to be a member of G is to be a “member” of rr , to be identical
 1317 to either role-type. To identify a plurality’s “members” is to identify that of
 1318 which it’s a plurality; to identify a *group*’s members isn’t to identify the role-
 1319 types of which it’s a plurality. To claim that role-types are members of groups
 1320 would significantly strain our—admittedly pre-theoretical—conception of
 1321 membership. Rather, it’s to identify the individuals that *play* the roles of
 1322 which it’s a plurality, that “are” the role-tokens of those role-types in the way
 1323 that Jonathan “is” the department’s chair. As a result, there are grounds for
 1324 claiming that whereas “membership” is extensional, membership proper is
 1325 *non*-extensional, and, so, that membership isn’t “membership” (contra (ii)).

1326 As a result, the claim that to be a member of G is to be a “member” of ab is
 1327 neutral both with respect to what a and b are and with respect to what it is for
 1328 the individuals that are a group’s members to be grouped. We can suppose that
 1329 a and b are particular individuals rather than the roles they play *and* that to be
 1330 a member of G is to be a “member” of ab . Still, it’s compatible with this view
 1331 that the relation between G and ab isn’t identity. Again, on FUNCTIONALISM,
 1332 it’s realization. In other words, it’s compatible with (ii) that G exists because
 1333 rr does, that G is a plurality of these role-types that’s realized by a plurality of
 1334 individuals that play them at particular times and in particular worlds (contra
 1335 (i)). And this allows that G itself might be realized by different pluralities at
 1336 different times and in different worlds. As a result, the functionalist pluralist

On the one hand, “department chair” has a reading on which it refers to the particular individual that happens to play the relevant role at a particular time and in a particular world (e.g., Jonathan). On the other hand, “department chair” has a reading on which it refers to an arbitrary individual; in particular, to the individual—whoever they are—that plays the relevant role at particular times and in particular worlds.

But we’ve independent reason to interpret claims in which member terms like “department chair” figure *de dicto*. For if “department chair” refers to Jonathan, the claim that Jonathan *is* the department chair is necessarily true; Jonathan is necessarily self-identical. Similarly, if “chief justice” refers to Roberts and “outfielder” to Martinez, the claims that Roberts is the chief justice and Martinez an outfielder are necessarily true. But, of course, that can’t be.

This can be avoided by using member terms like “department chair” (e.g., “chief justice” and “outfielder”) to refer to role-types and, so, to arbitrary individuals. We can say that Jonathan “is” the department’s chair because he plays the role it refers to within a particular social system; in other words, that he “is” a role-token of that role-type. He realizes it at a particular time, in a particular world, and within a particular social system.

1337 needn't reject **EXTENSIONALITY** to reply to the arguments from **DIFFERENCE**
1338 and **SAMENESS**.³³

1339 To summarize: **ROLES** entails that (2) and (5) are false because it distin-
1340 guishes being grouped (via (A)) from being a member of (via (C)), where
1341 **BEING A MEMBER** is non-extensional (via (B)). Pluralism is back in busi-
1342 ness.³⁴

1342 **4 Functionalism Monism vs. Functionalism Pluralism: A**
1344 **Challenge**

1345 That's interesting enough. Whether the argument that follows is successful,
1346 we'll have made some progress: the arguments from **DIFFERENCE** and **SAME-**
1347 **NESS** presuppose an account of pluralism that we shouldn't accept, and, so,
1348 they're unsound. But that does nothing to recommend pluralism. That plu-
1349 ralism is invulnerable to the arguments from **DIFFERENCE** and **SAMENESS** is
1350 one thing. That we should accept it is another.

1351 In this section, I give it a try. I argue that the fact that groups are partly
1352 individuated by the structured social systems in which they're embedded (per
1353 **FUNCTIONALISM**) is an obstacle for functionalist monism. Because monism
1354 recommends the view that groups are exclusively individuated by the relations
1355 their members realize "internally," its conception of the kinds of things groups

33 This marks a crucial difference between Uzquiano (2018)'s pluralism and **ROLES**. Whereas he denies **EXTENSIONALITY**, I don't.

34 Before proceeding, a disclaimer. One might worry that while **ROLES** isn't vulnerable to the arguments from **DIFFERENCE** and **SAMENESS**, it's vulnerable to arguments from nearby metaphysical constraints. In particular, we can distinguish **DIFFERENCE** and **SAMENESS**—which are claims about the individuals that are a group's members—from **Difference*** and **Sameness***—which are claims about a group's structure; in this case, about the relevant role-types:

DIFFERENCE*. Groups can have different structures at different times and in different worlds.

SAMENESS*. Different groups can have the same structures at the same times and in the same worlds.

Provided we accept **EXTENSIONALITY**, in other words, the arguments from **DIFFERENCE*** and **SAMENESS*** immediately arise. Ultimately, I'm inclined to say that functional kinds are hierarchical and, so, that groups can survive important structural changes by realizing "lower-level" kinds at different times and in different worlds. But because consideration of this issue exceeds the scope of this paper, I set it aside. See Pagano (ms), Faller (2021), and Wilhelm (2022) for accounts of persistence that claim to be pluralistic.

are is a liability. To the extent that we favor **FUNCTIONALISM**, then we've reason to disfavor functionalist monism and to favor functionalist pluralism.

A note. I'm not claiming that the challenge I raise for functionalist monism is dispositive. Functionalism monism is a powerful view of the metaphysics of groups, and it has powerful resources. Rather, I'm claiming that it's a *meaningful* challenge for functionalist monism, both independently and because of its implications for functionalist pluralism. As a result, I offer it to both functionalist monists and functionalist pluralists. It represents a significant point of disagreement among them that's worth exploring.

4.6.1 *Internal and External Structure*

Let's return to the view that groups are embedded in structured social systems. In a series of influential papers, Ritchie (2013, 2015, 2020) defends an account of groups according to which groups are—*deep breath*—elements-realizing-social-structure. Groups are singular things, and their elements (e.g., their members) are arranged in particular ways.³⁵ As a result, her account substantiates the arranging/arrangement distinction I introduced in section 1: a group's elements are arranged in *realizing* an arrangement, and arrangements are the social structures they realize. The Red Sox's members realize a particular social structure consisting in part of PITCHER and CATCHER.

Ritchie's is an exceptionally insightful account of groups, innovative and rich with nuance. But one of its innovations is especially important in understanding both groups and the structured social systems in which they're embedded. She distinguishes what she calls "internal" and "external" ways the individuals that are a group's members realize social structure. In particular, she claims that a group is internally structured when precisely its members are arranged in particular ways and externally structured when it and/or its members and other groups and/or *their* members are arranged in particular ways. For instance, Bertrand Russell University is internally structured relative to both the College of Liberal Arts and the Department

35 It's an open question whether things like books, buildings, and buses occupy nodes, too. Certainly, they aren't *members* of groups but, perhaps, parts of them. For instance, Fine (2020) claims that things like buildings, buses, and basins are *parts* of groups. In particular, he proposes that they're "spatial" rather than "temporal" parts of groups rather than that they're members. Of course, one needn't accept that buildings, buses, and basins are parts of groups either. For instance, one might accept that buildings, buses, and basins *enable* the relevant connections among groups and/or their members.

Nonetheless, I'll focus on their members as Ritchie does.

1385 of Philosophy, the College of Liberal Arts is externally structured relative
 1386 to the Bertrand Russell University and internally structured relative to the
 1387 Department of Philosophy, and the Department of Philosophy is internally
 1388 structured relative to its members and externally structured relative to both
 1389 the College of Liberal Arts and Bertrand Russell University.

1390 That's quite plausible. But there are several things to emphasize about
 1391 Ritchie's account. First, she claims that groups are internally and externally
 1392 structured *relative* to both their members and other groups and/or their mem-
 1393 bers. The Department of Philosophy is externally structured relative to both
 1394 the College of Liberal Arts and Bertrand Russell University, yes. But it's ex-
 1395 ternally related relative to the HeXaeities, with which it's co-extensive, too.
 1396 There are uncountably many social structures particular groups realize and,
 1397 so, uncountably many ways they're structured relative to one another.

1398 Second, and again, Ritchie develops the internal/external distinction by
 1399 appealing to Shapiro (1997)'s account of mathematical structure. Shapiro
 1400 distinguishes "systems" and "structures" and claims that a system is a "collec-
 1401 tion of [particular elements] with certain relations" (Shapiro 1997, 73), and a
 1402 structure is "the abstract form of a system, highlighting the interrelationships
 1403 among the [particular elements]" (Shapiro 1997, 74). Analogously, Ritchie ac-
 1404 cepts that whereas groups are systems whose elements are arranged in this or
 1405 that way, a group's structure is the arrangement their elements realize. Groups
 1406 are internally and externally structured in whatever ways the arrangements
 1407 they realize specify.

1408 Shapiro represents structures hypergraphically in terms of "nodes"—or
 1409 positions things occupy—and edges—or relations that link them. Ritchie
 1410 does, too, claiming that

1411 the structure of a group can be represented with nodes [...] and
 1412 edges connecting nodes to other nodes. The edges of a structure
 1413 capture the relations that hold between nodes. Since all members
 1414 of a group are related to some degree, each node in structure S is
 1415 connected to every other node in S. (Ritchie 2013, 268)

1416 As a result, she precisifies her internal/external distinction hypergraphically:
 1417 a group is internally structured when "all the relevant nodes are occupied by
 1418 its members and every member occupies some node or other" (Ritchie 2020,
 1419 409), and externally structured if and only if its elements "occupy only some
 1420 node/s" of the relevant internal structures and when "other nodes [...] are
 1421 occupied by entities or systems that are not" among them (Ritchie 2020, 410).

1422 Third, the internal/external distinction doesn't tell us how particular groups
 1423 are individuated. But we can and should ask how they are. In particular, we
 1424 can and should ask why—that is, in virtue of what³⁶—the College of Liberal
 1425 Arts is externally structured relative to Bertrand Russell University rather than
 1426 to the ALE and internally structured relative to the Department of Philosophy
 1427 rather than to the Red Sox in the ways it is. Again, the internal/external
 1428 distinction doesn't say which of these social structures is privileged relative
 1429 to the College of Liberal Arts. But, of course, exactly one is: the department
 1430 of philosophy structure. As a result, we'll want more.

4.2 A Challenge

1432 Luckily, **FUNCTIONALISM** gives us more. Suppose the individuals that are the
 1433 Red Sox's members are precisely the individuals that are the Department of
 1434 Philosophy's members. But, of course, they're different groups, and **FUNC-**
 1435 **TIONALISM** accounts for the fact that they're different groups by appealing
 1436 to the functions they serve. It says that whereas the Red Sox function in one
 1437 way—the baseball team way—the Department of Philosophy functions in
 1438 another way—the department of philosophy way—whatever these amount
 1439 to.

1440 As a result, **FUNCTIONALISM** implies that we can't account for the fact
 1441 that the Red Sox and the Department of Philosophy are different groups by
 1442 appealing to the ways the relevant individuals are arranged. We've supposed
 1443 that the individuals that are members of both groups are arranged in both
 1444 ways. Nonetheless, the individuals that are the Red Sox's members aren't
 1445 arranged in being, e.g., teachers, nor are the individuals that are members of
 1446 the Department of Philosophy's arranged in being pitchers. In other words,
 1447 although the individuals that are the Red Sox's members are teachers, they
 1448 don't do philosophy *as* members of a baseball team. Likewise, although the
 1449 individuals that are the Department of Philosophy's members are pitchers,
 1450 they don't play baseball *as* members of a department of philosophy. However,
 1451 we *can* account for the fact that the Red Sox and the Department of Philosophy
 1452 are different groups by appealing to the different functions they serve.

1453 Here it's especially important to emphasize that, according to **FUNCTION-**
 1454 **ALISM**, the Department of Philosophy is the kind of group it is because the
 1455 collection of individuals that are its members serve a particular function

36 There's a causal "why question" I mean to avoid asking in this context.

1456 *within a particular social system.* The individuals that are members of the Red
1457 Sox don't play baseball simpliciter. Rather, they play baseball *in the ALE.* If
1458 we can't appeal to the ALE in individuating the Red Sox, we can't distinguish
1459 it from either departments of philosophy with the same members or—more
1460 importantly—from other baseball teams (e.g., The Dodgers). Similarly, the
1461 individuals that are members of the Department of Philosophy don't do phi-
1462 losophy simpliciter. Again, if we can't appeal to the College of Liberal Arts,
1463 we can't distinguish it from either baseball teams with the same members
1464 or—again, more importantly—from other departments of philosophy (e.g., at
1465 David Lewis University). As a result, **FUNCTIONALISM** implies that a group's
1466 external structure plays a distinctive role in individuating them. In particular,
1467 it implies that the Red Sox and the Department of Philosophy are different
1468 groups because each is embedded in different social systems that thereby
1469 structure them.

1470 Now for the important point. The functionalist's emphasis on external
1471 structure seems to be in tension with the view that the individuals that are
1472 a group's members are singular entities. For if groups are “unified wholes,”
1473 they have identifiable boundaries that mark them off from one another. In
1474 particular, they're marked off by their intrinsic rather than their extrinsic
1475 properties. Indeed, that's what intrinsic properties are: properties things have
1476 that don't “mention” other things. But because **FUNCTIONALISM** entails that
1477 group kinds are extrinsic to the collections of individuals that realize them, it
1478 seems in conflict with the view that they're singular entities that exclude the
1479 groups that are external to them.

1480 Although Ritchie isn't my primary target, let's consider an explicitly func-
1481 tionalist version of her account and see whether it has the resources to respond.
1482 Surely, her conception of internal/external structure captures the fact that
1483 the Department of Philosophy is externally structured relative to the College
1484 of Liberal Arts in being bound by its charter. Per Shapirian structuralism,
1485 the relevant complex of relations is there, and we abstract it. But as singular
1486 entities with identifiable boundaries, it's not clear that she's entitled to the
1487 view that particular groups are partly individuated by their external structures.
1488 Again, as she suggests, whereas a group is internally structured when “all
1489 the relevant nodes are occupied by its members and every member occupies
1490 some node or other” (Ritchie 2020, 409), it's externally structured if and only
1491 if its elements “occupy *only some node/s*” of the relevant internal structures
1492 and when “other nodes [...] are occupied by entities or systems *that are not*”
1493 *among them* (Ritchie 2020, 410, emphasis added). And it's precisely this that

1494 makes a group's external structure "stand outside" the singular entities—the
 1495 groups—that are thereby externally related to it. In other words, the view
 1496 that groups are singular entities seems to imply that they're individuated
 1497 exclusively by their internal structures. But that's the problem.

1498 There's a good question about how singular entities are individuated, of
 1499 course.³⁷ There are certainly accounts of singular entities that don't have this
 1500 result. (Fine's is one of them, and I'll consider it shortly.) But because Ritchie's
 1501 account of groups has it that we abstract structures from whatever social systems
 1502 are there already, it's difficult to see how she might individuate precisely
 1503 the "right" social systems (i.e., the Department of Philosophy) rather than
 1504 others (e.g., the Department of Philosophy + the Red Sox) without privileging
 1505 their internal structures. In other words, it's difficult to see how she isn't
 1506 committed to the view that a particular group is individuated exclusively by
 1507 its internal structure when what's there to be extracted is a tangle of relations,
 1508 both internal and external, only some of which unify the group in question.

1509 Luckily, Ritchie has options, and each is worthy of significant consideration.
 1510 Again, I don't claim that this challenge to functionalist monism is dispositive,
 1511 only that it's worth considering. I'll consider one.

1512 She might accept that the Department of Philosophy and the College of
 1513 Liberal Arts *are* asymmetrically, *internally* related. It's certainly true that
 1514 they stand in a kind of asymmetric relation. There's a function they serve
 1515 that entails it, and that's realized when the Department of Philosophy is
 1516 bound by its charter. But that seems to require that the College of Liberal
 1517 Arts and Department of Philosophy *aren't* different groups. In particular, it
 1518 seems to require that the realization of structure to which the Department of
 1519 Philosophy corresponds *is* the realization of structure to which the College of
 1520 Liberal Arts since the function they serve unifies them.

1521 And this problem compounds the further up the hierarchy we go. For to
 1522 retain the view that the relevant functions are served by whatever singular
 1523 entities they unify, we seem compelled to search out ever larger social systems
 1524 to accommodate the view that the groups to which they correspond have
 1525 identifiable boundaries. Again, the Department of Philosophy will become
 1526 a member of the College of Liberal Arts such that the members of each are

37 This corresponds to a familiar issue, namely, the possibility of Max Black objects. Something that's worth considering is whether there are Max Black *groups*. If so, the monist might have resources to resist this challenge. But the issue is precisely how to square FUNCTIONALISM—that individuates groups by appeal to extrinsic properties—with monism—that individuates groups exclusively by appeal to their intrinsic properties.

1527 in fact internally related. Moreover, the College of Liberal Arts will become
 1528 a member of Bertrand Russell University such that all of their members are
 1529 internally related, too. But, again, that robs us of the view that these are
 1530 different groups. Whether this commits Ritchie to the existence of a single
 1531 group—society itself, say—is beside the point.³⁸ The point is that in order to
 1532 accommodate the view that groups are individuated by whatever functions
 1533 unify their members, she'll commit herself to an implausible view of their
 1534 interrelations.³⁹

1535 Functionalist Fineanism recommends a different response, one that might
 1536 be available to Ritchie, too. (However, her failure to account for the role group
 1537 kinds play in individuating groups remains a problem.) But although Fine
 1538 doesn't have the problem I've raised for Ritchie's account, he has a relevantly
 1539 similar problem.

1540 Importantly, Fine can reasonably deny that groups are individuated solely
 1541 by the relations their members realize internally; in particular, because there
 1542 are no restrictions on the content of the principles of embodiment they mani-
 1543 fest. Again, he might accept that though we individuate the department by
 1544 its relation to the university, the university isn't part of that *thing*, the de-
 1545 partment. In particular, he might insist that because the relevant principles
 1546 of embodiment are functional, particular extrinsic—or, in Ritchie's sense,

38 Interestingly, this suggests a kind of priority monism with respect to social goings-on. According to Schaffer (2010), a single thing—the universe—grounds everything else there is. The universe is fundamental. Similarly, the functionalist might accept that a single thing—society—is fundamental relative to the social and, so, with respect to groups. And in the way Schaffer appeals to quantum entanglement to justify his priority monism, the “social priority monist” might appeal to “social entanglement” to justify theirs. And that's an issue worth exploring, monism and pluralism aside.

39 She has another, better option. (I think there's problems with it, too, but it's important to emphasize that it isn't a problem for Ritchie's monism but for her commitment to Shapiro's view of structure.) As a result, I'll mention it, then set it aside. She might accept a plausible distinction between parthood and membership (*pace Ruben 1985*), and she might insist that both a group's members and the groups of which they're *part* are individuated. For the wider social systems of which particular groups are embedded are plausibly groups of which they're *part* rather than of which they're members. That's a significantly better option and one that's worth pursuing. However, it, too, risks a kind of arbitrariness. Again, it doesn't say which social systems are groups and which are the groups of which they're *part*. Certainly, it implies that the Department of Philosophy and the College of Liberal Arts are different groups; each serves a different function. But it also implies that they *aren't* different groups. As before, there is a function they *both* serve; there are many. It implies both that they are and aren't different groups. If being a member of a group concerns realizing the relevant functional kind, then even the membership/parthood distinction is inadequate to capture the view that they're different groups.

1547 external—relations are needed to pick out the particular collections of indi-
 1548 viduals they unify. He's entitled to use Ritchie's distinction in that way.

1549 For instance, he might say that to the extent that the Red Sox are a baseball
 1550 team, the individuals that are its members are unified by the relations that
 1551 define BASEBALL TEAM. Nonetheless, they're individuated by their relations
 1552 to, e.g., the ALE—and, so, to the MLB—and to the City of Boston, too, because
 1553 the relations between them are what make the Red Sox the unique instance
 1554 of the kind it is; the very group it is. In other words, whereas the internal
 1555 relations that make them a baseball team unify them, the external relations
 1556 that make them *the* baseball team in question—the Red Sox—don't.

1557 Nonetheless, this response makes Finean principles of embodiment intol-
 1558 erably arbitrary.⁴⁰ In particular, it suggests that if a group is individuated both
 1559 by the structures it realizes internally and externally, there isn't a principled
 1560 distinction between a particular group and the groups to which it's externally
 1561 related. This is a version of the challenge raised for Ritchie. For if, in order to
 1562 individuate the relevant collections of individuals, principles of embodiment
 1563 appeal to relations that aren't definitional of the kinds of groups in question, it
 1564 will be difficult to say which groups are which and why. That the department is
 1565 unified by a principle that appeals to relations that don't unify the individuals
 1566 that are its members—in this case, to the university—is at best stipulative. In
 1567 other words, if variable embodiment is unifying, it's not clear why that which
 1568 is externally related to that which is internal to a particular group doesn't
 1569 have as much a right to be counted as part of the same group. As a result, it's
 1570 not clear that he's justified in claiming that genuine unification occurs.

1571 Relatedly, this response makes it impossible to tell whether to prefer Fine's
 1572 monism to Uzquiano's pluralism. Again, Uzquiano claims that the relevant
 1573 principles of embodiment don't stamp out singular but plural "entities." But
 1574 since each assumes that principles of embodiment are either singular or
 1575 plural, it's difficult to know how to decide between them. For both Fine and
 1576 Uzquiano accept that for variable embodiments to be identical is for "them" to
 1577 embody the same principle of embodiment. But if principles of embodiment
 1578 are individuated by their modal profiles—as Fine (1999, 70) and Uzquiano

40 There are other responses available to Fine, too. For instance, he might accept that both internal and external relations *are* definitional of groups such that it's necessary that particular groups are embedded in the social systems in which they in fact are. But this gives rise to a host of other problems, in particular concerning the possibility that groups can have different *structures* at different times and in different worlds. Ultimately, solving that problem requires an account of group structure itself. Again, see Pagano (ms) for an attempt.

1579 (2018, 442)’s remarks suggest—it’s not clear why we should think that a given
 1580 group embodies a plural *rather* than a singular condition.⁴¹ (This is as much
 1581 a problem for Uzquiano as it is for Fine, of course, but I’ll set that aside.)

4.3 *Functionalist Pluralism: Redux*

1583 However, this isn’t a problem for the functionalist pluralist; in particular, for
 1584 **ROLES**. For given the distinction between internal and external structure,
 1585 we can accept that groups are structured by the internal relations among
 1586 the roles that ground their existence. We can accept that some of the roles
 1587 departments of philosophy realize depend on others. For instance, we can
 1588 accept that the role of being an associate professor is tied to the role of being
 1589 an assistant professor in the way the role of being a pain is tied to that of
 1590 being a wince. And though the proponent of **ROLES** accepts that the roles
 1591 in question are interrelated, they get to deny that groups are *unified* by the
 1592 relations among them. In particular, they have principled reasons to deny that
 1593 groups are individuated solely by these relations and to accept that they’re
 1594 at least partly individuated by the social systems in which they’re embedded.
 1595 Their pluralism is precisely what vindicates their **FUNCTIONALISM**.

1596 But, again, there are details to sort out. And, again, how a functionalist
 1597 pluralist ought to conceive of structured social systems is important. Here’s
 1598 what I’m inclined to say. As I suggested in section 1, **FUNCTIONALISM** ranges
 1599 over social systems. Given what I’ve said here, then, we might think of social
 1600 systems as consisting of clusters of role-tokens, each of which corresponds
 1601 to a group.⁴² How tightly pluralities of roles cluster will correspond to the
 1602 specificity of the functions they realize. For instance, the Red Sox play roles

41 A related problem arises because of Fine’s assumption that groups are *either* rigid or variable
 embodiments. For it’s important to emphasize that the difference between what I’ve called
 collective and individual embodiment isn’t inconsequential; in particular, because there are
 considerations that count in favor of individual embodiment. For there seem to be groups that
 are neither rigid nor variable embodiments. For instance, the Jimi Hendrix Experience (JHE)
 seems to be a group that has some of its members rigidly—namely, Jimi Hendrix—and some
 of its members variably—namely, everyone else. As a result, though there’s no time or world
 at which Jimi Hendrix isn’t a member of JHE, there are both times and worlds at which other
 individuals are its other members. But that’s incompatible with JHE being either a rigid or variable
 embodiment. However, it’s perfectly reasonable to accept that some groups are constituted by
both rigid and variable roles. JHE might be constituted both by a role that only Jimi Hendrix can
 play and by roles that many individuals can play.

42 I say *either/or* because, in addition to disagreeing about what structure is, we can meaningfully
 disagree about what social systems are.

1603 that are clearly defined by the function they serve—again, to play baseball in
1604 a particular way within a particular set of institutions. However, genders—for
1605 instance, women—play roles that aren’t as clearly defined and that interact
1606 with different roles—for instance, with race and class roles—in complicated
1607 ways.

1608 But, again, because clusters of roles aren’t singular entities, we can indi-
1609 viduate them both by the structures they realize internally—that is, by the
1610 relations among the roles in question and because of which they can be said
1611 to cluster—and/or the structures they realize externally—that is, by their
1612 relation to other clusters. For instance, we can individuate the Department of
1613 Philosophy by identifying the roles the relevant individuals realize. And we
1614 can identify these by identifying the function they realize within the relevant
1615 set of institutions—and, ultimately, the maximal social system—in question.
1616 The Department of Philosophy is the department it is because it does philoso-
1617 phy in a particular way within a broader social system within which the other
1618 groups to which it’s related are embedded, too.

1615 5 Conclusion

1620 In this paper, I’ve argued that the arguments from **DIFFERENCE** and **SAME-**
1621 **NESS** are unsound. They obscure both the distinction between the defini-
1622 tional and ontological questions and between **BEING GROUPED** and **BEING A**
1623 **MEMBER**. I’ve articulated a version of functionalist pluralism—what I called
1624 **ROLES**—that bears this out.

1625 Moreover, I’ve argued that once we make these distinctions, we see that the
1626 crucial question is whether the grouping relation is monistic or pluralistic, in
1627 particular, whether or not grouping is unifying. I’ve argued that if groups are
1628 one, the grouping relation *is* unifying and that this raises an important diffi-
1629 culty for the functionalist monist, namely, the problem of how to individuate
1630 groups. I’ve argued that if groups are many, this problem doesn’t arise.

1631 Although the implications for pluralism are clear, one of my aims is to
1632 generate interest in **FUNCTIONALISM** about groups, whether monistic or
1633 pluralistic. As I’ve suggested, there are important details about which we
1634 might reasonably disagree. Nonetheless, I hope to have shown we have reason
1635 to attend to them and, so, to treat **FUNCTIONALISM** as a viable metaphysical
1636 framework for theorizing about groups.*

* THANKS

Emilie Pagano
 University of Texas at Austin
 emilie.pagano@gmail.com

References

- 1637
 1638
 1639
- 1640
- 1641 BAKER, Lynne Rudder. 2007. *The Metaphysics of Everyday Life: An Essay in Practical*
 1642 *Realism*. Cambridge Studies in Philosophy. Cambridge: Cambridge University
 1643 Press, doi:10.1017/cbo9780511487545.
- 1644 BAXTER, Donald L. M. 1988. "Identity in the Loose and Popular Sense." *Mind* 97(388):
 1645 575–582, doi:10.1093/mind/xcvii.388.575.
- 1646 BEALER, George. 1997. "Self-Consciousness." *The Philosophical Review* 106(1): 69–117,
 1647 doi:10.2307/2998342.
- 1648 BICKLE, John. 2020. "Multiple Realizability." in *The Stanford Encyclopedia of Philoso-*
 1649 *phy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of
 1650 Language and Information, [https://plato.stanford.edu/archives/sum2020/entries/](https://plato.stanford.edu/archives/sum2020/entries/multiple-realizability/)
 1651 [multiple-realizability/](https://plato.stanford.edu/archives/sum2020/entries/multiple-realizability/).
- 1652 BLOCK, Ned. 1978. "Troubles with Functionalism." in *Minnesota Studies in the Phi-*
 1653 *losophy of Science, Volume IX: Perception and Cognition: Issues in the Foundations*
 1654 *of Psychology*, edited by C. Wade SAVAGE, pp. 261–325. Minneapolis, Minnesota:
 1655 University of Minnesota Press. Reprinted in Block (2007, 63–101), doi:10.7551/mi
 1656 [tpress/2111.003.0006](https://doi.org/10.7551/mitpress/2111.003.0006).
- 1657 —. 2007. *Consciousness, Function, and Representation. Collected Papers, Volume 1*.
 1658 Cambridge, Massachusetts: The MIT Press, doi:10.7551/mitpress/2111.001.0001.
- 1659 CORREIA, Fabrice. 2017. "Real Definitions." in *Philosophical Issues 27: Metaphysics*,
 1660 edited by Jonathan SCHAFFER, pp. 52–73. Malden, Massachusetts: Wiley-Blackwell,
 1661 doi:10.1111/phis.12091.
- 1662 DASGUPTA, Shamik. 2014. "On the Plurality of Grounds." *Philosophers' Imprint* 14(20),
 1663 <http://hdl.handle.net/2027/spo.3521354.0014.020>.
- 1664 DORR, Cian. 2016. "To Be F Is to Be G." in *Philosophical Perspectives 30: Metaphysics*,
 1665 edited by John HAWTHORNE and Jason TURNER, pp. 39–134. Hoboken, New Jersey:
 1666 John Wiley and Sons, Inc., doi:10.1111/phpe.12079.
- 1667 EFFINGHAM, Nikk. 2010. "The Metaphysics of Groups." *Philosophical Studies* 149(2):
 1668 251–267, doi:10.1007/s11098-009-9335-4.
- 1669 EPSTEIN, Brian. 2015. *The Ant Trap: Rebuilding the Foundations of the Social Sciences*.
 1670 Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199381104.001.0001.
- 1671 EVNINE, Simon J. 2016. *Making Objects and Events: A Hylomorphic Theory of Artifacts,*
 1672 *Actions, and Organisms*. Oxford: Oxford University Press, doi:10.1093/acprof:
 1673 [oso/9780198779674.001.0001](https://doi.org/10.1093/acprof:oso/9780198779674.001.0001).

- 1674 FALLER, August. 2021. "How Groups Persist." *Synthese* 198(8): 7149–7163, doi:[10.1007/s11229-019-02514-0](https://doi.org/10.1007/s11229-019-02514-0).
- 1675
- 1676 FINE, Kit. 1999. "Things and their Parts." in *Midwest Studies in Philosophy 23: New*
1677 *Directions in Philosophy*, edited by Peter A. FRENCH and Howard K. WETTSTEIN, pp.
1678 61–74. Boston, Massachusetts: Blackwell Publishers, doi:[10.1111/1475-4975.00004](https://doi.org/10.1111/1475-4975.00004).
- 1679 —. 2020. "The Identity of Social Groups." *Metaphysics* 3(1): 81–91, doi:[10.5334/met.45](https://doi.org/10.5334/met.45).
- 1680 FODOR, Jerry A. 1968. *Psychological Explanation: An Introduction to the Philosophy of*
1681 *Psychology*. New York: Random House.
- 1682 GRIFFITH, Aaron M. 2020. "Realizing Race." *Philosophical Studies* 177(7): 1919–1934,
1683 doi:[10.1007/s11098-019-01291-3](https://doi.org/10.1007/s11098-019-01291-3).
- 1684 GUALA, Francesco. 2016. *Understanding Institutions: The Science and Philosophy of*
1685 *Living Together*. Princeton, New Jersey: Princeton University Press, doi:[10.1515/9781400880911](https://doi.org/10.1515/9781400880911).
- 1686
- 1687 HASLANGER, Sally. 1995. "Ontology and Social Construction." *Philosophical Topics*
1688 23(2): 95–125, doi:[10.5840/philtopics19952324](https://doi.org/10.5840/philtopics19952324).
- 1689 —. 2000. "Gender and Race: (What) Are They? (What) Do We Want Them to Be?"
1690 *Noûs* 34(1): 31–55, doi:[10.1111/0029-4624.00201](https://doi.org/10.1111/0029-4624.00201).
- 1691 —. 2012. *Resisting Reality: Social Construction and Social Critique*. New York: Oxford
1692 University Press, doi:[10.1093/acprof:oso/9780199892631.001.0001](https://doi.org/10.1093/acprof:oso/9780199892631.001.0001).
- 1693 —. 2016. "What Is a (Social) Structural Explanation?" *Philosophical Studies* 173(1):
1694 113–130, doi:[10.1007/s11098-014-0434-5](https://doi.org/10.1007/s11098-014-0434-5).
- 1695 HAWLEY, Katherine. 2017. "Social Mereology." *Journal of the American Philosophical*
1696 *Association* 3(4): 395–411, doi:[10.1017/apa.2017.33](https://doi.org/10.1017/apa.2017.33).
- 1697 HORDEN, John and LÓPEZ DE SA, Dan. 2021. "Groups as Pluralities." *Synthese* 198(11):
1698 10237–10271, doi:[10.1007/s11229-020-02715-y](https://doi.org/10.1007/s11229-020-02715-y).
- 1699 JENKINS, Katharine. 2017. "What Women Are For: Pornography and Social Ontology."
1700 in *Beyond Speech: Pornography and Analytic Feminist Philosophy*, edited by Mari
1701 MIKKOLA, pp. 91–112. Oxford: Oxford University Press, doi:[10.1093/acprof:oso/9780190257910.003.0005](https://doi.org/10.1093/acprof:oso/9780190257910.003.0005).
- 1702
- 1703 KOONS, Robert C. 2014. "Staunch vs. Faint-Hearted Hylomorphism: Toward an Aris-
1704 totelian Account of Composition." *Res Philosophica* 91(2): 151–177, doi:[10.11612/r](https://doi.org/10.11612/resphil.2014.91.2.1)
1705 [esphil.2014.91.2.1](https://doi.org/10.11612/resphil.2014.91.2.1).
- 1706 KOSLICKI, Kathrin. 2018. *Form, Matter, Substance*. Oxford: Oxford University Press,
1707 doi:[10.1093/oso/9780198823803.001.0001](https://doi.org/10.1093/oso/9780198823803.001.0001).
- 1708 LEMOS, Noah M. 2005. "The Bearers of Intrinsic Value." in *Recent Work on Intrinsic*
1709 *Value*, edited by Toni RÖNNOW-RASMUSSEN and Michael J. ZIMMERMAN, pp.
1710 181–190. Library of Ethics and Applied Philosophy n. 17. Berlin: Springer, doi:[10](https://doi.org/10.1007/1-4020-3846-1_14)
1711 [.1007/1-4020-3846-1_14](https://doi.org/10.1007/1-4020-3846-1_14).
- 1712 LEWIS, David. 1980. "Mad Pain and Martian Pain." in *Readings in the Philosophy of*
1713 *Psychology. Volume I*, edited by Ned BLOCK, pp. 216–222. Language and Thought.

- 1714 Cambridge, Massachusetts: Harvard University Press. Reprinted, with a postscript
 1715 (Lewis 1983b), in Lewis (1983a, 122–130), doi:10.4159/harvard.9780674594623.c24.
 1716 —. 1983a. *Philosophical Papers, Volume 1*. Oxford: Oxford University Press, doi:10.109
 1717 3/0195032047.001.0001.
- 1718 —. 1983b. “Postscript to Lewis (1980).” in *Philosophical Papers, Volume 1*, pp. 130–132.
 1719 Oxford: Oxford University Press, doi:10.1093/0195032047.003.0009.
- 1720 MACKINNON, Catharine A. 1993. *Only Words*. Cambridge, Massachusetts: Harvard
 1721 University Press, doi:10.2307/j.ctvj2xs7.
- 1722 MARMODORO, Anna. 2013. “Aristotle’s Hylomorphism Without Reconditioning.” *Philo-*
 1723 *sophical Inquiry* 37(1/2): 5–22, doi:10.5840/philinquiry2013371/28.
- 1724 MCCULLAGH, Mark. 2000. “Functionalism and Self-Consciousness.” *Mind and Lan-*
 1725 *guage* 15(5): 481–499, doi:10.1111/1468-0017.00146.
- 1726 MILLS, Charles W. 1997. *The Racial Contract*. Ithaca, New York: Cornell University
 1727 Press, doi:10.7591/9780801471353.
- 1728 PUTNAM, Hilary. 1975. *Mathematics, Matter and Method. Philosophical Papers, Volume*
 1729 *1*. Cambridge: Cambridge University Press. Second edition: Putnam (1979).
 1730 —. 1979. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. 2nd ed.
 1731 Cambridge: Cambridge University Press. First edition: Putnam (1975), doi:10.101
 1732 7/cbo9780511625268.
- 1733 RITCHIE, Katherine. 2013. “What Are Groups?” *Philosophical Studies* 166(2): 257–272,
 1734 doi:10.1007/s11098-012-0030-5.
- 1735 —. 2015. “The Metaphysics of Social Groups.” *Philosophy Compass* 10(5): 310–321,
 1736 doi:10.1111/phc3.12213.
- 1737 —. 2020. “Social Structures and the Ontology of Social Groups.” *Philosophy and Phen-*
 1738 *omenological Research* 100(2): 402–424, doi:10.1111/phpr.12555.
- 1739 ROSEN, Gideon, BYRNE, Alex, COHEN, Joshua, HARMAN, Elizabeth and SHIFFRIN,
 1740 Seana Valentine, eds. 2018. *The Norton Introduction to Philosophy*. 2nd ed. New
 1741 York: W.W. Norton & Co.
- 1742 RUBEN, David-Hillel. 1985. *The Metaphysics of the Social World*. London: Routledge &
 1743 Kegan Paul.
- 1744 RUSSELL, Bertrand Arthur William. 1903. *The Principles of Mathematics*. London:
 1745 Taylor & Francis. Second edition: Russell (1937), third edition: Russell (2020).
 1746 —. 1937. *The Principles of Mathematics*. 2nd ed. London: George Allen & Unwin.
 1747 Second edition of Russell (1903), with a new introduction; third edition: Russell
 1748 (2020).
 1749 —. 2020. *The Principles of Mathematics*. 3rd ed. London: Routledge. Third edition of
 1750 Russell (1903), doi:10.4324/9780203822586.
- 1751 SCHAFFER, Jonathan. 2010. “Monism: The Priority of the Whole.” *The Philosophical*
 1752 *Review* 119(1): 31–76, doi:10.1215/00318108-2009-025.
- 1753 SEARLE, John R. 1995. *The Construction of Social Reality*. London: The Free Press.

- 1754 —. 2010. *Making the Social World: The Structure of Human Civilization*. Oxford: Oxford
1755 University Press, doi:10.1093/acprof:osobl/9780195396171.001.0001.
- 1756 SHAPIRO, Stewart. 1997. *Philosophy of Mathematics: Structure and Ontology*. Oxford:
1757 Oxford University Press, doi:10.1093/0195139305.001.0001.
- 1758 SHOEMAKER, Sydney S. 1984. *Identity, Cause and Mind: Philosophical Essays*. 1st ed.
1759 Oxford: Oxford University Press. Second, expanded edition: Shoemaker (2003).
- 1760 —. 2003. *Identity, Cause and Mind: Philosophical Essays*. 2nd ed. Oxford: Oxford
1761 University Press. Second, expanded edition of Shoemaker (1984), doi:10.1093/os
1762 o/9780199264698.001.0001.
- 1763 SIDER, Theodore. 2001. *Four-Dimensionalism: An Ontology of Persistence and Time*.
1764 Oxford: Oxford University Press, doi:10.1093/019924443X.001.0001.
- 1765 THOMASSON, Amie L. 2019. “The Ontology of Social Groups.” *Synthese* 196(12):
1766 4829–4845, doi:10.1007/s11229-016-1185-y.
- 1767 TOOLEY, Michael. 2001. “Functional Concepts, Referentially Opaque Contexts, Causal
1768 Relations, and the Definition of Theoretical Terms.” *Philosophical Studies* 105(3):
1769 251–279, doi:10.1023/a:1010381127525.
- 1770 UZQUIANO, Gabriel. 2018. “Groups: Toward a Theory of Plural Embodiment.” *The*
1771 *Journal of Philosophy* 115(8): 423–452, doi:10.5840/jphil2018115825.
- 1772 WILHELM, Isaac. 2020. “The Stage Theory of Groups.” *Australasian Journal of Philoso-*
1773 *phy* 98(4): 661–674, doi:10.1080/00048402.2019.1699587.
- 1774 —. 2022. “Pluralities, Counterparts, and Groups.” *Philosophical Studies* 179(7):
1775 2133–2153, doi:10.1007/s11098-021-01755-5.
- 1776 WITT, Charlotte. 2011. *The Metaphysics of Gender*. New York: Oxford University Press,
1777 doi:10.1093/acprof:oso/9780199740413.001.0001.
- 1778 YOUNG, Iris Marion. 1990. *Justice and the Politics of Difference*. Princeton, New Jersey:
1779 Princeton University Press. Second edition Young (2012).
- 1780 —. 2012. *Justice and the Politics of Difference*. 2nd ed. Princeton, New Jersey: Princeton
1781 University Press. With a foreword by Danielle S. Allen, doi:10.2307/j.ctvcmm4g4q.

PROOF

1782

Causal Inference from Big Data? A Reply to Pietsch (2021)

SERENA GALLI

1783 In his book *Big Data* (2021), Wolfgang Pietsch defends the view that vari-
1784 ational induction, which stands in the tradition of Mill's methods, allows
1785 us to derive conclusions about causal relationships from observational
1786 data and that the algorithms that are most successfully applied for big
1787 data analysis implement variational induction. In his opinion, the fact
1788 that the analysis of big data by machine learning algorithms enables
1789 reliable predictions and effective interventions in the world supports the
1790 assumption that these algorithms correctly identify causal relationships.
1791 In the first part of the paper, I argue that attempts to infer causal relation-
1792 ships from observational data by variational induction face fundamental
1793 difficulties. Furthermore, I contend that these difficulties are not due
1794 to the specific way in which the method is spelled out but are manifes-
1795 tations of a general underdetermination problem. In the second part, I
1796 consider Pietsch's claim that the practical benefit of big data approaches
1797 indicates that variational induction implemented by machine learning
1798 algorithms generates causal knowledge. I provide a critical assessment
1799 of his notion of causal knowledge, and I argue that his conclusion relies
1800 on an inaccurate depiction of scientific practice.

1801 In developing his views on variational induction, Pietsch relies on a difference-
1802 making account of causation. More specifically, he defines causal relevance
1803 and causal irrelevance as three-place relations between two variables and a
1804 given context or background:

1805 In a context B , in which a condition A and a phenomenon C occur,
1806 A is causally relevant to C , in short $A \mathcal{R} C \mid B$, iff the following
1807 counterfactual holds: if A had not occurred, C would also not
1808 have occurred.

1809 In a context B , in which a condition A and a phenomenon C occur,
1810 A is causally irrelevant to C , in short $A \mathcal{J} C \mid B$, iff the following

1811 counterfactual holds: if A had not occurred, C would still have
 1812 occurred. (Pietsch 2016b, 5)

1813 The truth value of the defining counterfactual statement is assessed in terms
 1814 of difference-making, taking into account instances that are or were realized
 1815 in our world (Pietsch 2016a, 11).¹ Methodologically, this assessment rests on
 1816 the framework of variational induction, which stands in the tradition of Mill's
 1817 (1889, 253ff) methods and comprises two key methods, namely, the method of
 1818 difference and the strict method of agreement. To determine if a circumstance
 1819 A is causally relevant for a phenomenon C with respect to background B, the
 1820 *method of difference* must be employed:

1821 If two instances with the same background B are observed, one
 1822 instance, in which circumstance A is present and phenomenon
 1823 C is present, and another instance, in which circumstance A is
 1824 absent and phenomenon C is absent, then A is causally relevant
 1825 to C with respect to background B, iff B guarantees homogeneity.
 1826 (Pietsch 2021, 33)²

1827 Simply put, the homogeneity of the background ensures that all the circum-
 1828 stances that are potentially causally relevant for phenomenon C are held fixed,
 1829 except for circumstance A, whose influence on phenomenon C is explicitly
 1830 studied.³

1831 In contrast, the *strict method of agreement* allows us to identify relations of
 1832 causal irrelevance:

1833 If two instances with the same background B are observed, one
 1834 instance, in which circumstance A is present and phenomenon
 1835 C is present, and another instance, in which circumstance A is
 1836 absent and phenomenon C is still present, then A is causally
 1837 irrelevant to C with respect to background B, iff B guarantees
 1838 homogeneity. (Pietsch 2021, 33)

-
- 1 This conception of counterfactual statements differs fundamentally from traditional counterfactual approaches to causation, such as those advanced by Lewis, who analyzes the truth conditions of counterfactual statements by referring to possible worlds (1973, 560–561).
- 2 In principle, causal relationships between continuous variables can be established likewise by extending the framework of variational induction by the method of concomitant variation (Pietsch 2021, 34). In the following, I will be concerned with binary variables exclusively.
- 3 I examine the homogeneity condition more closely in the context of epiphenomena. For a detailed discussion, cf. Pietsch (2021, 33–34; 2016b, 11–13).

1839 For variational induction to yield reliable results, several conditions must
 1840 be fulfilled. Most importantly, (i) all variables that are potentially causally
 1841 relevant for the phenomenon of interest must be known, and (ii) the dataset
 1842 must contain a sufficiently large number of observations covering all relevant
 1843 constellations of the variable values. Pietsch acknowledges that, due to the fact
 1844 that he makes these presuppositions, his account is (what he calls) *externally*
 1845 *theory-laden*. However, he contends that his account avoids *internal theory-*
 1846 *ladenness*, i.e., assumptions about causal connections between the variables
 1847 considered. In other words, he claims to avoid the kind of theory-ladenness
 1848 that is distinctive of hypothesis-driven approaches.⁴ With the framework of
 1849 variational induction, by contrast, the causal structure of the phenomenon of
 1850 interest is supposed to be elaborated from the data alone.⁵

1851 **1 Causal Inference by Variational Induction and the** 1852 **Underdetermination Problem**⁶

1853 To point out the fundamental difficulty of inferring causal relationships by
 1854 variational induction, I first focus on more complex causal structures, namely,
 1855 (i) *symmetric overdetermination and preemption* and (ii) *epiphenomena*. Then,
 1856 I evaluate whether the (iii) *directionality* of the relation of causal relevance
 1857 can be established or not. For this assessment, I take for granted that the
 1858 above-mentioned conditions for variational induction are met. In particular,
 1859 I assume that every possible constellation of variable values that could have
 1860 been generated by the causal structure in question is indeed observed and,
 1861 moreover, that the set of observations involves neither measurement errors
 1862 nor accidentally correlating variables.

1863 (i) Let us consider the following dataset consisting of observations 1–4,
 1864 which all share the same background B:

4 A prominent advocate of such an approach is Pearl, who maintains that “causal questions can never be answered from data alone” and that answering those questions “require[s] us to formulate a model of the process that generates the data, or at least some aspects of that process,” also in the context of big data (Pearl and Mackenzie 2018, 351).

5 If the requirements for variational induction are met, “then there are enough data to avoid spurious correlations and to map the causal structure of the phenomenon without further internal theoretical assumptions about the phenomenon” (Pietsch 2015, 910–911). See also Pietsch (2021, 65–66).

6 Woodward uses the term *underdetermination problem* to refer to the circumstance that, given a set of variables, different causal structures encompassing these same variables can generate an identical pattern of correlations and conditional correlations (2003, 106–107).

- 1865 Observation 1: $A_1 \& A_2 \& C$
- 1866 Observation 2: $A_1 \& \neg A_2 \& C$
- 1867 Observation 3: $\neg A_1 \& A_2 \& C$
- 1868 Observation 4: $\neg A_1 \& \neg A_2 \& \neg C$

1869 Further, let us suppose that the causal relationships that do, in fact, underlie
 1870 these observations are those depicted by the model in figure 1a. In this model,
 1871 we have two potential causes of C, A_1 and A_2 , with A_1 preempting A_2 when
 1872 both obtain. How does Pietsch propose to deal with this dataset? *Preempted*
 1873 or *alternative causes* such as A_2 require other circumstances, in this case A_1 ,
 1874 to be absent in order to have an impact on the phenomenon of interest. He
 1875 claims that such alternative causes can be singled out based on the method
 1876 of difference and the method of strict agreement too, but the background B
 1877 must be specified by an additional condition X, the *preempting cause*:

1878 A is an ‘alternative cause’ to C with respect to background B, iff
 1879 there exists an X such that A is causally relevant to C with respect
 1880 to a background $B \& \neg X$, but causally irrelevant to C with respect
 1881 to a background $B \& X$ (i.e., C is always present in $B \& X$). (Pietsch
 1882 2021, 34)

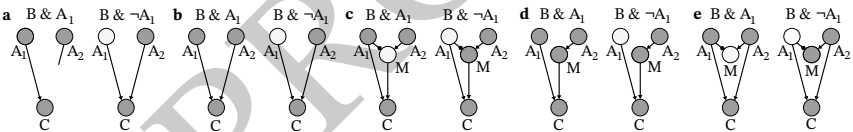


Figure 1: Symmetric overdetermination and preemption. In scenario (a), the exact mechanism of preemption is not specified and, therefore, symbolized by the prematurely terminated line originating from A_2 .

1883 Since A_2 is a preempted cause of C, it is causally relevant to C with respect to B
 1884 only in the absence of A_1 , which can be deduced by contrasting observations 3
 1885 and 4: $A_2 \mathcal{R} C \mid B \& \neg A_1$. Comparing observations 1 and 2, A_2 turns out to be
 1886 causally irrelevant for C with respect to $B \& A_1$, relying on the strict method of
 1887 agreement ($A_2 \mathcal{I} C \mid B \& A_1$). Yet, two problems arise from Pietsch’s approach
 1888 for the identification of alternative causes: First, A_1 , which is the *preempting*
 1889 *cause* of C, is determined to be causally irrelevant for C when applying the
 1890 strict method of agreement to observations 1 and 3. As seen before, it is only

1891 in the absence of A_2 that the causal relevance between A_1 and C with regard
 1892 to background B can be established ($A_1 \mathcal{R} C \mid B \ \& \ \neg A_2$, observations 2 and 4).
 1893 However, by definition, the impact of A_1 on C does not depend on the value
 1894 of the variable A_2 . Thus, variational induction fails to depict the asymmetry
 1895 between the preempting and the preempted cause. Second, the above-listed
 1896 observations 1–4 are compatible with another underlying causal structure,
 1897 namely a model of symmetric overdetermination, as displayed in figure 1b.
 1898 Hence, without any prior knowledge about the causal connections between
 1899 the variables, it is impossible to decide which causal structure really underlies
 1900 the observed constellations of variable values.

1901 Then, a common feature of big data is its high dimensionality, meaning that
 1902 each observation includes numerous different variables. So, it could be put
 1903 forward that this problem only arises because the dataset is not sufficiently
 1904 complex and not enough variables were regarded. For example, introduc-
 1905 ing the variable M , which mediates the causal effect of A_2 on C and whose
 1906 instantiation is prevented in the presence of A_1 , would definitively allow
 1907 to distinguish between the case of preemption and the case of symmetric
 1908 overdetermination, as depicted in figures 1c and 1d.⁷

1909 Observation 5: $A_1 \ \& \ A_2 \ \& \ \neg M \ \& \ C$
 1910 Observation 6: $A_1 \ \& \ \neg A_2 \ \& \ \neg M \ \& \ C$
 1911 Observation 7: $\neg A_1 \ \& \ A_2 \ \& \ M \ \& \ C$
 1912 Observation 8: $\neg A_1 \ \& \ \neg A_2 \ \& \ \neg M \ \& \ \neg C$

1913 Observation 5 is indeed not compatible with the model of symmetric overde-
 1914 termination in figure 1d because, according to that model, the variable A_1 has
 1915 no impact on the other cause A_2 or its mediating variable M . Yet, these four
 1916 observations are consistent with another model of symmetric overdetermina-
 1917 tion, where the instantiation of M depends both on the presence of A_2 and
 1918 the absence of A_1 , but M does not mediate the causal effect of A_2 on C , as
 1919 illustrated in figure 1e. Hence, including more variables does not solve, but,
 1920 at best, deflects the above-mentioned difficulties.

1921 (ii) In connection with *epiphenomena*, similar problems arise. Epiphenom-
 1922 ena, such as A_4 in figure 2c, have a common cause with the phenomenon of
 1923 interest but have no causal impact on it. Let us consider another dataset con-

⁷ According to Lewis' terminology, the causal model displayed in figure 1c is an example of early preemption (1986b, 200).

1924 sisting of observations 9 and 10, which, in turn, share the same background
 1925 B:

1926 Observation 9: $A_3 \ \& \ A_4 \ \& \ C$

1927 Observation 10: $\neg A_3 \ \& \ \neg A_4 \ \& \ \neg C$

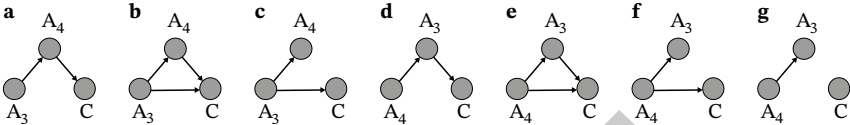


Figure 2: Epiphenomena and directionality of causal connections.

1928 Then, let us suppose that these observations, which are compatible with all
 1929 the models depicted in figure 2a-f, were generated by the causal model in 2c.
 1930 Since A_3 and A_4 covary, neither the method of difference nor the strict method
 1931 of agreement can be applied to determine whether A_4 stands in a relation of
 1932 causal relevance or irrelevance to C with respect to background B .⁸ By contrast,
 1933 A_3 proves to be causally relevant to C with respect to B , which, in this case,
 1934 does guarantee homogeneity.⁹ On behalf of Pietsch, it could be put forward
 1935 that this problem can be circumvented by considering the combination of
 1936 the variables A_3 and A_4 instead of examining them separately. Following
 1937 this approach, the method of difference establishes either the conjunction
 1938 or the disjunction of A_3 and A_4 to be causally relevant for C with respect to
 1939 background B ($A_3 \ \& \ A_4 \ \mathcal{R} \ C \ | \ B$ or $A_3 \ \vee \ A_4 \ \mathcal{R} \ C \ | \ B$).¹⁰ Still, in all cases,
 1940 variational induction fails to establish that A_4 is causally irrelevant to C
 1941 with respect to B .

1942 Pietsch acknowledges the issue that algorithms employing variational in-
 1943 duction may mistakenly single out epiphenomena as causally relevant for a

8 A_3 , which is potentially causally relevant for C as well, cannot be held fixed, as it strictly covaries with A_4 . Therefore, B does not guarantee homogeneity.

9 B guarantees homogeneity with respect to the relationship between A_3 and C if “only circumstances that are causally irrelevant to C can change” or that “lie on a causal chain through A_3] to C or that are effects of circumstances that lie on this causal chain” (Pietsch 2021, 33–34). Since A_4 is an effect of A_3 , B does guarantee homogeneity, although A_4 cannot be held fixed. However, presuming that A_4 is connected to A_3 in this way contradicts Pietsch’s claim that his account avoids assumptions about causal connections between the variables considered.

10 Such Boolean expressions are, as Pietsch maintains, a possible result of variational induction (2021, 50).

1944 phenomenon of interest, although they lack the difference-making character
 1945 of a cause (Pietsch 2021, 55; 2016a, 153–154).¹¹ However, he attributes it to the
 1946 fact that either the dataset is incomplete or the algorithm does not fully imple-
 1947 ment variational induction. By contrast, this example demonstrates that this
 1948 erroneous conclusion is not due to missing observations because it is drawn
 1949 despite considering all observations compatible with a given causal model.
 1950 On the other hand, it cannot be ascribed to the algorithmic implementation
 1951 either, as the manual, non-algorithmic application of variational induction
 1952 does not satisfyingly deal with epiphenomena either.¹²

1953 (iii) Finally, the fact that observations 9–10 could have been generated by a
 1954 causal structure with A_3 being a cause of A_4 (figures 2a–c) or a causal structure
 1955 with A_4 being a cause of A_3 (figures 2d–f) demonstrates that the *direction of*
 1956 *the relation of causal relevance* cannot be established by variational induction.
 1957 Besides, the same holds true for phenomenon C, which could as well be a cause
 1958 and not an effect of variable A_3 or A_4 if not predefined as the phenomenon
 1959 of interest. To solve this problem, Pietsch has suggested introducing a time
 1960 index for the phenomenon of interest and the variables examined (2014,
 1961 424). Yet, from a conceptual point of view, this seems like an ad hoc solution
 1962 since the truth condition he specifies for the counterfactual defining causal
 1963 relevance fails to capture the asymmetry in the relation of causal relevance.
 1964 Additionally, this solution is not practicable for cross-sectional data, where all
 1965 the variables are recorded at the same time and, accordingly, the timepoint of
 1966 their occurrence remains unknown.

1967 In light of these difficulties, Pietsch’s assertion that it is possible “to deter-
 1968 mine the true causal relationships by means of variational induction” seems
 1969 to be unwarranted (2021, 61). In the causal discovery literature, it is well es-
 1970 tablished that, given the causal Markov condition and the causal faithfulness
 1971 condition, certain features of the underlying causal structure can be deduced
 1972 from the probability distribution in the data. But, aside from special cases, it

11 Strictly speaking, he refers to *proxies*, which I take to be the equivalent of epiphenomena.

12 Needless to say, some of such wrong conclusions can be traced back to issues regarding data acquisition. For instance, a sampling error can result in an accidental correlation between variables that is not present in the population from which the sample was drawn. Let us suppose that observations 9–10 were generated by the causal structure displayed in figure 2g. In this case, the observed correlation between A_3 and A_4 , on one side, and C, on the other side, must have occurred by chance. Yet, by employing variational induction, the conjunction or disjunction of A_3 and A_4 is mistakenly identified as causally relevant for C with respect to B, and this misattribution can be recognized as such and corrected only when analyzing another, possibly larger dataset devoid of this accidental correlation.

1973 is not possible to uniquely determine the true causal structure.¹³ In my view,
 1974 variational induction similarly faces the problem of underdetermination as
 1975 it rests ultimately on the analysis of patterns of (conditional) dependencies
 1976 and independencies in the data. That is to say, variational induction aims at
 1977 identifying the constellation of variables V that has difference-making char-
 1978 acter with respect to background B : If this exact configuration of variables
 1979 is present, C is always present as well (i.e., $P(C | V) = 1$); in its absence, C is
 1980 always absent as well (i.e., $P(C | \neg V) = 0$). From this dependence between
 1981 C and V , a relation of causal relevance is inferred ($V \mathcal{R} C | B$). Thus, the
 1982 procedure of variational induction can be viewed as the comparison of the
 1983 conditional probabilities of C rather than the comparison of individual obser-
 1984 vations. From the pattern of conditional probabilities based on observations 9
 1985 and 10, for example, it can be deduced that it is either the conjunction or the
 1986 disjunction of A_3 and A_4 that makes a difference for the value of the variable
 1987 C with respect to background B (i.e., $V = (A_3 \& A_4)$ or $V = (A_3 \vee A_4)$).¹⁴ While
 1988 these two Boolean expressions are highly useful for predicting the value of C ,
 1989 the pattern of dependencies is, as already stated for the direct comparison of
 1990 individual observations, compatible with all the models depicted in figures
 1991 2a–f. Accordingly, Pietsch’s claim that “the difference-making circumstances
 1992 identified by variational induction are exactly the circumstances that need to
 1993 be manipulated or intervened upon in order to change a phenomenon” does
 1994 not seem justified. Although a given configuration of circumstances might
 1995 unequivocally determine the value of the phenomenon of interest C in an
 1996 observational setting, the exact connection between these circumstances and
 1997 C remains elusive. Therefore, successful intervention strategies cannot be
 1998 deduced from the Boolean expression of these circumstances. While in 2a,
 1999 2d, 2e, and 2f a single intervention on A_4 is an effective way of manipulating
 2000 the value of C , in 2c this is clearly not the case. The Boolean expression may

13 These two conditions are so-called *bridge principles*, which are required to connect the observations of a given set of variables to the underlying causal model that generated these observations. More specifically, the causal Markov condition allows the inference from a probabilistic dependence between two variables to a causal connection, whereas the causal faithfulness condition allows the inference from a probabilistic independence to causal separation. Cf. Eberhardt (2017, 82–85). For a discussion of underdetermination in causal inference in relation to different success criteria and background assumptions, see Zhang (2009).

14 Since the configurations $(A_3 \& \neg A_4)$ and $(\neg A_3 \& A_4)$ do not occur in a purely observational setting, these two possibilities cannot be distinguished.

2001 encompass the necessary and together sufficient circumstances for *observing*
 2002 phenomenon C but not for *producing* it.¹⁵

2002 **Objectives of Big Data Analysis and Causal Knowledge**¹⁶

2004 Pietsch distinguishes two central functions of big data approaches, namely,
 2005 *prediction* and *intervention*, and claims that the exertion of both requires some
 2006 access to causal knowledge.¹⁷ Arguably, the view that causal knowledge is
 2007 indispensable for effectively manipulating a phenomenon of interest is hardly
 2008 contested. However, causal knowledge is usually not considered a prerequisite
 2009 for predictive success.¹⁸ In that regard, it is useful to touch upon Pietsch's
 2010 notion of causal knowledge, which bears on his distinction of direct and indi-
 2011 rect causal connections: If a certain variable is causally relevant for another
 2012 variable in a given context, as it is the case for A_3 and C in figure 2c, the rela-
 2013 tionship between those two variables constitutes a *direct causal connection*,
 2014 as Pietsch suggests. If, by contrast, two variables are not causally relevant for
 2015 one another but are related via a common cause, then there exists an *indirect*
 2016 *causal connection* between these two variables, as it is the case for A_4 and C in

-
- 15 As an anonymous reviewer pointed out, a promising way of dealing with this problem of underdetermination is the appeal to theoretical virtues such as parsimony. For example, Forster et al. introduce the principle of frugality that favors those causal structures with the fewest causal connections (2018). I fully agree that, technically, the procedure of variational induction could be combined with an algorithm that ranks the possible causal structures in terms of simplicity. Yet, this constraint regarding the total number of causal connections involves an assumption about the causal connections between the variables since causal models with more numerous connections, such as 2b or 2e, are dismissed in favor of models with fewer connections, although perfectly compatible with the data. Therefore, this strategy is *internally theory-laden* and not reconcilable with the concept of variational induction as a purely data-driven approach. An alternative strategy to determine the true causal structure is experimentation. For a detailed discussion of experimentation as a means for resolving underdetermination, cf. Eberhardt (2013).
- 16 In this section, which is concerned with variational induction as a means of causal inference from big data specifically, I acknowledge Pietsch's claim that the most successful algorithms are based on variational induction without further examination.
- 17 Rather than distinguishing between different functions, I would propose to differentiate between two questions that are to be answered by big data analysis. To specify *intervention* as a function of big data approaches presupposes what is under consideration. Besides, it remains unclear how to discern which function, prediction or intervention, is exerted in a given case.
- 18 For example, Woodward maintains that accurate predictions can be made based on correlations solely; furthermore, he points out that "inferences from effect to cause are often more reliable than inferences from cause to effect" (2003, 31–32).

2017 figure 2c, which are both effects of A_3 (Pietsch 2021, 55).¹⁹ Certainly, only for
 2018 successful interventions upon a phenomenon of interest must it be known
 2019 whether there is a direct causal connection between that same phenomenon
 2020 and the variable that is to be manipulated or not. But, as he argues, for accu-
 2021 rate predictions, either such a direct causal connection or an indirect causal
 2022 connection between the phenomenon of interest and a potential predictor
 2023 variable is required. Thus, when a machine learning algorithm singles out a
 2024 variable as a promising predictor variable for a given phenomenon of interest,
 2025 the algorithm thereby generates causal knowledge to a certain degree. In my
 2026 opinion, this broad notion of causal knowledge allowing for different degrees
 2027 is particularly problematic in three respects:

2028 (i) First of all, as a cause (usually) correlates with the phenomenon of
 2029 interest, so does an epiphenomenon of this cause. The first correlation is
 2030 indicative of a (direct) causal connection, whereas the second is indicative
 2031 of a common cause. As discussed for epiphenomena, variational induction
 2032 does not allow us to distinguish between a correlation ascribable to a direct
 2033 causal connection and a correlation ascribable to a common cause. It follows
 2034 that not only variational induction but also the analysis of (conditional) cor-
 2035 relations yields causal knowledge in this wide sense. Accordingly, it does not
 2036 seem consistent to specify correlation as a contrasting notion for causation.
 2037 Furthermore, since the procedure of variational induction makes use of the
 2038 pattern of dependencies in the data, it does not even allow for a distinction
 2039 between correlations that are indicative of some sort of causal connection and
 2040 purely accidental correlations. Therefore, it remains unclear in what sense
 2041 big data algorithms are capable of delimiting causation from correlation, as
 2042 Pietsch maintains.²⁰

2043 (ii) This broad notion of causal knowledge stands in tension with Pietsch's
 2044 claim that the primary function of causal knowledge is to guide us on how to
 2045 effectively intervene in the world (2021, 54). If the knowledge of an indirect
 2046 causal connection between two variables is regarded as causal knowledge
 2047 as well, having access to causal knowledge in this wide sense does not help
 2048 to discriminate between effective and ineffective strategies to manipulate a
 2049 phenomenon of interest.

19 Pietsch's distinction of direct and indirect causal connections differs from the conventional view, according to which the difference between direct and indirect causal connection results from the absence or presence of a mediating variable. See, for example, Woodward (2003, 55).

20 "By relying on variational induction, big data approaches are to some extent able to distinguish causation from correlation" (Pietsch 2021, 57).

2050 (iii) And, finally, it risks obscuring the distinction between questions of
2051 prediction and questions of intervention, which are addressed in scientific
2052 practice: Causal knowledge in the strict sense, that is to say, knowledge about
2053 relations of causal relevance and irrelevance in a given set of variables, is no
2054 precondition for predictions. Thus, the fact that an algorithm implementing
2055 variational induction yields accurate predictions cannot be cited in support
2056 of the view that variational induction is capable of establishing causal rela-
2057 tionships. Conversely, interventions indeed depend on causal knowledge in
2058 the strict sense. If such an algorithm truly did enable us to efficiently inter-
2059 vene in the world, this could speak in favor of Pietsch’s view that variational
2060 induction is capable of inferring direct causal connections. As an example,
2061 he refers to “algorithms [that] are designed to determine the best medicine
2062 to cure a certain cancer” (Pietsch 2021, 54).²¹ In fact, there are a number
2063 of studies that relied on machine learning in order to predict the response
2064 to a given drug. In a recently published work, the tumor tissue of patients
2065 with breast cancer was analyzed with different methods at diagnosis (Sam-
2066 mut et al. 2022). Patients were subsequently treated with chemotherapy, and
2067 treatment response was evaluated. Using a machine learning algorithm, the
2068 authors built a model to predict the response to chemotherapy, which was
2069 based on the molecular profile of the tumor as well as clinicopathological
2070 features, and model performance was successfully validated on a different
2071 dataset. Amongst other things, they drew the conclusion that patients pre-
2072 dicted to show a poor response to standard-of-care chemotherapy should be
2073 enrolled in clinical trials investigating novel therapies. Therefore, the results
2074 of this big data approach may allow for better stratification of patients that
2075 will or will not benefit from conventional chemotherapy and are inasmuch
2076 action-guiding. However, Pietsch maintains that these algorithms are, more-
2077 over, designed to *determine the best treatment* for a given cancer, in this way
2078 allowing us to effectively intervene upon the phenomenon of interest, namely,
2079 tumor growth. For the sake of argument, let us suppose the algorithm revealed
2080 that three signaling pathways are hyperactive in tumors poorly responding to
2081 chemotherapy compared to tumors displaying a good treatment response. But,
2082 as outlined above, it is impossible to determine if (or which of) these three
2083 pathways are indeed driving tumor growth and which are rather an epiphe-
2084 nomenon of the cause of excessive tumor growth or even a consequence
2085 thereof. Accordingly, the question whether one of these hyperactive signaling

21 He does not cite any specific publication to underpin his assertion.

2086 pathways truly constitutes a promising therapeutic target or not cannot be
 2087 answered based on the observational data alone but requires experimentation.
 2088 Besides, in order to successfully intervene in the world, it is essential not
 2089 only to identify the causes of the phenomenon but also to understand how
 2090 these causes can be manipulated, specifically which drug effectively targets
 2091 a given pathway (compare figure 3). This kind of causal knowledge may be
 2092 generated in randomized controlled trials or *in vitro* studies but, again, cannot
 2093 be derived from observational data. To trace back the practical benefit of big
 2094 data approaches to the generation of causal knowledge by the algorithms
 2095 used does not accurately reflect the scientific practice, which builds upon
 2096 different sources of knowledge to determine effective interventions.

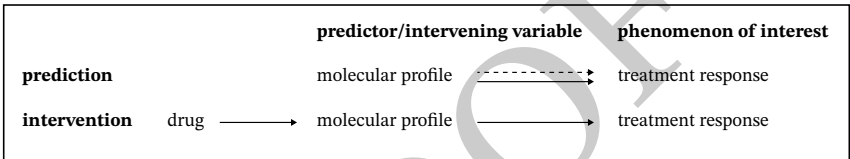


Figure 3: Prediction of and intervention upon a phenomenon of interest. Indirect causal connections are represented by dashed arrow lines, direct causal connections by solid arrow lines.

2093 3 Conclusions

2098 In my view, variational induction fails to elucidate causal structures involving
 2099 preemption, symmetric overdetermination, or epiphenomena, establishing
 2100 causal relationships between variables that actually are conjoined in a relation
 2101 of causal irrelevance and vice versa. Furthermore, the direction of the relation
 2102 of causal relevance cannot be specified by variational induction, which poses
 2103 a problem for even the simplest causal models possible. These shortcomings
 2104 are neither specific to the method of variational induction nor ascribable to
 2105 an imperfect dataset with missing observations, an insufficient number of
 2106 observed variables, or any measurement errors. Rather, the attempt to infer
 2107 causal relationships from observational data (including big data) itself faces
 2108 important limitations: Since for a given set of observations multiple underlying
 2109 causal structures are usually conceivable, generally it is impossible to uniquely
 2110 determine the true causal model from this set of observations alone without

2111 endorsing any background assumptions about or having any prior knowledge
2112 of the causal relationships between the variables involved.

2113 Pietsch's notion of causal knowledge explains, at least partially, why he
2114 reaches another assessment of variational induction as a method for generat-
2115 ing causal knowledge. Supposing a broader notion of causal knowledge, he
2116 seems to have in mind a less strict success criterion: It suffices for variational
2117 induction to approximate causal relationships, namely, to determine if there
2118 is any causal connection, direct or indirect, between two variables. This could
2119 create the appearance that the conflicting assessment of variational induction
2120 as a means to infer causal relationships is merely due to two divergent, equally
2121 plausible notions of causal knowledge. However, in my opinion, Pietsch's
2122 broad notion of causal knowledge is problematic because it blurs the distinc-
2123 tion between causation and correlation and between the prerequisites for
2124 prediction and for intervention.

2125 If the practical benefit of big data approaches cannot be attributed to the
2126 elucidation of causal relationships, an alternative explanation is needed. The
2127 identification of predictive markers may indeed improve patient care by spar-
2128 ing patients who are unlikely to respond to the adverse reactions of an ineffec-
2129 tive treatment. Randomized controlled trials can yield negative results only
2130 because patients are not selected appropriately. This could be obviated by a
2131 more adequate patient stratification based on reliable predictor variables. The
2132 analysis of the molecular profile of a tumor can generate promising hypothe-
2133 ses about chemoresistance in a relatively unbiased way, which may prove to
2134 be true in experimental assays. Undoubtedly, the results of machine learning
2135 algorithms contain very valuable information, which, in conjunction with
2136 knowledge derived from other sources, provide reasons to act in a certain way.
2137 In this sense, Pietsch is right in stating that precisely data-rich sciences such as
2138 medicine are fundamentally concerned with difference-making relationships
2139 and that the correlations unveiled by machine learning algorithms certainly
2140 do not replace causation. But, although such results of big data analysis can
2141 be action-guiding and aid in singling out potentially effective interventions,
2142 this should not be taken as a confirmation of the claim that machine learning
2143 algorithms indeed elucidate the causal structure underlying the observational
2144 dataset.*

* I thank Peter Schulte and two anonymous reviewers for their helpful comments on an earlier version of this paper.

Serena Galli

 0000-0002-9630-8645

University of Zurich

serenacarolina.galli@uzh.ch

References

- 2149
- 2150 EBERHARDT, Frederick. 2013. "Experimental Indistinguishability of Causal Structures." *Philosophy of Science* 80(5): 684–696, doi:10.1086/673865.
- 2151 —. 2017. "Introduction to the Foundations of Causal Discovery." *International Journal of Data Science and Analytics* 3(2): 81–91, doi:10.1007/s41060-016-0038-6.
- 2152
- 2153 FORSTER, Malcolm R., RASKUTTI, Garvesh, STERN, Reuben and WEINBERGER, Naf-tali. 2018. "The Frugal Inference of Causal Relation." *The British Journal for the Philosophy of Science* 69(3): 821–848, doi:10.1093/bjps/axw033.
- 2154
- 2155 LEWIS, David. 1973. "Causation." *The Journal of Philosophy* 70(17): 556–567. Reprinted, with a postscript (Lewis 1986b), in Lewis (1986a, 159–172), doi:10.2307/2025310.
- 2156
- 2157 —. 1986a. *Philosophical Papers, Volume 2*. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.
- 2158
- 2159 —. 1986b. "Postscript to Lewis (1973)." in *Philosophical Papers, Volume 2*, pp. 172–213. Oxford: Oxford University Press, doi:10.1093/0195036468.003.0006.
- 2160
- 2161 MILL, John Stuart. 1843. *A System of Logic, ratiocinative and inductive*. London: Long-mans, Green & Co.
- 2162
- 2163 —. 1889. *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. London: Long-mans, Green & Co. First edition: Mill (1843).
- 2164
- 2165 PEARL, Judea and MACKENZIE, Dana. 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books, <http://bayes.cs.ucla.edu/WHY/>.
- 2166
- 2167 PIETSCH, Wolfgang. 2014. "The Structure of Causal Evidence Based on Eliminative Induction." *Topoi* 33(2): 421–435, doi:10.1007/s11245-013-9190-y.
- 2168
- 2169 —. 2015. "Aspects of Theory-Ladenness in Data-Intensive Science." *Philosophy of Science* 82(5): 905–916, doi:10.1086/683328.
- 2170
- 2171 —. 2016a. "The Causal Nature of Modeling with Big Data." *Philosophy & Technology* 29(2): 137–171, doi:10.1007/s13347-015-0202-2.
- 2172
- 2173 —. 2016b. "A Difference-Making Account of Causation." Unpublished manuscript, <https://philsci-archive.pitt.edu/id/eprint/11913>.
- 2174
- 2175 —. 2021. *Big Data*. Elements in the Philosophy of Science. Cambridge: Cambridge University Press, doi:10.1017/9781108588676.
- 2176
- 2177 SAMMUT, Stephen-John, CRISPIN-ORTUZAR, Mireia, CHIN, Suet-Feung, PROVENZANO, Elena, BARDWELL, Helen A., MA, Wenxin, COPE, Wei, et al. 2022. "Multi-Omic Ma-
- 2178
- 2179
- 2180
- 2181

- 2182 chine Learning Predictor of Breast Cancer Therapy Response.” *Nature* 601(7894):
2183 623–629, doi:10.1038/s41586-021-04278-5.
- 2184 WOODWARD, James F. 2003. *Making Things Happen: A Theory of Causal Explanation*.
2185 Oxford Studies in the Philosophy of Science. Oxford: Oxford University Press,
2186 doi:10.1093/0195155270.001.0001.
- 2187 ZHANG, Jiji. 2009. “Underdetermination in Causal Inference.” *Studies in Logic* 2(4):
2188 16–47.

PROOF

PROOF

Retro-Closure Principle and Omniscience

CIRO DE FLORIO & ALDO FRIGERIO

Todd and Rabern (2021) have argued that if we assume that future contingents are untrue and if we accept the Retro-closure principle ($p \rightarrow \mathbf{PF}(p)$), then the existence of a temporal omniscient entity becomes metaphysically impossible. Since the truth of a metaphysical and theological theory should not be dependent on questions of temporal semantics, Todd and Rabern conclude that, if one wishes to maintain that future contingents are untrue, one must abandon the Retro-closure principle. The aim of this paper is to propose a temporal semantics system in which future contingents are untrue, the Retro-closure principle is valid, and the possibility of the existence of an omniscient and temporal being is guaranteed.

The future is uncertain. We do not know if it will rain tomorrow or if a drought will persist. Over centuries, philosophers have been wondering whether this uncertainty concerns only the epistemic dimension—and, therefore, it is only the result of our ignorance—or whether it is rooted in the ontological dimension, such that the world itself is at least partly indeterminate. Those who believe that the future is—at least to a certain extent—indeterminate should account for this intuition through an adequate semantics of future tense sentences. Let us assume that the actual state of the world and its natural laws do not determine the weather of tomorrow. According to many future tense semantics, “It will rain” is not true in this situation. However, suppose that time passes, and the following day, it rains. What would be the truth value of the proposition expressed yesterday by the sentence, “It will rain tomorrow,” if considered from today’s perspective, given that it rains today? Many may have the intuition that this proposition is true today. Following Todd and Rabern (2021), we call the content of this intuition the Retro-closure principle (RCP). In fact, several future tense semantics—though not all—account for this intuition.

2219 Todd and Rabern (2021) have advanced an interesting argument against
 2220 what they call Open-closurism (OC), which is the conjunction of two theses:
 2221 1) Open futurism (OF), following their terminology, which refers to the view
 2222 that future contingents are untrue (i.e., false or neither true nor false); and 2)
 2223 RCP. Theirs is an indirect argument: they do not argue against OC directly but
 2224 instead show that OC is incompatible with the possibility of an omniscient
 2225 temporal being. Even though we do not want to be committed to the existence
 2226 of an omniscient temporal being, it is odd that semantic principles would
 2227 determine a complex metaphysical question, such as the possibility of an
 2228 omniscient entity. Todd and Rabern's conclusion is that we have to drop OC
 2229 and, consequently, those OF semantics that validate RCP.

2230 In this paper, we present a future tense semantics that accepts both OF and
 2231 RCP without implying the impossibility of an omniscient temporal being. As
 2232 a consequence, Todd and Rabern's objection fails within this semantics or
 2233 semantics similar in kind. As we will see below, one of the main features of our
 2234 semantics is the presence of two indices, the first of which indicates the time
 2235 at which a formula is evaluated, while the second indicates the perspective
 2236 from which it is evaluated.

2237 The paper is structured as follows: In section 1, we put forward some con-
 2238 siderations about the plausibility of RCP. In section 2, Todd and Rabern's
 2239 objection is presented in detail. In section 3, the branching time semantics is
 2240 introduced, and the difficulties of Thomason's supervaluationism in respond-
 2241 ing to Todd and Rabern's objection are considered. In section 4, a particular
 2242 OC semantics is presented. It is shown that this semantics allows for the
 2243 possibility of an omniscient temporal being, thus refuting Todd and Rabern's
 2244 argument. Finally, section 5 contains some closing remarks.

2245 1 Intuitions in Favour of RCP

2246 As indicated above, OC is—at least prima facie—a plausible view. In this
 2247 section, we would like to further elaborate on our intuitions about OC and
 2248 show that OC appears to be quite natural when it is “referred,” so to speak,
 2249 to the past and present; the extension to the future case is quite natural, and,
 2250 therefore, the sceptic about OC has much work to do.¹

1 In this paper, we will defend a particular version of OC for which future contingents are neither true nor false. There is another version of OC for which future contingents are false. The combination of this view with RCP seems to us much less plausible, so we will ignore this version of OC here.

2251 Suppose today is December 9th, 2022, and the world is such that it is not
2252 determined whether it will rain tomorrow. However, it is determined that
2253 tomorrow, Paul will believe the proposition expressed by “It’s raining today.” In
2254 other words, the current state of the world and the laws of nature determine a
2255 particular configuration of the neurons of Paul’s brain, such that on December
2256 10th, 2022, Paul believes the proposition expressed by the indexical sentence
2257 “It’s raining today.” Is Paul’s belief of tomorrow correct today? In other words,
2258 when Paul believes tomorrow that it is raining that day, will he have a correct
2259 belief? Intuitively, we can respond neither positively nor negatively to this
2260 question. The answer depends on how things will go: if it rains tomorrow, then
2261 Paul’s belief will be correct; if it does not rain, it will be incorrect. However,
2262 since it is not determined today whether it will rain tomorrow, it is also not
2263 determined whether Paul’s belief will be correct. Tomorrow, however, it will
2264 rain, or it will not rain, and Paul’s belief will be either correct or not. Therefore,
2265 the correctness of Paul’s belief in the proposition expressed by the sentence “It
2266 is raining today,” uttered on December 10th, is determined by what happens
2267 on December 10th and cannot be determined before that date.

2268 Let us now make a change to our initial scenario. Suppose that today is
2269 December 9th, but in this case, Paul is not determined to believe something
2270 tomorrow; rather, he believes something the day after tomorrow, on December
2271 11th. In particular, he is determined to believe the proposition expressed by
2272 the sentence “It rained yesterday.” Let us assume again that on December
2273 9th, it is undetermined whether tomorrow, December 10th, it will rain or not.
2274 Once again, we can ask whether Paul’s belief of December 11th that it had
2275 rained the day before would be correct. Once again, however, today, December
2276 9th, we can neither say that it is correct nor that it is incorrect because it is not
2277 determined whether it will rain tomorrow on December 10th. Only tomorrow
2278 will the correctness of that belief be determined. Tomorrow, we will be able
2279 to say whether Paul’s belief of the day after tomorrow will be correct or not.
2280 In fact, the correctness of Paul’s belief on the day after tomorrow depends on
2281 what happens tomorrow. However, today, it is indeterminate what will happen
2282 tomorrow, and, therefore, it is also indeterminate whether Paul’s belief will be
2283 correct. Nevertheless, time passes, and the world determines itself. Tomorrow,
2284 it will rain, or it will not rain. Consequently, from tomorrow’s perspective,
2285 Paul’s belief of December 11th, that it rained the day before, will be correct
2286 or incorrect. Therefore, the correctness of Paul’s belief in the proposition
2287 expressed by the sentence “It rained yesterday,” uttered on December 11th, is

2288 determined by what happens on December 10th and cannot be determined
2289 before that date.

2290 Now, suppose that Paul believes today, December 9th, that it will rain
2291 tomorrow. Again, it is not determined today whether it will rain tomorrow
2292 or not. Is Paul's belief correct or incorrect? By symmetry with the previous
2293 cases, we would say that from today's perspective, his belief is neither correct
2294 nor incorrect. The correctness of Paul's belief depends on what will happen
2295 tomorrow, but today, what will happen tomorrow is indeterminate; therefore,
2296 the correctness of Paul's belief is not determined either. However, the world
2297 moves forward and determines itself. Tomorrow, it will rain, or it will not rain.
2298 On the basis of what will happen tomorrow, it will be possible to say whether
2299 Paul's belief of December 9th was correct or not. If it rains, it was correct;
2300 otherwise, it was incorrect. Therefore, from tomorrow's perspective, Paul's
2301 belief will have been correct or incorrect. The correctness of Paul's belief in
2302 the proposition expressed by the sentence "It will rain tomorrow," uttered on
2303 December 9th, is determined by what happens on December 10th and cannot
2304 be determined before that date.

2305 Those who deny **OC** must deny that it is indeterminate today whether Paul's
2306 belief that it will rain tomorrow is correct or incorrect and that instead, from
2307 tomorrow's perspective, it can be said that it was correct or not. To deny this,
2308 they must either 1) deny that the correctness of Paul's beliefs is determined
2309 in the course of time or 2) deny the symmetry between the future case and
2310 the present and past cases. Let us consider these two possibilities in detail:

- 2311 (1) One could reject that the correctness of Paul's belief in the proposition
2312 expressed by the sentence "It is raining today," uttered on December
2313 10th, is determined only when the present is December 10th and cannot
2314 be determined before that date. This can be done in at least two ways:
2315 (i) Either it is claimed that it is already true today, December 9th, that
2316 such a belief is correct or incorrect; or (ii) it is denied that tomorrow,
2317 on December 10th, Paul's belief becomes correct or not. Since it is
2318 indeterminate today, on December 9th, whether it will rain tomorrow
2319 or not, (i) is only possible assuming that one of the possible futures is the
2320 true one. In other words, (i) is only possible when assuming a Thin Red
2321 Line semantics of the future. This means that **OC** must be abandoned.
2322 We will ignore this solution because our goal here is to defend the
2323 consistency of **OC** and, thus, the compatibility of **OF** with **RCP**, and not
2324 to argue for **OF**. Then, (ii) is completely implausible: suppose it rains

2325 tomorrow, and that Paul believes it rains. How can we deny that Paul's
2326 belief is correct? Suppose that someone says to Paul, "It is true that you
2327 believe that it is raining today, and it is true that it is raining today, but
2328 your belief is not correct because it was uncertain yesterday whether it
2329 would rain today, and therefore it was also uncertain yesterday whether
2330 your belief would be correct or not." Paul, and we with him, would find
2331 this reasoning absurd. Paul would probably say, "Regardless of how
2332 things were yesterday, it is a fact that it is raining today and that I believe
2333 it is raining today. Therefore, my belief about today is correct *today*."

- (2) More plausibly, the symmetry between past and present cases, on the
2334 one hand, and the future case, on the other hand, might be denied. It
2335 is true that the correctness of Paul's belief of December 10th, that it is
2336 raining that day, is determined on December 10th, and it is true that
2337 the correctness of Paul's belief of December 11th, that it rained the day
2338 before, is determined on December 10th. However, the correctness of
2339 Paul's belief of December 9th that it will rain the following day is not
2340 determined on December 10th. This amounts to saying that while the
2341 correctness of Paul's belief that it is raining today is determined by what
2342 happens today and that the correctness of Paul's belief that it rained
2343 yesterday is determined by what happened yesterday, the correctness of
2344 Paul's belief that it will rain tomorrow is not determined by what will
2345 happen tomorrow. However, it is hard to see why the case of the future
2346 should not be similar to the present and past cases.

2347 The only plausible argument to deny the symmetry between these cases
2348 could be this: in the first two cases, the passage of time determines the
2349 correctness of a present or future belief; in the third case, the passage
2350 of time determines the correctness of a past belief. However, the past is
2351 fixed and unchangeable; therefore, the flow of time cannot determine
2352 something in the past. Nevertheless, as Todd and Rabern (2021, 106)
2353 point out, a change in the correctness of a belief is an extrinsic change
2354 (or a so-called "Cambridge change"), not an intrinsic one. When the
2355 past is said to be fixed, it is usually assumed that it is fixed with respect
2356 to intrinsic changes. However, extrinsic changes do not seem to be barred
2357 by the fixity of the past. For instance, World War I acquired the property
2358 of having ended 21 years before the outbreak of World War II in 1939.
2359 However, acquiring this property poses no problem for the fixity of the
2360

2361 past because it is an extrinsic property. Similarly, it is not a problem for
 2362 the fixity of the past that a past belief becomes correct after the fact.²

2363 A further argument can be advanced in favour of OC; something similar to
 2364 this principle seems to be valid for other verb tenses as well, particularly for
 2365 the progressive. Bonomi (1997) gives the following example. Suppose that Leo
 2366 has just begun a journey in France. In the first stage, he drives from Milan to
 2367 Dijon, where he arrives on July 14th at a quarter to three p.m. He does not
 2368 stop there because he plans to spend his first night in France in one of the
 2369 following cities: Besançon, Metz, or Paris. Actually, he must make a decision
 2370 since three different routes correspond to these alternatives. However, at a
 2371 quarter to three p.m., while driving around the Dijon ring road, he has not
 2372 yet decided where to go; he is thinking it over because, for several reasons, all
 2373 these cities attract him in exactly the same way. In this situation, the following
 2374 propositions are arguably untrue if evaluated at a quarter to three p.m.:

- 2375 (i) Leo is going to Besançon.
- 2376 (ii) Leo is going to Metz.
- 2377 (iii) Leo is going to Paris.

2378 Since Leo has not yet decided where to go, none of these propositions is
 2379 privileged with respect to the others. Suppose, however, that Leo ultimately
 2380 decides to go to Besançon, where he arrives two hours later. Moreover, suppose
 2381 that traffic police have photographed Leo on the Dijon ring road at a quarter
 2382 to three p.m. The day after, someone asks, “What was Leo doing yesterday, at a
 2383 quarter to three p.m., when the picture was taken?” Intuitively, the following
 2384 answer seems to be correct:

- 2385 (iv) Leo was going to Besançon.

2386 The analogy with OC is obvious: while from the perspective of July 14th at a
 2387 quarter to three p.m., it is not true that Leo was going to Besançon at that time,
 2388 from the perspective of July 15th, it was true that Leo was going to Besançon
 2389 on July 14th at a quarter to three p.m. The progressive, moreover, has an
 2390 obvious connection with the future. According to the standard semantics
 2391 of the progressive (Landman 1992; Portner 1998), Prog(e) is true if a first

2 The correctness or incorrectness of past beliefs about the future are soft facts—that is, past facts that depend on future facts. For the distinction between soft and hard facts, cf. Hoffman and Rosenkrantz (1984), Todd (2013), and De Florio and Frigerio (2018).

2392 stage of event e has occurred so far, and if all will go on normally without
 2393 any unexpected interruptions, e will be completed in the future. Thus, our
 2394 intuitions about the truth values of (i) to (iv) provide further support for OC.

2392 OC and Omniscience

2396 Our intuitions in favour of OC are very strong; this means that our formal
 2397 semantics should validate it. Otherwise, we would be in the presence of a
 2398 conceptual tension; although the principle is intuitively valid, our logic fails to
 2399 characterise it. Indeed, many of the major semantic frameworks for branching
 2400 time account for OC. Specifically, it is valid in Thomason's supervaluationism
 2401 (cf. Thomason 1970) because $p \rightarrow \mathbf{PF}p$ holds in all histories, and supervalua-
 2402 tionism universally quantifies on all histories. It is also valid in MacFarlane's
 2403 relativism (cf. MacFarlane 2003, 2014), which indeed seems to be designed to
 2404 account for the intuition that $\mathbf{F}p$ is untrue with respect to yesterday's assess-
 2405 ment context but true with respect to today's assessment context. However,
 2406 OC is not valid in Peircean or Thin Red Line (TRL) semantics—in Peircean
 2407 semantics, OF is valid, but RCP is not, whereas in TRL semantics, neither OF
 2408 nor RCP are valid³—and this might be a problem for these semantics.

2409 Todd and Rabern (2021) are not convinced by OC. For this reason, they
 2410 construct an ingenious argument against it. Their strategy is indirect: if OC is
 2411 supposed to be valid, then counterintuitive consequences regarding the logic of
 2412 divine omniscience follow. In other words, the validity of OC is incompatible
 2413 with the existence of an omniscient and temporal entity. Todd and Rabern
 2414 are not committed to the actual existence of an omniscient entity, but—and
 2415 this is their point of argument—it is very odd that a question concerning
 2416 the semantics of future statements has metaphysical consequences for the
 2417 existence of omniscient entities. Our intuition suggests that the conditions for
 2418 the possibility of omniscient supernatural entities should not be related to the
 2419 structure of the future and RCP. However, if Todd and Rabern are right, and
 2420 if OC is accepted, there can be no omniscient temporal entity. Since this is a
 2421 metaphysical thesis, it follows that OC is a suspicious doctrine. Incidentally,
 2422 this is a point in favour of semantics that do not validate OC. In particular, if

3 For the invalidity of RCP in TRL semantics, cf. Belnap and Green (1994). It is possible to amend the TRL semantics in order to validate RCP. For instance, the TRL can be relativized to instants of time (cf. McKim and Davis 1976; Øhrstrøm 2009), or the TRL can be initialized at the post-semantic level rather than at the semantic level (cf. Iacona 2014; Wawer 2014; Wawer and Malpass 2020). We will ignore these issues here.

2423 **OF** is supposed, this constitutes a point in favour of Peirceanism, which does
2424 not validate **RCP**.

2425 Therefore, let us examine Todd and Rabern's argument in detail. We may,
2426 for ease of presentation, sometimes refer to the omniscient entity as "God";
2427 we also assume a doxastic operator of belief (**B**), which we intend to always
2428 refer to the omniscient subject, whereby Bp indicates that God believes p .
2429 Todd and Rabern establish the principle of omni-accuracy as follows:

2430 **OMNI-ACCURACY.** $p \leftrightarrow Bp^4$

2431 The **OMNI-ACCURACY** principle thus establishes a double implication between
2432 p and God's belief that p : every divine belief is true, and there is no truth
2433 that is not believed by God. Using **OMNI-ACCURACY** and **RCP**, we obtain the
2434 following:

2435 (v) $p \rightarrow P_1B(F_1p)^5$

2436 As a result, if it is true that it is raining today, then yesterday, God believed it
2437 would rain today. However, suppose yesterday it was not determined whether
2438 it would rain today. Then yesterday, God could not believe that today it would
2439 rain. In fact, yesterday F_1p was untrue because today's weather was indeter-
2440minate, and, therefore, God could not believe it.

2441 Advocates of **OC** have two possibilities at this point:

2442 (1) They can claim that yesterday it was not true that God believed F_1p ,
2443 but that today it is true that God believed it. This is equivalent to stating
2444 that the past changes with the passage of time. This does not seem

4 An aspect of **OMNI-ACCURACY** seems to be problematic; if p is true, then it is true that God believes p ; analogously, in the case in which it is false (and therefore $\neg p$ is true), we have that God does not believe p . If p is indeterminate, how does the principle behave? One might answer that it is indeterminate whether God believes p , but this seems strange; if p is an indeterminate proposition, then God should not believe it. However, the principle remains silent on this point. It could be argued that it is for this reason that Todd and Rabern introduce the principle of **OMNI-CORRECTNESS** (see below in the main text). However, this move does not seem to solve the problem of the indeterminacy of divine belief. Suppose that the truth value of p is undetermined. It would, therefore, be the case that $T(p)$ is also indeterminate (because of the Tarskian T-schema). However, through the **OMNI-CORRECTNESS** principle, we obtain that $B(p)$ is indeterminate, and, therefore, the point raised against **OMNI-ACCURACY** is reiterated.

5 F_1 and P_1 are metric temporal operators. Their semantics is straightforward: fixing the day as the unit of time, P_1p means that yesterday, it was true that p ; analogously, F_1p means that tomorrow, it will be true that p .

2445 acceptable. As mentioned earlier, we can assume that the correctness of
 2446 beliefs changes over time because correctness concerns the relationships
 2447 between beliefs and states of affairs in the world. Since changing these
 2448 relationships does not imply any intrinsic change in beliefs, they do
 2449 not appear to have any impact on the fixity of the past. However, the
 2450 change that seems required here does not concern the correctness of
 2451 beliefs. Rather, what is required here is an intrinsic change of the past:
 2452 while at time t_1 it is true that God did not believe at t_0 that it would rain
 2453 (where $t_0 < t_1$), at a later time t_2 it is true that God believed at t_0 that it
 2454 would rain. This intrinsic change in the past is clearly in conflict with
 2455 the fixity of the past.

- 2456 (2) They can deny **OMNI-ACCURACY**. This amounts to denying the very
 2457 *possibility* of an omniscient being. However, it seems strange that a
 2458 future semantic theory could imply the denial of such a possibility.
 2459 As the authors observe, “In general, one could argue that a semantic
 2460 theory—a theory concerned with the logic and compositional structure
 2461 of the language—ought not to settle certain substantive non-semantic
 2462 questions” (Todd and Rabern 2021, 116).

2463 Therefore, it seems that **OC** must be denied. If we accept **OF**, **RCP** must
 2464 be denied. Todd and Rabern discuss another possible defence by advocates
 2465 of **RCP**. Instead of assuming **OMNI-ACCURACY**, the defender of **RCP** could
 2466 assume Omni-correctness:

2467 **OMNI-CORRECTNESS.** $Tp \leftrightarrow B(p)$

2468 where T is the truth predicate; in other words, an omniscient being believes
 2469 that p iff p is true. They could then insist that yesterday, it was not true that
 2470 F_1p . In other words, they could argue that today, $p \rightarrow P_1F_1p$ holds, but
 2471 $p \rightarrow P_1TF_1p$ does not hold. Since it was not true that it would rain yesterday,
 2472 the principle of **OMNI-CORRECTNESS** is not violated. This seems to reconcile
 2473 **OC** with the possibility of an omniscient being.⁶

2474 Todd and Rabern do not find this solution convincing. In fact, they find it
 2475 strange that it could be said that yesterday it was the case that it would rain
 2476 today but that yesterday it was not true that it would rain today. In their view,

6 Notice that this notion of truth is not the one defended by Thomason (1970), who proposed a completely transparent treatment of T : $t \models Tp$ iff $t \models p$.

2477 Very plausibly, if one is moved by the backward-looking intuition
 2478 that, given that a sea-battle has occurred, it was always going to
 2479 occur, it seems that one should likewise be moved by the intuition
 2480 that given that a sea-battle has occurred, it was always true—
 2481 which is not to say determined!—that it was going to occur. (Todd
 2482 and Rabern 2021, 114)⁷

2483 Todd and Rabern also reject MacFarlane’s relativism. According to MacFarlane,
 2484 it is necessary to evaluate propositions on the basis of both the context
 2485 of utterance and the context of assessment: a statement such as “It will rain
 2486 tomorrow” uttered on December 9th has December 9th as its context of utterance.
 2487 However, it can have different contexts for assessment. When evaluated
 2488 with respect to December 9th, it is neither true nor false, but when evaluated
 2489 with respect to December 10th, it is true or false. Todd and Rabern state that
 2490 apart from the technicalities with which this relativistic intuition is implemented,
 2491 it remains true that “insofar as the Open-closurist view has a notion
 2492 of truth that vindicates the (updated) Retro-closure principle, they will have
 2493 to accept the conclusion that *God was genuinely ignorant*. Something was *true*
 2494 (in the relevant sense) that God didn’t believe” (2021, 115). From the point of
 2495 view of December 10th, it was true on December 9th that it would rain the
 2496 following day, and since God did not believe on December 9th that it would
 2497 rain on December 10th, there was something true that God did not believe.

2498 We believe that the two-dimensional semantics we propose in this paper—
 2499 inspired by MacFarlane’s intuitions—has all the resources to demonstrate
 2500 that from OC, it does not follow that God failed to know something true. It is
 2501 possible to assume OC and still not deny the possibility of an omniscient being.
 2502 We will show this in section 4. In the next section, we introduce the branching
 2503 time semantics and illustrate the difficulties of traditional supervaluationism
 2504 in responding to Todd and Rabern’s objection.

2505 3 Branching Time, Supervaluationism, and RCP

2506 As we have seen above, Todd and Rabern are convinced that there is no way
 2507 out for the OC adherent who accepts the possibility of the existence of an
 2508 omniscient entity. Indeed, as we shall see, supervaluationists (*à la* Thomason)

7 Todd and Rabern do not explicitly mention the principle that would be abandoned if one embarked on this strategy. This would amount to denying the Tarskian T-scheme for which φ iff $T\varphi$. In our opinion, this move has unsustainable theoretical costs.

2509 can formally accept the two conditions without contradiction. However, the
 2510 theoretical cost that they have to pay is high; for this reason, we will develop
 2511 an alternative semantics to supervaluationism that is able to account for our
 2512 intuitions about omniscience and that validates **OC**. First, we present the key
 2513 ingredients of a branching time semantics; then, we reconstruct the argument
 2514 from a supervaluationist perspective and show that, although not inconsistent,
 2515 the supervaluationist is nevertheless forced to accept very strange conclusions.

3.1 Branching Time

2517 A branching time structure⁸ is a couple consisting of a non-empty set of time
 2518 instants and an order relation defined on them: $\mathcal{B} = \langle \mathbb{T}, < \rangle$. Intuitively, the
 2519 instants are possible instantaneous states of the world, and $<$ is the relation
 2520 of temporal precedence. This relation is, therefore, asymmetric and transitive
 2521 and satisfies (at least) the conditions of Backward Linearity (**BL**) and Historical
 2522 Connectedness (**HC**).

$$2523 \quad \text{BL. } \forall t \forall t_1 \forall t_2 ((t_1 < t \wedge t_2 < t) \rightarrow (t_1 = t_2 \vee t_1 < t_2 \vee t_2 < t_1))$$

2524 In words, two instants of the past of t are either identical or ordered by $<$; this
 2525 implies that for every instant t , there is one and only one past history.

$$2526 \quad \text{HC. } \forall t_1 \forall t_2 \exists t (t \leq t_1 \wedge t \leq t_2)$$

2527 **HC** asserts that all the instants are connected in the past.

2528 Maximal subsets of instants linearly ordered in t are referred to as histories
 2529 (h)—the possible courses of events around the world. Ours is a propositional
 2530 language that includes a possible infinite set of propositional variables (**Var**)
 2531 and two temporal operators, **P** and **F**. It is useful, as we will see in a moment,
 2532 to exploit metric temporal operators, such as \mathbf{P}_n and \mathbf{F}_n . As indicated earlier,
 2533 \mathbf{P}_n means “ n units of time before the instant of evaluation,” and \mathbf{F}_n means “ n
 2534 units of time after the instant of evaluation.”

2535 Now, let us see how to define formula evaluations in our semantics. Here,
 2536 we will use an Ockhamist framework in which a formula φ is evaluated with
 2537 respect to a time t and a history h . For the formal feature of the order relation
 2538 among instants, any instant in the structure has only one past history but

8 For a classical presentation of branching time, see Belnap, Perloff and Xu (2001). Readers familiar with this literature can go directly to section 3.2.

2539 one or more future histories, depending on whether there is branching in the
 2540 future of that instant.

2541 The use of histories becomes crucial in the case of the evaluation of state-
 2542 ments in the future. Suppose there are two histories branching off from instant
 2543 t_0 , namely, h_1 and h_2 . In h_1 , certain things happen, and therefore certain for-
 2544 mulas are true, while in h_2 , things go differently and, consequently, other
 2545 formulas are true. How can we interpret the proposition “It will rain” ($\mathbf{F}\varphi$)
 2546 evaluated at t_0 if in one history it rains the following day and in the other
 2547 history it does not? A very natural solution might be to relativise truth condi-
 2548 tions to histories, as in Ockhamist semantics. Therefore, today, it is true that
 2549 in the future, it will rain in history, say, h_1 , while it is false that in the future,
 2550 it will rain in history h_2 :

$$\begin{aligned} \mathcal{M}, t/h_1 \models^{\text{ock}} \mathbf{F}\varphi & \quad \text{iff} \quad \exists t' > t, \mathcal{M}, t'/h_1 \models^{\text{ock}} \varphi \\ \mathcal{M}, t/h_1 \not\models^{\text{ock}} \mathbf{F}\varphi & \quad \text{iff} \quad \neg \exists t' > t, \mathcal{M}, t'/h_1 \models^{\text{ock}} \varphi \end{aligned}$$

2551 On the Ockhamist semantics, only if a future branch is specified can a truth
 2552 value be ascribed to a formula. The intuitive problem with this theory consists
 2553 in the fact that there is no trace of such specification of possible branches in
 2554 our everyday talk about the future, which we would like to model (Wawer
 2555 2014, 366).

2556 In the literature, there are two large families of answers to this problem.
 2557 The first, the *Open Futurist* semantics, denies that future contingents can be
 2558 true. Most Open Futurist semantics state that the truth value of a future tense
 2559 statement depends (in a sense to be specified) on what happens in all the
 2560 histories that stem from a certain instant. In the second family of views, on
 2561 the contrary, future contingents can be true. Usually, those who embrace this
 2562 thesis assume *linearist* semantics—that is, they believe that there is somehow
 2563 a privileged history and that the truth conditions of a future tense statement
 2564 concern only what happens in that history.

2565 As is widely known, Open Futurists can be roughly divided into *Peirceans*
 2566 and *Aristotelians*: for the former, greatly simplifying, future contingents are
 2567 all false since a future tense statement is true if and only if it is true in all
 2568 future histories, and by definition, a future contingent is true in some future
 2569 histories and false in others. Aristotelians, on the other hand, maintain that
 2570 future contingents are neither true nor false.

2571 It is not the purpose of this paper to precisely characterise the options on
 2572 the table nor to argue in favour of one solution or another (for this, we refer,
 2573 among many others, to Thomason 1970; Todd 2021). Our aim here is only
 2574 to demonstrate that Open Futurism and RCP are not in conflict with the
 2575 possibility of an omniscient being.

3.2 Supervenience and RCP

2577 Now, let us reconstruct Todd and Rabern’s argument from a supervenience
 2578 standpoint.

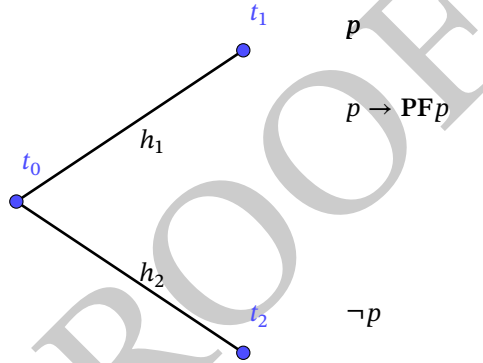


Figure 1: caption

2579 Supervenience distinguishes truth from super-truth. Formulas are eval-
 2580 uated with respect to instant/history pairs; they are then supervenient with
 2581 respect to instants only. The idea can be described as follows: Formula φ is
 2582 super-true at t if and only if φ is linearly true in every history that passes
 2583 through t ; formula φ is super-false at t if and only if φ is linearly false in every
 2584 history that passes through t ; finally, φ is indeterminate if and only if it is
 2585 neither super-true nor super-false.

2586 Figure 1 represents a fork made up of two histories (h_1 and h_2). Within
 2587 the first, p is true; within the second, it is not true. Since p is true at t_1 , in
 2588 accordance with RCP, it was true at t_0 that p would be true because, for every
 2589 history that passes through t_1 (that is, h_1), $\mathbf{PF}p$ is true at t_1 . Now, $t_0 \not\models^{\text{sup}} \mathbf{F}p$
 2590 holds because future contingents are never super-true. By OMNI-ACCURACY

2591 and logic, we have that $t_0 \not\models^{\text{sup}} B(\mathbf{F}p)$ —that is, it is not (super>true that God
 2592 believes at t_0 that it will rain.⁹ Therefore, supervaluationists who endorse
 2593 **OMNI-ACCURACY** are forced to accept both of the following semantic state-
 2594 ments:

- 2595 (vi) $t_0 \not\models^{\text{sup}} B(\mathbf{F}p)$
 2596 (vii) $t_0/h_1 \models^{\text{ock}} B(\mathbf{F}p)$

2597 Although (vi) and (vii) are not contradictory and do not violate **OMNI-**
 2598 **ACCURACY**, this is a weird situation for supervaluationists. They must claim
 2599 that at a given instant of time, it is not (super>true that God believes that
 2600 tomorrow it will rain, but it is true with respect to some history that God
 2601 believes that it will rain tomorrow. It is not straightforward how the advocate
 2602 of this view could account for these results in a coherent, general picture.

2603 In addition, there is a more general problem concerning the adequacy
 2604 of supervaluationist semantics in characterising epistemic predicates, such
 2605 as belief. It is well known that supervaluationism has been developed to
 2606 rigorously account for the semantics of vague predicates (cf., for instance,
 2607 *Fine 1975*). Now, it is not clear how this could apply to belief operator B. If
 2608 a subject is *agnostic* about p , then it seems natural to claim not only that it
 2609 is not true that she believes p (and that she believes $\neg p$) but also that it is
 2610 *false* that she believes p . However, this entails that $\neg B(p)$ is true. The export
 2611 of negation from the metatheory to the theory—that is, the passage from the
 2612 metatheoretical general claim $\mathcal{M} \not\models \varphi$ to $\mathcal{M} \models \neg\varphi$ —is invalid within the
 2613 supervaluationist semantics: from the fact that it is not (super>true that there
 2614 will be a sea-battle tomorrow, it does not follow that it is (super>true that
 2615 there will be no sea-battle. However, for the doxastic operator B, things look
 2616 different. Supposing that Emma is agnostic about the presence of beer in the
 2617 fridge, the supervaluationist semantics would force the following reading: it
 2618 is not (super>true that Emma believes that there is any beer in the fridge, but,
 2619 from that, it does not follow that it is false that Emma believes that there is
 2620 any beer in the fridge. Given our pre-theoretical stances about the semantics
 2621 of the verb *believe*, this seems to be rather odd. To solve this problem, perhaps
 2622 supervaluationists might postulate that if φ is indeterminate, then $\neg B(\varphi)$ is
 2623 true.

9 MacFarlane (*forthcoming*) correctly notices that supervaluationism and **OMNI-ACCURACY** imply $t_0 \not\models^{\text{sup}} B(\mathbf{F}p)$ and not $t_0 \models^{\text{sup}} \neg B(\mathbf{F}p)$, as Todd and Rabern seem to suppose. However, we find supervaluationism still problematic for the reasons given in the main text.

2624 There is a more serious problem for supervaluationists: the question arises
 2625 as to the truth value of formula $\mathbf{P}(\mathbf{B}(\mathbf{F}p))$ evaluated at t_1 . Supervaluationists
 2626 have two possibilities: either $t_1 \not\models^{\text{sup}} \mathbf{P}(\mathbf{B}(\mathbf{F}p))$ or $t_1 \models^{\text{sup}} \mathbf{P}(\mathbf{B}(\mathbf{F}p))$. However,
 2627 both seem to be problematic. If the former holds, then a fact of the past
 2628 (i.e., a divine belief) is indeterminate. This is against the intuition according
 2629 to which past facts are metaphysically determined. Therefore, the second
 2630 possibility must hold. However, in this case, Todd and Rabern's objection
 2631 seems cogent: yesterday, it was indeterminate (or false) that God believed
 2632 $\mathbf{F}p$ since $\mathbf{F}p$ is a future contingent, lacking a (super)truth value, but today,
 2633 it is determinate that yesterday God believed $\mathbf{F}p$. Therefore, endorsing this
 2634 view is quite implausible because it implies an intrinsic change in past facts.¹⁰
 2635 Therefore, we agree with Todd and Rabern's claim that supervaluationism
 2636 is in trouble with **OC** and the possibility of the existence of an omniscient
 2637 entity (Todd and Rabern 2021, 110–111). Now, the following question arises:
 2638 Is it possible to develop an **OC** semantics compatible with the existence of an
 2639 omniscient entity?

2640 **4 Retro-Believing and Retro-Truth**

2641 *4.1 Double-Indices Semantics*

2642 In the following, we develop an **OC** double-indices semantics. Our system
 2643 is, in a sense, inspired by intuitions surrounding MacFarlane's relativist se-
 2644 mantics; however, unlike MacFarlane's system, our semantics considers both
 2645 MacFarlane's contexts of evaluation and of assessment as semantic indices.¹¹
 2646 A formula is evaluated at a particular time and with respect to a perspective;
 2647 the perspective indicates the present time within the structure. We call this

10 MacFarlane (*forthcoming*) claims that Todd and Rabern's argument presupposes a substantive metaphysical claim: past and present beliefs are settled. However, it is difficult to see why past and present beliefs should not be fixed as any other past or present fact. MacFarlane appeals to Jackman (1999) to support the idea that past and present beliefs are not settled. However, Jackman believes that a past belief is not settled when it involves indeterminate meanings that are determined over time. Future uses ultimately determine past uses of a word. By Jackman's own admission, these cases might be quite rare. However, this would not be the case with divine past beliefs about future contingents, which should always be determined by what happens in the future, even when dealing with wholly determinate meanings.

11 We use this semantics since we believe that it makes our argument clearer. However, we assume that our argument could be formulated in any semantics or post-semantics that involves two evaluation indices.

2648 framework *perspectival semantics*. From a formal point of view, this means
 2649 that a formula is evaluated with respect to an instant and any history that is
 2650 included between that instant and the particular instant that is the present,
 2651 or the “now.” Therefore, we evaluate the truth value of a formula at a certain
 2652 instant *when* another instant has the property of being now. In other words,
 2653 the second index is the perspective from which we “see” the structure.¹² This
 2654 parameter is essential, as we will see soon, since it cuts off the histories against
 2655 which a formula is evaluated. In fact, the advancement of the world deter-
 2656 mines the future; time flows, and the bundle of possible available histories is
 2657 reduced. Today (t), it is indeterminate whether it will rain or not tomorrow,
 2658 but tomorrow, when day $t + 1$ will be now, the weather will no longer be
 2659 indeterminate.

2660 To clarify, let us take the above example: today, December 10th, it rains.
 2661 Assuming that today’s rain is a contingent feature of the world, is it indetermi-
 2662 nate on December 9th whether it will rain the next day? The answer would be
 2663 that it depends on the *perspective* from which we locate ourselves. If we place
 2664 the now on December 9th or at an instant preceding December 9th, then the
 2665 rain of the following day is an indeterminate event. The world has arrived—so
 2666 to speak—at just a certain point, and the future is open. But if we place the
 2667 now at an instant following December 9th, such as December 10th, the world
 2668 has been determined, and some histories are no longer available; in particular,
 2669 the possibility of a December 10th with no rain has expired. Therefore, *from*
 2670 *the perspective of December 10th*, it is determined on December 9th that it will
 2671 rain the following day.

2672 Why adopt a two-indices semantics? The grounding idea is that the ad-
 2673 vancement of the present prunes certain histories and leaves others open.
 2674 This feature is crucial when we evaluate the truth value of sentences that are
 2675 future-tensed but evaluated at a *past* instant. Since time has passed, some
 2676 histories are no longer available; they were open *before*, but not now because
 2677 things went a certain way. As said before, our framework explicitly favours no
 2678 specific metaphysics of time; consequently, we do not take a stance towards
 2679 the phenomenon of pruning. But let us consider that our linguistic practices
 2680 often refer both to available and no longer available histories. In other terms,
 2681 we evaluate the formula ψ at an instant t from a perspective t' . Sometimes,

12 One might wonder what the intended interpretation of perspectival semantics is and whether it favours an A-theoretic construal. Although we have argued elsewhere that perspectival semantics is an adequate framework for advocates of the A-theory, it is nevertheless wholly compatible with a purely indexical reading of the now.

2682 the instant of evaluation is connected to the perspective, that is, it lies in its
 2683 past or future; sometimes not. In the first case, we have a *factual* situation;
 2684 in the second, a *counterfactual* one. The two-indices semantics seems to be a
 2685 promising conceptual tool in order to characterise these scenarios.¹³

2686 From a formal point of view, our semantics accounts for this possibility by
 2687 using two temporal indices: the first is the evaluation instant, and the second
 2688 is the present. Therefore, let us consider the following expression:

$$2689 \text{ (viii) } \mathcal{M}, t_i, t_j \models^{\text{PRS}} \varphi$$

2690 (viii) should be read as follows: φ is true at t_i when the present is t_j . The former
 2691 index (t_i) refers to the instant at which the formula is evaluated. The latter
 2692 index (t_j) refers to the position of the present within the structure.

2693 The two instants can coincide. This would be a case in which we would
 2694 evaluate, say, φ at t when the world has arrived at t . In any case, the evaluation
 2695 instant (e.g., t) and the present (e.g., t') must be *connected*; in other words,
 2696 one of the following conditions must hold: $t < t'$ or $t \approx t'$ or $t' < t$.¹⁴ The
 2697 truth clause for atomic formulas is as follows:

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \varphi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \models^{\text{ock}} \varphi$$

2698 where \mathcal{H}_{t_i} is the set of histories that pass through the instant t_i ($\mathcal{H}_{t_i} = \{h | t_i \in$
 2699 $h\}$); therefore, $(\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j})$ is the intersection between the two sets of histories.
 2700 Here, we assume that the satisfiability operator is not bivalent in perspective
 2701 semantics. Therefore, we have the following:

$$\mathcal{M}, t_i, t_j \not\models^{\text{PRS}} \varphi \quad \Leftrightarrow \quad \neg \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \models^{\text{ock}} \varphi$$

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \varphi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \not\models^{\text{ock}} \varphi$$

13 For a more extensive defence and articulated exposition of this two-indices framework, cf. De Florio and Frigerio (2020).

14 A perspectival semantics in which the instant of evaluation and the now need not be connected has been exploited to provide a counterfactual semantics in De Florio and Frigerio (2020). We thank you, an anonymous referee, for this point.

2702 The following truth clauses are straightforward:

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \neg\varphi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \not\models^{\text{ock}} \varphi$$

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \varphi \wedge \psi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \models^{\text{ock}} \varphi$$

$$\text{and } \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), t_i/h \models^{\text{ock}} \psi$$

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \mathbf{P}\varphi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), \exists t' < t_i, t'/h \models^{\text{ock}} \varphi$$

2703 Clauses without temporal operators are a natural extension of the Ockhamist
 2704 linearist evaluation. As for the past case, the second index is vacuous. In
 2705 the past of the instant of evaluation, there is just one history, and, therefore,
 2706 the evaluation is linear. Things become more interesting in the future case
 2707 because the second index plays a significant role.

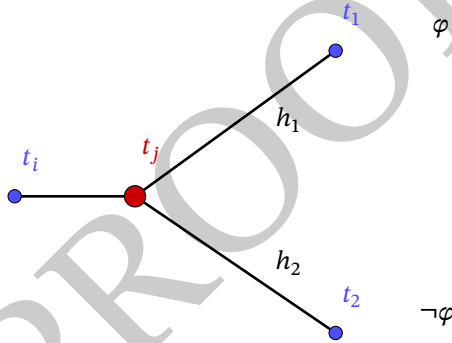


Figure 2: caption

2708 In figure 2, the instant of evaluation is t_i , while the now is at t_j ; the truth
 2709 conditions of $\mathbf{F}\varphi$ at t_i are the following:

$$\mathcal{M}, t_i, t_j \models^{\text{PRS}} \mathbf{F}\varphi \quad \Leftrightarrow \quad \forall h \in (\mathcal{H}_{t_i} \cap \mathcal{H}_{t_j}), \exists t' > t_i, t'/h \models^{\text{ock}} \varphi$$

2710 We now have two cases (see the figures below): the case in which the present
 2711 coincides (or precedes) the instant of evaluation, and the case in which the
 2712 present follows that instant:

2713 (I Case). In this case, the now (t_0) coincides with the point of evaluation.
 2714 Since $\mathcal{H}_{t_0} \cap \mathcal{H}_{t_0} = \mathcal{H}_{t_0}$, both branches (the φ -branch and the $\neg\varphi$ -branch) are
 2715 available. Therefore, $\mathcal{M}, t_0, t_0 \not\models^{\text{PRS}} \mathbf{F}\varphi$.

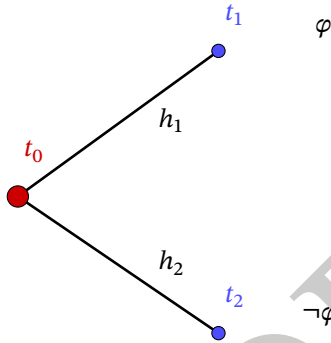


Figure 3: I Case

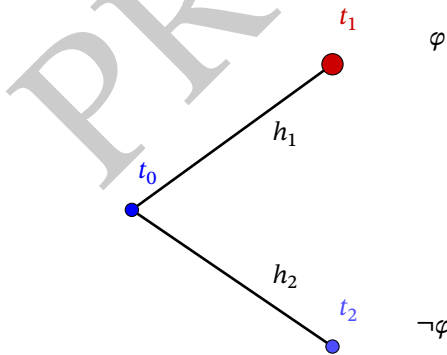


Figure 4: II Case

2716 (II Case). Here, the now is at t_1 and follows the instant of evaluation (t_0). The
 2717 history in which $\neg\phi$ is true is no longer available because $h_2 \notin \mathcal{H}_{t_0} \cap \mathcal{H}_{t_1}$.
 2718 Consequently, $\mathcal{M}, t_0, t_1 \models^{\text{PRS}} \mathbf{F}\phi$.

2719 In the double-indices framework, the truth values of propositions change
 2720 with the flow of time. This accounts for **RCP** and the intuitions reviewed in
 2721 section 1.¹⁵

4.2 The Possibility of Omniscience

2723 Is our perspectival semantics able to answer Todd and Rabern's challenge?
 2724 We think so, and in order to show this, we must elaborate on a further im-
 2725 portant point. The belief operator is usually characterised through Kripkean
 2726 semantics using doxastically possible worlds—namely, possible cognitive
 2727 representations of the subject. It is debatable whether this approach is ade-
 2728 quate to model divine beliefs; in fact, God has just one representation of the
 2729 world—the correct one. This fact is correctly grasped by Todd and Rabern's
 2730 **OMNI-ACCURACY** principle. However, there is another interesting point here.
 2731 The concept of belief is clearly representational: a subject believes that p in
 2732 that she has a representation of the (actual) world as a world in which p is
 2733 true. This holds true both for human subjects and for omniscient entities.
 2734 Believing that tomorrow p will be true has to do with the representation of
 2735 the future actual world.

2736 These reflections may seem to be trivial, but they are of a certain interest
 2737 if we take into account the double-indices analysis we provided. Since the
 2738 concept of belief is representational, it is natural to *centre* it on the now, where
 2739 we represent the present structure of the world. This does not mean that the
 2740 belief cannot concern the future or the past; we can believe that it will rain
 2741 tomorrow or that it rained yesterday, but we do so from the actual (present)
 2742 perspective, whereby if Emma believes (now) that it will rain (tomorrow),
 2743 this means that it is from the perspective of today that Emma believes that it
 2744 will rain tomorrow. Analogously, if Emma believed yesterday that it would
 2745 rain today, it is from the perspective of yesterday that Emma represented the
 2746 following day.

15 Our two-dimensional semantics has some similarity with the two-dimensional possibility frame-
 work proposed by Cariani (2024), even though in a rather different context. In his framework, too,
 the evaluations of formulas depend on two different possibilities, which correspond to moments
 in branching time semantics.

2747 This peculiar feature could be formally presented through a particular
 2748 semantic clause about the belief operator B; in a nutshell, when one construes,
 2749 at instant t and from the perspective t' , a formula in which the B operator
 2750 occurs, then one has to reformulate that interpretation, *moving* the now to
 2751 the instant of evaluation of the formula. The justification for this manoeuvre
 2752 follows from the fact that the belief operator must be fixed to the now. In
 2753 other words, we represent the world based on the information available at the
 2754 moment of the representation.

2755 This means that we cannot rightly construe a past doxastic state if we locate
 2756 the epistemic subject at a time different from that at which she represents
 2757 the world. When we look back and wonder what we believed in the past,
 2758 we must *backdate* the now, bringing back the perspective to the one that is
 2759 contemporaneous with the instant of evaluation.

2760 Formally, all of this can be characterised through the following *belief se-*
 2761 *mantic norm*:

$$2762 \text{BSN. } \mathcal{M}, t_i, t_j \models^{\text{prs}} B(\varphi) \Rightarrow \mathcal{M}, t_i, t_i \models^{\text{prs}} B(\varphi)$$

2763 where φ can have any logical complexity. A formula that describes a belief
 2764 attitude of any complexity must be evaluated from a perspective that coincides
 2765 with the instant evaluation of the formula. Therefore, (BSN) moves the index
 2766 of the perspective from t_j to t_i .¹⁶

2767 This point is important, and it deserves some clarification.¹⁷ (BSN) is a
 2768 principle that is not derived from other semantical axioms. On the contrary,
 2769 it is assumed as a formal counterpart of a conceptual reflection about the
 2770 very notion of belief. Let us consider an ascription of a doxastic state, such
 2771 as $B(x, \psi)$, whose intended meaning is: the doxastic agent x believes that ψ
 2772 is true. Well, ψ could have any complexity; in particular, it can be a future
 2773 or past tense sentence. So far, so good. But let us also consider x , that is, the
 2774 doxastic agent. In a temporal framework, x is located (so to speak) somewhere;
 2775 she believes something at a given time. Then, the perspective from which
 2776 to evaluate $B(x, \psi)$ must be centred on the actual temporal position of the
 2777 doxastic agent.

16 In the following, we also exploit the (BSN) principle for cases of dis-beliefs; this is reasonable, however, since they are representational attitudes toward the world. The idea is that if a proposition is untrue (perhaps because its truth value is indeterminate), then an omniscient entity does not believe it is true.

17 We want to thank an anonymous referee for having prompted these considerations.

2778 As a consequence, the **OMNI-ACCURACY** principle $\varphi \leftrightarrow B(\varphi)$ becomes the
 2779 following:

2780 **OMN-PRS.** For every $t, \mathcal{M}, t, t \models^{\text{PRS}} \varphi \leftrightarrow B(\varphi)$

2781 For every perspective, God believes what is true at the instant of that perspective.
 2782 For example, if $\text{now}(t_0)$, God believes every proposition that is true at t_0
 2783 from the perspective of t_0 . Notice that since φ can have any logical complexity,
 2784 it can contain any number of temporal operators. Therefore, at t_0 from the
 2785 perspective of t_0 , God believes what is true at the other points of the structure
 2786 from the perspective of t_0 . For instance, if φ is true at t_{-1} when $\text{now}(t_0)$, then
 2787 God believes at t_0 from the perspective of t_0 that $\mathbf{P}_1\varphi$. In symbols:

$$\mathcal{M}, t_0, t_0 \models^{\text{PRS}} \mathbf{P}_1\varphi \leftrightarrow B(\mathbf{P}_1\varphi)$$

2788 Therefore, God believes at any instant what is true at any point in the structure
 2789 from the perspective of that instant.¹⁸

2790 The fundamental ingredients of our semantics can be summarised as follows:
 2791 propositions are true at an instant of evaluation and with respect to a
 2792 perspective. This entails that, in some cases, the truth value of a proposition,
 2793 evaluated at instant t_i , is indeterminate from a certain perspective, while from
 2794 another perspective, it is true. Divine beliefs always track the truth from the
 2795 perspective of the evaluation because they are representations of the world at
 2796 a certain instant from the perspective of that instant.

2797 Let us now see how our perspectival semantics, enhanced with (**BSN**), is
 2798 able to answer Todd and Rabern’s objection. First, let us resume, in a semi-
 2799 formal fashion, Todd and Rabern’s argument. For convenience, we use metric
 2800 temporal operators. Consider figure 5.¹⁹ Since $t_1 \models \varphi$ and since, by hypothesis,

18 As previously noted, the **OMNI-ACCURACY** principle is silent about cases in which the truth value of the formula at play is indeterminate. We think it is plausible and in agreement with our intuitions about an omniscient entity (but, in general, this should be valid for any epistemic subject) that the following condition must be satisfied: if a proposition is indeterminate with respect to its truth value, then the omniscient entity does not believe it (as true). We translate this indeterminacy situation via the following meta-theoretical conditional:

$$\text{IND-BEL. If } \mathcal{M}, t_0, t_0 \not\models^{\text{PRS}} \varphi, \text{ then } \mathcal{M}, t_0, t_0 \models^{\text{PRS}} \neg B\varphi$$

(**OMN-PRS**) and **IND-BEL** describe, therefore, the relationships between propositions and their truth values from one side and God’s beliefs on the other side.

19 Todd and Rabern never specify the semantic system in which they carry out their arguments. For this reason, we assume that their satisfaction relation (\models) involves a quantification on times.

2801 RCP holds, we have that $t_1 \models \mathbf{P}_1 \mathbf{F}_1 \varphi$. However, $\mathbf{F}_1 \varphi$ is not true at t_0 because
 2802 it is future contingent. It follows that $t_0 \models \neg \mathbf{B}(\mathbf{F}_1 \varphi)$ (cf. footnote 12) because
 2803 God does not believe what is untrue, but since it is true at t_1 that at t_0 φ
 2804 would be true the following day, God should have believed that. Instead,
 2805 $t_1 \models \mathbf{P}_1(\mathbf{F}_1 \varphi \wedge \neg \mathbf{B}(\mathbf{F}_1 \varphi))$. In other words, let us locate ourselves at t_1 (when
 2806 φ is true). In this case, two things were true yesterday: on the one hand, φ
 2807 would be true the next day, and on the other hand, God did not believe that φ
 2808 would be true the next day. However, this means that God is ignorant about
 2809 the future truth of φ .

2810 It is easy to realise that, within our framework, we get two theoretically
 2811 interesting results. First, the framework allows us to distinguish the case of
 2812 the “genuine” future from the case of the retrospective future (i.e., the future
 2813 in the past). At the same time, Todd and Rabern’s argument is no longer
 2814 reproducible. Let us see why.

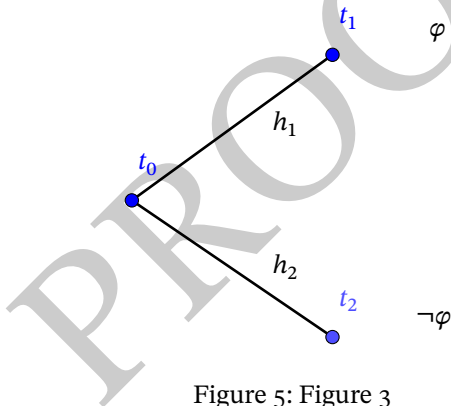


Figure 5: Figure 3

2815 From $\mathcal{M}, t_1, t_1 \models^{\text{PRS}} \mathbf{P}_1(\mathbf{F}_1 \varphi \wedge \neg \mathbf{B}(\mathbf{F}_1 \varphi))$, it follows that $\mathcal{M}, t_1, t_1 \models^{\text{PRS}} \mathbf{P}_1 \mathbf{F}_1 \varphi \wedge$
 2816 $\mathbf{P}_1 \neg \mathbf{B}(\mathbf{F}_1 \varphi)$ and then that $\mathcal{M}, t_1, t_1 \models^{\text{PRS}} \mathbf{P}_1 \mathbf{F}_1 \varphi$. We also have that $\mathcal{M}, t_1, t_1 \models^{\text{PRS}}$
 2817 $\mathbf{P}_1 \neg \mathbf{B}(\mathbf{F}_1 \varphi)$, whose meaning is: from today’s perspective, it was true that
 2818 yesterday, God did not believe that φ would be true. From this, it follows that
 2819 $\mathcal{M}, t_0, t_1 \models^{\text{PRS}} \neg \mathbf{B}(\mathbf{F}_1 \varphi)$ —that is, again, from today’s perspective, yesterday,
 2820 God did not believe that φ would be true. Is this sufficient to claim that God
 2821 was ignorant? No. God appears to be ignorant only because we see the world
 2822 from a metaphysically privileged standpoint, and we look at the world from
 2823 the perspective of how things actually happened.

2824 If we really want to locate ourselves at t_0 , we must also backdate the now;
 2825 we must, in other terms, rewind the tape of history. This is precisely what
 2826 (BSN) does: $\mathcal{M}, t_0, t_1 \models^{\text{PRS}} \neg B(\mathbf{F}_1\varphi)$ becomes $\mathcal{M}, t_0, t_0 \models^{\text{PRS}} \neg B(\mathbf{F}_1\varphi)$. Thus,
 2827 when the present was yesterday, God did not believe that φ would happen.
 2828 However, $\mathcal{M}, t_0, t_0 \not\models^{\text{PRS}} \mathbf{F}_1\varphi$ also holds because $\mathbf{F}_1\varphi$ is a future contingent.
 2829 Therefore, assuming the present version of OF, φ is neither true nor false at
 2830 t_0 . However, the fact that an omniscient entity does not believe what is not
 2831 (yet) true is not a problem for that entity's omniscience.²⁰

2832 Todd and Rabern's argument hinges on a theoretical passage that retro-
 2833 grades the truth and charges an omniscient being with the alleged failure
 2834 to grasp that truth. However, it is precisely a retrograded truth: it is a truth
 2835 only because the world has moved forward, and what was indeterminate is
 2836 now determinate. If we place ourselves at the temporal perspective of the
 2837 omniscient entity at t_0 (i.e., if we move the now to t_0), we naturally get that
 2838 the omniscient entity does not believe that φ would be the case since, from
 2839 t_0 's perspective, it is not true that φ will be the case.

2840 At the same time, from t_1 's perspective, it was true that φ would be true:
 2841 RCP entails $\mathcal{M}, t_0, t_1 \models^{\text{PRS}} \mathbf{F}_1\varphi$. However, this is not a problem for the possibil-
 2842 ity of omniscience because Todd and Rabern suppose that the belief of the
 2843 omniscient entity occurs when the present is t_0 , not when it is t_1 .

2844 In addition, we have that $\mathcal{M}, t_1, t_1 \models^{\text{PRS}} \mathbf{P}_1\mathbf{F}_1\varphi \wedge B(\mathbf{P}_1\mathbf{F}_1\varphi)$, from which it
 2845 follows that $\mathcal{M}, t_1, t_1 \models^{\text{PRS}} B(\mathbf{P}_1\mathbf{F}_1\varphi)$. In other words, from today's perspective,
 2846 God believes that yesterday, it was true that it would rain today. Actually, as
 2847 we have seen, it is true from today's perspective that yesterday, it was true that
 2848 it would rain today.

2849 To sum up, when now(t_0), God does not believe that $\mathbf{F}_1\varphi$ because she sees
 2850 the world from t_0 's perspective, from which $\mathbf{F}_1\varphi$ is untrue. When now(t_1),
 2851 God believes that $\mathbf{F}_1\varphi$ was true the previous day because she sees the world
 2852 from t_1 's perspective, from which the previous day it was true that φ would
 2853 be true the next day. Therefore, God's beliefs always track the truth. Time
 2854 flows, and with its advancing, the truth values of propositions change; an
 2855 omniscient God always believes at a time t what is true from the perspective

20 An anonymous referee suggests that Todd and Rabern's argument shows the incompatibility between the Retro-closure principle and the idea of *permanent omniscience*. The idea is as follows: in our framework, we have cases in which an omniscient being realises—so to speak—that in the past, there were true propositions not known. However, this does not happen in our framework since the omniscient being realises that in the past, He did not believe φ , but, nevertheless, φ was indeterminate from yesterday's perspective. Only from today's perspective, φ was true.

2856 of that time. Put differently, God's beliefs are changing in the same way that
 2857 the truth values of propositions change due to the flowing of time. However,
 2858 this is what is reasonable to demand of an omniscient entity.

2859 Obviously, our argument can be opposed by objecting (BSN). However, this
 2860 does not seem to be a plausible objection if we assume that God is temporal,
 2861 as Todd and Rabern do. Such a God is temporally located, and He, therefore,
 2862 knows the world from the perspective of the present. If we embrace the Open
 2863 theism view, His beliefs evolve when tracking the evolution of the world. If
 2864 this concept of God is assumed, it is quite natural to endorse (BSN): God is an
 2865 entity located within the present who, therefore, knows the world from the
 2866 present's perspective, like human beings.

2867 Of course, we could have different theistic views. We could hypothesise that
 2868 God is timeless and that He knows the evolution of the world from an eternal
 2869 perspective. Within these views, (BSN) can be safely discharged. However, if
 2870 these views are presupposed, Todd and Rabern's argument is no longer valid:
 2871 being timeless, God neither remembers nor anticipates anything. Rather, He
 2872 sees the whole unfolding of the world from His eternal standpoint, and thus,
 2873 He knows all that happens at every instant. If God is timeless, it is meaningless
 2874 to wonder whether $t_0 \models^{Prs} P(B(Fp))$ holds since the belief operator B cannot
 2875 be within the scope of the temporal operators and, in general, cannot be
 2876 evaluated with respect to an instant.²¹

2877 Therefore, either one assumes that God is temporal, and then Todd and
 2878 Rabern's argument does not succeed since it is reasonable to claim that (BSN)
 2879 holds, or one assumes that God is timeless, and then (BSN) does not hold.
 2880 In the latter case, however, Todd and Rabern's argument cannot even be
 2881 formulated since their argument presupposes a God located within time.

2885 5 Conclusions

2883 In our semantic framework, there is no instant at which a formula is true, and
 2884 the omniscient entity does not believe it to be true. The theoretical cost we

21 It is reasonable to require that a timeless God knows the truth values of propositions relatively at every instant of time and every temporal perspective, that is, from any "now." For this omniperspectival view of God's knowledge, see De Florio and Frigerio (2019, chap. 6). This view seems to be naturally connected with a B-theoretic metaphysics of time, where all the "presents" exist on a par. If one wants to keep together a timeless God and an A-theoretic metaphysics, one needs to appeal to non-standard A-theories, like Fragmentalism. On this, again, see De Florio and Frigerio (2019, chap. 6).

2885 have to pay is the acceptability of (BSN); specifically, the principle according
 2886 to which the ascription of belief to a subject at an instant is constrained by the
 2887 state of the universe at that time and cannot be legitimately forward dated.
 2888 We think this is a highly affordable cost based on a reasonable theoretical
 2889 proposal. We conclude that Todd and Rabern’s argument fails to show that
 2890 OF is incompatible with the possibility of an omniscient entity. OF, therefore,
 2891 remains a viable alternative in the tense semantics market.*

2892
2893
2894
2895
2896
2897
2898
2899

Ciro De Florio

 0000-0002-5503-7463

Università Cattolica of Milan, Italy

ciro.deflorio@unicatt.it

Aldo Frigerio

 0000-0003-4698-2110

Università Cattolica of Milan, Italy

aldo.frigerio@unicatt.it

2900 References

- 2901 BELNAP, Nuel D., Jr. and GREEN, Mitchell S. 1994. “Indeterminism and the Thin
 2902 Red Line.” in *Philosophical Perspectives 8: Logic and Language*, edited by James
 2903 E. TOMBERLIN, pp. 365–388. Atascadero, California: Ridgeview Publishing Co.,
 2904 doi:10.2307/2214178.
- 2905 BELNAP, Nuel D., Jr., PERLOFF, Michael N. and XU, Ming. 2001. *Facing the Future:
 2906 Agents and Choices in Our Indeterminist World*. Oxford: Oxford University Press,
 2907 doi:10.1093/oso/9780195138788.001.0001.
- 2908 BONOMI, Andrea. 1997. “The Progressive and the Structure of Events.” *Journal of
 2909 Semantics* 14(2): 173–205, doi:10.1093/jos/14.2.173.
- 2910 CARIANI, Fabrizio. 2024. “Modeling Future Indeterminacy in Possibility Semantics.”
 2911 Unpublished manuscript, <https://philpapers.org/archive/CARMFI.pdf>.
- 2912 DE FLORIO, Ciro and FRIGERIO, Aldo. 2018. “Two Kinds of Soft Facts.” *Grazer
 2913 Philosophische Studien* 95(1): 34–53, doi:10.1163/18756735-000031.
- 2914 —. 2019. *Divine Omniscience and Human Free Will: A Logical and Metaphysical Anal-
 2915 ysis*. Palgrave Frontiers in Philosophy of Religion. London: Palgrave Macmillan,
 2916 doi:10.1007/978-3-030-31300-5.

* The authors gratefully thank two anonymous referees from *Dialectica* for their constructive comments and recommendations, which definitely helped to improve a previous version of this paper.

- 2917 —. 2020. “The Thin Red Line, Molinism, and the Flow of Time.” *Journal of Logic,*
2918 *Language, and Information* 29(3): 307–329, doi:10.1007/s10849-019-09304-4.
- 2919 FINE, Kit. 1975. “Vagueness, Truth, and Logic.” *Synthese* 30(3/4): 265–300, doi:10.100
2920 7/bf00485047.
- 2921 HOFFMAN, Joshua and ROSENKRANTZ, Gary S. 1984. “Hard and Soft Facts.” *The*
2922 *Philosophical Review* 93(3): 419–434, doi:10.2307/2184544.
- 2923 IACONA, Andrea. 2014. “Ockhamism Without Thin Red Lines.” *Synthese* 191(12):
2924 2633–2652, doi:10.1007/s11229-014-0405-6.
- 2925 JACKMAN, Henry. 1999. “We Live Forwards but Understand Backwards: Linguistic
2926 Practices and Future Behavior.” *Pacific Philosophical Quarterly* 80(2): 157–177,
2927 doi:10.1111/1468-0114.00078.
- 2928 LANDMAN, Fred. 1992. “The Progressive.” *Natural Language Semantics* 1(1): 1–32,
2929 doi:10.1007/BF02342615.
- 2930 MACFARLANE, John. 2003. “Future Contingents and Relative Truth.” *The Philosophical*
2931 *Quarterly* 53(212): 321–336, doi:10.1111/1467-9213.00315.
- 2932 —. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford
2933 University Press, doi:10.1093/acprof:oso/9780199682751.001.0001.
- 2934 —. forthcoming. “Why Future Contingents Are Not All False.” *Analytic Philosophy*,
2935 doi:10.1111/phib.12334.
- 2936 MCKIM, Vaughan R. and DAVIS, Charles C. 1976. “Temporal Modalities and the
2937 Future.” *Notre Dame Journal of Formal Logic* 17(2): 233–238, doi:10.1305/ndjfl/10
2938 93887526.
- 2939 ØHRSTRØM, Peter. 2009. “In Defence of the Thin Red Line: A Case for Ockhamism.”
2940 *Humana.Mente* 3(8): 17–32, [https://www.humanamente.eu/index.php/HM/issu](https://www.humanamente.eu/index.php/HM/issue/view/8)
2941 [e/view/8](https://www.humanamente.eu/index.php/HM/issue/view/8).
- 2942 PORTNER, Paul H. 1998. “The Progressive in Modal Semantics.” *Language* 74(4):
2943 760–787, doi:10.2307/417002.
- 2944 THOMASON, Richmond H. 1970. “Indeterminist Time and Truth-Value Gaps.” *Theoria*
2945 36(3): 246–281, doi:10.1111/j.1755-2567.1970.tb00427.x.
- 2946 TODD, Patrick. 2013. “Soft Facts and Ontological Dependence.” *Philosophical Studies*
2947 164(3): 829–844, doi:10.1007/s11098-012-9917-4.
- 2948 —. 2021. *The Open Future: Why Future Contingents Are All False*. Oxford: Oxford
2949 University Press, doi:10.1093/oso/9780192897916.001.0001.
- 2950 TODD, Patrick and RABERN, Brian. 2021. “Future Contingents and the Logic of Tem-
2951 poral Omniscience.” *Noûs* 55(1): 102–127, doi:10.1111/nous.12294.
- 2952 WAWER, Jacek. 2014. “The Truth about the Future.” *Erkenntnis* 79(suppl., 3): 365–401,
2953 doi:10.1007/s10670-013-9454-3.
- 2954 WAWER, Jacek and MALPASS, Alex. 2020. “Back to the Actual Future.” *Synthese* 197(5):
2955 2193–2213, doi:10.1007/s11229-018-1802-z.

PROOF

Review of Schofield (2021)

AGNÈS BAEHNI

2956

2957 SCHOFIELD, Paul. 2021. *Duty to Self: Moral, Political, and Legal Self-Relation*.
2958 Oxford: Oxford University Press.

2959 Paul Schofield's *Duty to Self* (2021) is an excellent contribution to recent moral
2960 philosophy. It is a much-needed addition to a literature that has, up until now,
2961 largely ignored the possibility of reflexive moral relationships. Thorough
2962 and challenging, the book is an indispensable read for students and scholars
2963 with an interest in ethics, metaethics, and political philosophy. In this review,
2964 I outline what I perceive to be the book's main contributions and discuss
2965 some areas of concern about Schofield's innovative framework. It should be
2966 noted from the outset that Schofield does not claim to tell us what we owe to
2967 ourselves but rather seeks to establish the possibility of moral duties to the
2968 self. As he insists: "The project is most concerned with the metaphysics of
2969 morals, rather than the content of morality" (p. 18).

2970 In its traditional 'social' understanding, the term 'moral' is applied to our
2971 relationships with others and is only rarely applied to our relationships with
2972 ourselves. It is often taken for granted that the normative dimension of the self-
2973 relationship, how one should act towards oneself, is rather to be understood in
2974 terms of prudence. Schofield questions this assumption by focusing on moral
2975 duties to oneself. As we shall see, he is the first to offer a compelling way out of
2976 the objection raised by Marcus Singer in the early 1960s, according to which
2977 the notion of moral duties to oneself is incoherent. In a nutshell, Singer's
2978 (1959, 1963) argument is that we cannot be bound by a duty if we are at any
2979 time free to release ourselves from it. We typically think of duties as binding
2980 because the power to release us from them lies with another individual. Since
2981 we are always free to release ourselves from what we owe to ourselves, Singer
2982 argues, there can be no genuine duties to the self. This is the 'waivability
2983 objection'.

2984 The first chapter, "On the Significance of Duties to the Self," seeks to eluci-
2985 date whether duties to the self can be moral obligations. Drawing on Stephen
2986 Darwall's (2006) influential insights, Schofield argues that having a genuine

2987 moral duty to ourselves entails that we adopt a second-personal stance to-
 2988 wards ourselves. So, the main issue is “whether a person can have duties
 2989 second-personally directed to herself, of the kind had by a person standing on
 2990 another’s foot” (p. 26). Singer’s waivability objection is based on the idea that
 2991 we cannot relate morally to ourselves in this way. While using Singer’s objec-
 2992 tion as his target, Schofield sets aside the possibility of conceiving our moral
 2993 relationships to ourselves on a monadic model of morality, like virtue ethics
 2994 (p. 23).¹ This is where his approach is most ambitious: it aims to understand
 2995 our moral relationships to ourselves and others on a unified model.

2996 A first way to reply to the waivability objection is introduced in chapter
 2997 3 (“Defending Duties to the Self Part 1: Duties Across Time”). Schofield’s
 2998 idea, roughly, is that we can understand our relationships to ourselves second-
 2999 personally by making use of the notion of temporal perspective. To illustrate
 3000 this idea, Schofield discusses the case of a smoker. Let us call him Alan. Why
 3001 should we think that Alan owes it to himself to quit smoking? Schofield’s
 3002 answer is that Alan can adopt a second-personal stance towards himself by
 3003 paying attention to the legitimate demands that might be issued from one of
 3004 his temporal perspectives. We can think, for instance, of his perspective
 3005 in 20 years’ time. When so doing, Alan may be confronted with a conflict
 3006 between his present and his future interests, i.e., the interests of his future
 3007 perspective. Alan does not want to quit smoking now since he finds it enjoyable
 3008 and is not suffering (yet) from any side effects. By contrast, judging from his
 3009 future perspective, he should quit smoking, as continuing to smoke will be
 3010 detrimental to his health in the long run. As Schofield explains, “these cross-
 3011 temporal conflicts between various of a person’s interests enable us to tell a
 3012 story about *intrapersonal* generation of duties that parallels in its essentials
 3013 an *interpersonal* story” (p. 67).

3014 The waivability objection does not threaten duties owed to others because
 3015 the power to release us from our obligations lies in another’s hand. Under-
 3016 standing duties to ourselves in terms of obligations towards future perspectives
 3017 allows reconsidering the idea that we can always release ourselves from such
 3018 duties: we cannot when the power to do so lies with another temporal per-
 3019 spective. A question still remains: Does a person knowing that she will die
 3020 soon from an incurable disease have no duty to herself? To analyze duties to
 3021 ourselves as duties over time may not do justice to all these duties.

1 As Schofield observes, in virtue ethics, each virtue “supplies a standard for good action under which all persons are evaluable, without necessarily putting the subjects into normative contact with others” (p. 23).

3022 This limitation is addressed in chapter 4 (“Defending Duties to the Self Part
3023 2: Duties at a Moment”), which seeks to establish the existence of duties to
3024 ourselves *at a time*. When considering duties at a time, Schofield introduces
3025 the notion of a *practical perspective*. The idea that we can occupy different
3026 practical perspectives stems from the observation that we all wear different
3027 hats: we are philosophers, parents, sportsmen and women, chess players, etc.
3028 Sometimes, our duties as philosophers conflict with the duties of another of
3029 our practical perspectives; for example, we wonder whether we should go
3030 to the chess club or grade philosophy papers. As Schofield puts it, “When
3031 an individual addresses another, a person might address *herself* from the
3032 perspective of one of her practical identities, issuing demands that will be
3033 received from a perspective associated with a different practical identity”
3034 (p. 107). Again, the conflict between different perspectives’ interests helps
3035 explain how we can have moral duties to ourselves at a time. We cannot
3036 release ourselves from our duty to go to the chess club because the power to
3037 do so lies with a different practical perspective, that of a chess player.

3038 Schofield’s proposal is promising, but it also raises some concerns having
3039 to do with the notion of a person. Quoting Schofield:

3040 It has been my aim to give an account of duties [...] while main-
3041 taining that the person herself is the locus of moral value. But
3042 one might doubt whether I’ve succeeded in this. Talk of personal
3043 identities and of second-personal interactions between them will
3044 suggest to some readers a picture on which multiple “selves” con-
3045 stitute a person at a moment. (p. 124)

3046 Indeed, the idea that we adopt several practical and temporal perspectives in
3047 our practical deliberations intimates a kind of division that may threaten the
3048 unity presupposed by the idea of a person: “We’re thus left to worry: What if
3049 the cost of gaining a second person within is losing the person altogether?”
3050 (p. 206).

3051 Schofield goes on to answer this worry by insisting that the adoption of
3052 the second-personal standpoint only requires the capacity to occupy different
3053 practical or temporal perspectives. His proposal is not premised on the claims
3054 that one is composed of different selves interacting with one another at a
3055 time or that “person-stages, or time-slices” (p. 169) relate to one another over
3056 time. Perspectives are epistemic stances, not metaphysical entities, and it
3057 is always the person who relates second-personally to herself through the

3058 adoption of different perspectives. Schofield rightly emphasizes here a point
3059 of tension in his proposal, having to do with the notion of perspective. We
3060 might fear that by focusing on the task of telling us what these perspectives
3061 are not, Schofield leaves us wondering about what they exactly are and how
3062 we can access them. So, while his proposal is the most articulate response
3063 to the waivability objection to date, the key notion of a perspective remains
3064 somewhat elusive.

3065 I suggested that some people may not have the kind of distant temporal
3066 perspectives required to ground duties across time. One may also fear that the
3067 appeal to practical perspectives is insufficient to explain why we have duties
3068 to ourselves at a time. Consider, for instance, a genius artist who is fulfilled by
3069 committing herself exclusively to her art or a monk who is perfectly content
3070 with his life of devotion. Arguably, these individuals do not have other practi-
3071 cal perspectives, and there is thus no conflict of interest. Granted that they
3072 exist, can Schofield's proposal acknowledge that people who are completely
3073 coherent, unified, or focused on the present are in a moral relationship with
3074 themselves? At first sight, the answer is "no" since his view presupposes the
3075 capacity to look at ourselves from a distance, or to see ourselves as someone
3076 else, so to speak. The same seems to be true of our moral relationships with
3077 other people: I have a moral duty to be nice to my friends even though I *want*
3078 to be nice to them anyway. We can have moral duties to the self and to others
3079 without conflicting interests.

3080 A second worry is that Schofield's focus on practical and temporal per-
3081 spectives might be blurring the nature of the issue at stake. There might be
3082 alternative ways to gloss the necessary second-personal stance than by refer-
3083 ence to temporal and practical perspectives. For instance, we sometimes
3084 regard ourselves second-personally when considering that what we did was
3085 morally wrong. The duty of self-respect may not be explained in terms of
3086 demands issued by one of our practical or temporal perspectives but only by
3087 reference to a moral perspective.


3088 This brings me to my final point. Given its reliance on the notion of per-
3089 spective, Schofield's proposal may not have the resources to explain two *prima*
3090 *facie* central and interdependent aspects of the reflexive moral relationships:
3091 its relations to the aims of being happy and of becoming the best version
3092 of ourselves. Indeed, to which practical or temporal perspective do we owe
3093 the pursuit of happiness and of ethical or intellectual development? Given
3094 that these goods are desirable from any perspective, reference to perspectives
3095 seems here both unnecessary and insufficient in explaining why we have such

3096 duties to ourselves.² This is where Schofield’s model of moral self-relationship
 3097 meets its limits. As I observed, Schofield is explicit that he does not seek to
 3098 establish precisely what we owe to ourselves. Nevertheless, it seems reasonable
 3099 to expect the metaphysics of the moral self-relationship to be consistent with
 3100 how we intuitively conceive of the content of this relationship.

3101 Chapters 5 and 6 of the book shift from the moral to the political sphere.
 3102 Having shown how we can relate morally to ourselves, Schofield goes on to
 3103 show that we can also relate politically to ourselves. In chapter 5 (“Might There
 3104 Be Self-Directed Political Duties? Troubles for State Paternalism”), he lays
 3105 out compelling reasons for thinking that we have no self-regarding political
 3106 duties before arguing, in chapter 6 (“Defending Political Duties to the Self: The
 3107 Possibility of Liberal Paternalism”), that we in fact have self-directed duties
 3108 of right and justice and that paternalism is among the state’s functions. His
 3109 aim, ultimately, is to challenge “a conception on which politics is for others
 3110 exclusively” (p. 140).

3111 In the seventh and final chapter (“Practical Philosophy After Duties to
 3112 Self”), Schofield draws some important implications of his proposal, among
 3113 which the fact that it should not be taken “as a mere addendum to whatever
 3114 ethical theory is already in place” (p. 195). On the contrary, he insists that
 3115 two difficult and hitherto ignored questions now arise. First, what to do when
 3116 what we morally owe to ourselves conflicts with the moral interests of others?
 3117 Second, what should we do when our own interests conflict with one another?
 3118 These questions have been largely ignored so far, and practical philosophy
 3119 will certainly be busy in the coming years trying to answer them. We should
 3120 be grateful to Schofield for his stimulating proposal, which provides a stable
 3121 foundation for future debates on these issues.

3122 Agnès Baehni

3123  0000-0002-7605-9346

3124 Université de Genève

3125 Agnes.Baehni@unige.ch

3126 References

3127 DARWALL, Stephen L. 2006. *The Second-Person Standpoint: Morality, Respect, and*
 3128 *Accountability*. Cambridge, Massachusetts: Harvard University Press.

2 Yuliya Kanygina makes a similar point in Kanygina (2022).

- 3129 KANYGINA, Yuliya. 2022. "Duties to Oneself and Their Alleged Incoherence." *Aus-*
3130 *tralasian Journal of Philosophy* 100(3): 565–579, doi:[10.1080/00048402.2021.1909](https://doi.org/10.1080/00048402.2021.1909)
3131 [086](https://doi.org/10.1080/00048402.2021.1909086).
- 3132 SCHOFIELD, Paul. 2021. *Duty to Self: Moral, Political, and Legal Self-Relation*. Oxford:
3133 Oxford University Press, doi:[10.1093/oso/9780190941758.001.0001](https://doi.org/10.1093/oso/9780190941758.001.0001).
- 3134 SINGER, Marcus George. 1959. "On Duties to Oneself." *Ethics* 69(3): 202–205, doi:[10.1](https://doi.org/10.1086/291210)
3135 [086/291210](https://doi.org/10.1086/291210).
- 3136 —. 1963. "Duties and Duties to Oneself." *Ethics* 73(2): 133–142, doi:[10.1086/291439](https://doi.org/10.1086/291439).

PROOF

Review of Lepine (2023)

STEVE HUMBERT-DROZ

3137

3138 LEPINE, Samuel. 2023. *La nature des émotions. Une introduction partisane.*
3139 Analyse et philosophie. Paris: Librairie philosophique Jean Vrin.

3140 We are now in the age of affectivism (Dukes et al. 2021): while emotions have
3141 long been contrasted with cognition, they are now seen as a central element
3142 of our rational life. Lepine (2023) joins this paradigm, arguing that emotions
3143 are cognitive states, a source of axiological knowledge, and even an essential
3144 component of values.

3145 Lepine's original contribution consists of an extremely cautious and impres-
3146 sive interweaving of psychological and philosophical discussions of emotions
3147 as well as of values. We may take from *La nature des émotions* a set of en-
3148 tangled statements: (i) emotions are *cognitive states* distinguished by their
3149 *evaluative* nature; (ii) they are sui generis psychological modes that *focus our*
3150 *attention* and prepare our body for action; (iii) they are evaluative since we
3151 can ascribe a *correspondence* between the emotion and the value instanti-
3152 ated by the emotion's intentional object (i.e., correctness conditions); (iv) the
3153 correctness and justification conditions of emotions *partially depend on the*
3154 *background motivations* on which every emotion is based because (v) values
3155 depend on some non-evaluative properties of external objects as well as on
3156 the *agents' motivations*. These different points fit together to form the most
3157 comprehensive introduction to emotions I've read since Deonna and Teroni's
3158 *The Emotions* (2012). Let us examine how.

3159 Chapter (1) outlines the so-called "naive features" of emotions: automaticity,
3160 valence, intentionality, direction of fit, cognitive and motivational bases, etc.
3161 Chapter (2) focuses on the opposition between emotion and cognition. The
3162 notion of cognition is discussed with reference to the debate between Zajonc
3163 (1984) and Lazarus and Folkman (1984). Lepine (2023) endorses the appraisal
3164 theory of emotions in psychology (Lazarus' view), arguing that (i) emotions are
3165 *cognitive states* insofar as stimuli processing in emotions makes them available
3166 for semantic processing (p. 61), the *evaluative nature* of the processing being
3167 the mark of emotions in cognition (pp. 59–60).

3168 Chapters (3) and (4) explore the evaluative nature of emotions. Lepine
 3169 first wonders whether emotions constitute a natural kind. Chapter (3) sets
 3170 out Griffiths's (1997) well-known objections to the idea that emotions share
 3171 essential properties, coupled with considerations coming from constructivist
 3172 theories (e.g., Russell 2003). Lepine concludes, in line with the appraisal
 3173 theory, that "emotions would indeed constitute a natural kind insofar as each
 3174 of them shares the function of detecting a specific core relational theme"
 3175 (p. 87; I translate all quotations).¹

3176 Chapter (4) then explores philosophical theories of emotions to understand
 3177 further the notion of a *core relational theme*. According to Lepine, this notion
 3178 is analogous to what philosophers call "formal objects" (pp. 108–109). A dog, a
 3179 steep cliff, losing my money in the stock market, and my Ph.D. supervisor are
 3180 objects I mention when answering the question, What are you afraid of?—they
 3181 are the *intentional objects* of my fear. What do they have in common (when
 3182 my fear is appropriate)? They all instantiate the same evaluative property,
 3183 being dangerous. Danger is, thus, the *formal object* of fear. How are formal
 3184 objects connected to emotions? Lepine follows Deonna and Teroni's (2012,
 3185 2015, 2024) attitudinal view of emotions, according to which (ii) emotions are
 3186 sui generis *psychological modes* (i.e., they are reducible neither to judging nor
 3187 to perceiving, and so on)² constituted by unified bodily feelings that prepare
 3188 the subject for action—e.g., fear prepares me to flee. Most importantly, (iii) the
 3189 content of emotions need not be evaluative; the evaluative nature of emotions
 3190 lies in the *fittingness relation* between fear, shame, pride..., and the evaluative
 3191 properties instantiated by the intentional object of these emotions. In other
 3192 words, the content of my emotion doesn't need to go beyond a non-evaluative
 3193 representation of the intentional object—"the dog," "the steep cliff," etc.

3194 Lepine illustrates this relationship by interpreting the attitudinal theory
 3195 through Cummins's (1996) analysis of psychological attitudes. Attitudes are
 3196 characterized by their cognitive function;³ the attitude sets a target and pro-
 3197 cesses its content in a way that is correct when the target is reached:

-
- 1 We may regret that Lepine does not raise as vigorous objections to appraisal theory as he does to other ones. For example, it is unclear whether appraisals cause or constitute emotions (Moors 2013; Roseman and Smith 2001).
- 2 Chapter (4) also discusses perceptual, judgmental, mixed views, etc. These discussions are close to those of Deonna and Teroni (2012, chaps. 5–6). It is unfortunate that some recent approaches (e.g., Mitchell 2021; Müller 2019) are not covered.
- 3 Here, the term "function" refers to a representational function à la Dretske. In this sense, emotions can be said to represent values. Yet the attitudinal view argues that values do not feature in the *content* of emotions.

3198 My emotion of fear implies an evaluation of the dog as danger-
3199 ous: I apply to the content of my representation (the dog) a target
3200 (dangerousness). And this emotion is correct if it targets a state
3201 of affairs that fits with the mobilized representation, i.e., if dan-
3202 gerousness (the target) is indeed one of the properties of the
3203 represented dog (the content). (p. 134)

3204 A refinement of the attitudinal theory is then developed to circumvent the
3205 common objection that emotions are not reducible to bodily feelings (Goldie
3206 2000; Mitchell 2021). Lepine insists on the way emotions *direct our attention*
3207 on aspects of the intentional object and considers that, at least in the case of
3208 mild emotions, we need to focus on this rather than on bodily feelings (p. 139).
3209 He argues that attention is a necessary component as well, and sometimes
3210 more salient than the bodily component (p. 141, see also Deonna and Teroni
3211 2015).

3212 This elegant (and plausible!) refinement makes sense of the idea that bodily
3213 feelings are directed toward the external world—since they accompany and
3214 are calibrated by an attentional mechanism. In addition, just as Brady (2013),
3215 Lepine can explain how emotions, while not representing values in their
3216 content, lead us to focus on natural properties in the world that constitute the
3217 basis for an understanding of values (p. 143).

3218 Chapter (5) focuses on the correctness conditions of emotions and, there-
3219 fore, on what it means for an emotion to fit a value. In a new and stimulating
3220 way, Lepine contrasts an *independentist* view with a *motivational* view of
3221 correctness. According to the first view—attributed to Tappolet (2016) and
3222 D’Arms and Jacobson (2000)—correctness is determined solely by the natural
3223 properties instantiated by the intentional object of the emotion. For instance,
3224 my sense of humor is irrelevant to determining the correctness of my amuse-
3225 ment at Julie’s joke; the only thing we need to consider is whether the joke is
3226 objectively funny. Lepine rejects this analysis and suggests that (iv) the cor-
3227 rectness of emotions also depends on their *congruence with our “background*
3228 *motivations”*—i.e., desires, preferences, feelings, character traits...—as well
3229 as the coherence of these motivations, as we shall see later (p. 192).

3230 There are psychological and axiological elements in Lepine’s view. On the
3231 psychological side, motivations are considered as a necessary base of emotions
3232 (p. 163, see also Baier 2004; Roberts 2003). On the axiological side, Lepine
3233 adopts (v) a form of (neo-) *sentimentalism* according to which values are un-
3234 derstood in terms of appropriate emotions (p. 151)—e.g., injustice is what

deserves indignation (Brentano 1902). Thus, if appropriate emotions underlie values, and if emotions are based on motivations, then (by transitivity) motivations underlie values (at least partially). This leads to a *subjectivist view* of values: it is only when x 's natural properties are prized by humans (or sentient creatures) that x deserves pro-attitude A and thus is good.

Note that Lepine considers that motivations are themselves subject to correctness conditions: motivations are considered correct until proven otherwise, i.e., until they contradict our experience, higher-order beliefs, or (social, prudential, moral...) norms (p. 180). It is thus difficult to possess racist motivations without encountering contradictions (p. 182). Motivationalism is thus a naturalist view on value that occupies a middle ground between “raw” subjectivism and naïve realism.

Finally, in chapter (6), Lepine argues that the justification of emotions is also impacted by motivations. The independentist argues that emotions are justified when we can mention natural properties on which the corresponding value may supervene (“Why am I afraid of that dog? Did you see its teeth, its bloodshot eyes, its lowered tail?!”). The motivationalist replies that subjective motivations also play a justificatory role—as we shall see later.

In the very last section of *La nature des émotions*, Lepine attempts to demonstrate that emotions, despite their subjective nature, are a trustworthy tool for evaluative judgment (p. 224). Contrary to the idea that emotions are prone to many “false positives” (see Goffin 2023), Lepine argues that even when we “confabulate” to justify our (inappropriate) emotions, these confabulations are relatively plausible (p. 218) and should not prevent us from trusting our emotions most of the time (p. 224).

Considering the author's clear view on emotions and values, I may suggest only a few challenges aimed at extending the discussion he proposed.

Among these challenges, we might mention that Lepine does not substantiate his parallel between core relational themes and formal objects. Yet, as Teroni (2023) points out, psychologists have a hard time reconciling the core relational theme or “molar value” (such as injustice, dangerousness, sublime...) and the “molecular values” targeted by each appraisal check (such as relevance, urgency, power...). Since psychologists tend to subjectivize molecular values, this might confer an advantage to the motivational view, as long as the shifts between molecular value, molar value, and, finally, formal object are conceptually possible.

Another challenge concerns both psychological and axiological motivationalism. The two aspects seem inseparable in Lepine's mind, and this, in my

3273 opinion, implies that adopting motivationalism carries a lot of presupposi-
3274 tions. For instance, we find very little argument against the realist approaches
3275 to values except the evocation of one aspect of the queerness of values (Mackie
3276 1977): isn't it odd that an "objective" property of the world (value) has moti-
3277 vating or normative power (pp. 169–170)? The motivational view dissolves
3278 this issue, but it is far from being the only solution (see Enoch 2011), and it
3279 comes with costs and concessions.

3280 Consider the cost at the psychological level. A direct and acknowledged
3281 consequence of motivationalism is that there can be no such thing as emo-
3282 tional discoveries (p. 167, p. 173). Just as Sartre (1940) claimed that we can
3283 never discover anything with imagination other than what we've put into
3284 it, Lepine maintains that we never acquire new motivations by feeling an
3285 emotion. This is questionable. Consider Pablo being forced to attend opera—a
3286 musical genre he has no motivation to listen to. However, this time, he is
3287 touched; from now on, he is willing to come back every month and add opera
3288 playlists on Spotify. According to Lepine, Pablo's emotion necessarily arises
3289 from a pre-existing motivational basis. At a certain degree of generality, this is
3290 indisputable: maybe Pablo has a preference or a desire for music (in general),
3291 beauty, or pleasant moments. Yet, this seems to miss the point raised by schol-
3292 ars acknowledging the possibility of emotional discoveries. If we work with a
3293 fine-grained notion of motivation, we seem to acquire new specific interests
3294 or re-evaluate (quite radically) states of affairs thanks to our emotions. The
3295 only replies available to Lepine are either to assume that emotions that are not
3296 based on pre-existing motivations are inappropriate (see p. 185, p. 222) or that
3297 the relevant specific motivations are unconscious (see p. 188). This is unfortu-
3298 nate considering that the latter is painfully ad hoc, whereas the former clashes
3299 with cases of "outlaw emotions" that seem fitting even though they contrast
3300 with our personality (Silva 2021). Now, if Lepine rejects the fine-grained ap-
3301 proach, one might ask why congruence with background motivations should
3302 count as correctness conditions (p. 192): a condition that cannot be incorrect
3303 is incongruent with the common understanding of correctness conditions.

3304 At the axiological level, the motivational approach is convincing when
3305 we consider personal values (see Rønnow-Rasmussen 2007). For instance,
3306 my disappointment at a friend's betrayal may be justified by my twenty-year
3307 attachment to that friend (see Bell 2011). It is so because betrayal is the kind
3308 of value that depends on a relationship; it cannot be instantiated between
3309 two strangers. But when we consider impersonal values, motivationalism
3310 loses its panache. Epistemic values, for example, hardly seem to depend on

3311 our motivations. Invoking subjective motivation to justify an epistemic emo-
3312 tion comes across as odd—my astonishment at discovering a mathematical
3313 demonstration cannot be justified by my passion for Pythagoras. This provides
3314 a psychological explanation of the emotion (someone who does not share my
3315 passion for Pythagoras would not have been astonished), not a *justification*,
3316 which depends, e.g., on whether there are reasons to think that the proof is
3317 sound and noteworthy.

3318 Lepine is aware that motivationalism blurs the justification/explanation
3319 contrast (p. 200) and might reply that impersonal values are based on interests
3320 shared by all mankind (see p. 174). I disagree because I think that impersonal
3321 values render the world better for human beings, even when they currently
3322 have no motivation to promote them. In the world depicted by the movie
3323 *Idiocracy*—where the planet’s most foolish inhabitants have reproduced in
3324 large numbers to the point of engendering a society with no culture or histor-
3325 ical knowledge—nobody is motivated to acquire knowledge. I would not say,
3326 however, that knowledge has no value in this world. People are just wrong!
3327 You may say that they *should* be motivated by knowledge. Then, if knowledge
3328 possesses value not because people have motivations but because knowledge
3329 deserves to motivate them, we lose the motivationalist view on the way and
3330 go back to pure (neo-)sentimentalism.

3331 Samuel Lepine’s monograph is subtitled “une introduction partisane”: it is
3332 introductory in the noblest and most exciting sense of the word; it offers an
3333 overview of affective topics in philosophy and psychology without detracting
3334 from the precision and complexity of the debates. Written in crystal-clear
3335 French (guarantee without any trace of Sorbonnian style!), *La nature des*
3336 *émotions* results in a conceptually plausible and empirically supported defense
3337 of the appraisal theory, the attitudinal view (re-visited), and the motivational
3338 view (introduced here).

3339
3340
3341
3342
3343

Steve Humbert-Droz

 0000-0002-0680-649X

Université de Genève

Swiss Center for Affective Sciences, Geneva

Steve.Humbert-Droz@unige.ch

References

- 3344
- 3345 BAIER, Annette C. 2004. "Feelings That Matter." in *Thinking About Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. SOLOMON, pp. 200–213. Series
3346 in Affective Science. Oxford: Oxford University Press, doi:10.1093/oso/97801951
3347 53170.003.0014.
- 3348
- 3349 BELL, Macalester. 2011. "Globalist Attitudes and the Fittingness Objection." *The Philo-*
3350 *sophical Quarterly* 61(244): 449–472, doi:10.1111/j.1467-9213.2010.684.x.
- 3351 BRADY, Michael Sean. 2013. *Emotional Insight: The Epistemic Role of Emotional Expe-*
3352 *rience*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199685523.00
3353 1.0001.
- 3354 BRENTANO, Franz. 1889. *Vom Ursprung sittlicher Erkenntnis*. Leipzig: Duncker &
3355 Humblot. Second edition: Brentano (1921).
- 3356 —. 1902. *The Origin of the Knowledge of Right and Wrong*. Westminster: Archibald
3357 Constable & Co., Ltd. Translation of Brentano (1889) by Cecil Hague.
- 3358 —. 1921. *Vom Ursprung sittlicher Erkenntnis*. 2nd ed. Hamburg: Felix Meiner Verlag.
3359 Second edition of Brentano (1889), edited by Oskar Kraus.
- 3360 CUMMINS, Robert. 1996. *Representations, Targets, and Attitudes*. Cambridge, Mas-
3361 sachusetts: The MIT Press, doi:10.7551/mitpress/5887.001.0001.
- 3362 D'ARMS, Justin and JACOBSON, Daniel. 2000. "The Moralistic Fallacy: On the 'Appro-
3363 priateness' of Emotions." *Philosophy and Phenomenological Research* 61(1): 65–90,
3364 doi:10.2307/2653403.
- 3365 DEONNA, Julien Amos and TERONI, Fabrice. 2008. *Qu'est-ce qu'une émotion?* Chemins
3366 Philosophiques. Paris: Librairie philosophique Jean Vrin.
- 3367 —. 2012. *The Emotions: A Philosophical Introduction*. London: Routledge. Substantially
3368 enlarged edition and translation of Deonna and Teroni (2008), doi:10.4324/9780
3369 203721742.
- 3370 —. 2015. "Emotions as Attitudes." *Dialectica* 69(3): 293–311, doi:10.1111/1746-8361.1
3371 2116.
- 3372 —. 2024. "Emotions and Their Correctness Conditions: A Defense of Attitudinalism."
3373 *Erkenntnis* 89(1): 45–64, doi:10.1007/s10670-022-00522-0.
- 3374 DUKES, Daniel, ABRAMS, Kathryn, ADOLPHS, Ralph and AHMED, Mohammed E. 2021.
3375 "The Rise of Affectivism." *Nature Human Behaviour* 5(7): 816–820. Coauthored by
3376 many members of the CUSO Centre for Affective Sciences, University of Geneva,
3377 doi:10.1038/s41562-021-01130-8.
- 3378 ENOCH, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford:
3379 Oxford University Press, doi:10.1093/acprof:oso/9780199579969.001.0001.
- 3380 GOFFIN, Kris. 2023. "Better Scared than Sorry: The Pragmatic Account of Emotional
3381 Representation." *Erkenntnis* 88(6): 2633–2650, doi:10.1007/s10670-021-00470-1.
- 3382 GOLDIE, Peter. 2000. *The Emotions: A Philosophical Exploration*. Oxford: Oxford Uni-
3383 versity Press, doi:10.1093/0199253048.001.0001.

- 3384 GRIFFITHS, Paul E. 1997. *What Emotions Really Are: The Problem of Psychological*
 3385 *Categories*. Chicago, Illinois: University of Chicago Press.
- 3386 LAZARUS, Richard S. and FOLKMAN, Susan. 1984. *Stress, Appraisal, and Coping*. New
 3387 York: Springer, <https://archive.org/details/stressappraisal0000laza>.
- 3388 LEPINE, Samuel. 2023. *La nature des émotions: Une introduction partisane*. Analyse et
 3389 philosophie. Paris: Librairie philosophique Jean Vrin.
- 3390 MACKIE, John Leslie. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.
- 3391 MITCHELL, Jonathan. 2021. *Emotion as Feeling Towards Value: A Theory of Emotional*
 3392 *Experience*. Oxford: Oxford University Press, doi:10.1093/oso/9780192846013.001.
 3393 0001.
- 3394 MOORS, Agnes. 2013. "On the Causal Role of Appraisal in Emotion." *Emotion Review*
 3395 5(2): 132–140, doi:10.1177/1754073912463601.
- 3396 MÜLLER, Jean Moritz. 2019. *The World-Directedness of Emotional Feeling: On Affect*
 3397 *and Intentionality*. London: Palgrave Macmillan, doi:10.1007/978-3-030-23820-9.
- 3398 ROBERTS, Robert C. 2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge:
 3399 Cambridge University Press, doi:10.1017/cbo9780511610202.
- 3400 RØNNOW-RASMUSSEN, Toni. 2007. "Analysing Personal Value." *The Journal of Ethics*
 3401 11(4): 405–435, doi:10.1007/s10892-006-9012-1.
- 3402 ROSEMAN, Ira J. and SMITH, Craig A. 2001. "Appraisal Theory: Overview, Assump-
 3403 tions, Varieties, Controversies." in *Appraisal Processes in Emotion: Theory, Methods,*
 3404 *Research*, edited by Klaus R. SCHERER, Angela SCHORR, and Tom JOHNSTONE, pp.
 3405 3–19. Oxford: Oxford University Press, doi:10.1093/oso/9780195130072.003.0001.
- 3406 RUSSELL, James A. 2003. "Core Affect and the Psychological Construction of Emotion."
 3407 *Psychological Review* 110(1): 145–172, doi:10.1037/0033-295X.110.1.145.
- 3408 SARTRE, Jean-Paul. 1940. *L'Imaginaire: Psychologie phénoménologique de l'imagination*.
 3409 Paris: Gallimard.
- 3410 SILVA, Laura. 2021. "The Epistemic Role of Outlaw Emotions." *Ergo* 8(23): 664–691,
 3411 doi:10.3998/ergo.1160.
- 3412 TAPPOLET, Christine. 2016. *Emotions, Values, and Agency*. Oxford: Oxford University
 3413 Press, doi:10.1093/acprof:oso/9780199696512.001.0001.
- 3414 TERONI, Fabrice. 2023. "Evaluative Theories in Psychology and Philosophy of Emo-
 3415 tion." *Mind and Language* 38(1): 81–97, doi:10.1111/mila.12374.
- 3416 ZAJONC, Robert B. 1984. "On the Primacy of Affect." *American Psychologist* 39(2):
 3417 117–123, doi:10.1037/0003-066X.39.2.117.

3418

Review of Landgrebe and Smith (2022)

JONATHAN SIMON

3419

LANDGREBE, Jobst and SMITH, Barry. 2022. *Why Machines Will Never Rule the World. Artificial Intelligence Without Fear*. London: Routledge.

3420

3421

In this formidable book, Jobst Landgrebe and Barry Smith argue that no AI will ever attain human-level intelligence. The book is a challenging read, but it is full of important insights, its master argument is original, it is informed by an impressive array of sources, and it is timely. It merits philosophical attention. The book is also noteworthy because it is the collaboration of an engineer and a philosopher. Landgrebe runs an AI software company: he has an M.D. and a Ph.D. in biochemistry (although notably, his grandfather Ludwig was a famous phenomenologist). Smith is an expert on the Austrian phenomenological tradition and in formal and applied ontology (where, notably, he has pioneered a way to apply philosophical theory and method to data engineering). This kind of collaboration is vital but hard to achieve.

3422

3423

3424

3425

3426

3427

3428

3429

3430

3431

3432

3433

3434

3435

3436

3437

3438

3439

3440

3441

3442

3443

3444

3445

3446

3447

In very broad strokes, the book's master argument is this: Human-level intelligence requires coping with and getting on in environments that are complex dynamical systems—that is, environments that are open and chaotic and subject to feedback effects, with trends and statistics that change over time (think: the weather or the stock market). Data models of such complex dynamical systems are always mere approximations not good enough to enable long-term prediction in a complex world in constant flux (think: why you don't know whether it will rain next Saturday and why you can't reliably beat the stock market). But AI systems are just data models. So, in principle, they can't enable the sort of coping that humans are capable of. What about full-on emulations of the human neuro-cognitive system? That system, too, is a complex dynamical system, so no ensemble of algorithms based on data models can approximate it well enough for full-on emulation (in Landgrebe and Smith's terminology: well enough for a model that is both "adequate" and "synoptic"—where this means a model that enables predictions that are accurate enough for the task at hand).

3448 Here is how they proceed. Chapters 1–6 present their general picture of
 3449 intelligence and mindedness, as well as language and sociality, and why they
 3450 think that having these capabilities entails that we can cope *with* and also that
 3451 our neuro-cognitive systems *are* complex dynamical systems. Highlights from
 3452 these chapters include fresh insights on the mind-body problem in chapter
 3453 two, the authors’ breakdown of human intelligence into “primal” and “objec-
 3454 tifying” intelligence and their critique of reward-optimization conceptions of
 3455 intelligence in chapter three, and detailed analyses of language and social-
 3456 ity informed by both phenomenology and empirical work in chapters 4–6.
 3457 Chapters 7–8 deliver the linchpin of the master argument: the claim that AI
 3458 systems cannot adequately or synoptically model complex dynamical systems.
 3459 Chapters 9–12 argue that it follows from what they argue in chapters 1–8 that
 3460 AGI is impossible, that machines will not master human language or sociality,
 3461 and also that mind uploading is a waste of time, as are attempts to create
 3462 digital minds to carry on our civilization. Finally, chapter thirteen makes
 3463 positive recommendations, discussing what Landgrebe and Smith think AI is
 3464 good for and how they think it should be used.

3465 Let’s look closer at the linchpin of the master argument. In section 7.5.2,
 3466 Landgrebe and Smith enumerate seven key features of complex systems:

3467 CHANGE AND EVOLUTIONARY CHARACTER (pp. 126–128). Complex
 3468 systems evolve in various ways: the system’s boundaries can shift,
 3469 new elements come, and old elements go. In many cases, complex
 3470 systems can undergo changes in the *types* of elements they contain
 3471 or interactions they participate in.

3472 ELEMENT-DEPENDENT INTERACTIONS (pp. 128–129). Complex
 3473 systems typically have different kinds of functionally individuated
 3474 elements, e.g., the different roles played by proteins, kinases, and
 3475 ATP in phosphorylation (contrasted with the way that mass and ve-
 3476 locity are all you need to chart all of the interactions of a Newtonian
 3477 system). Elements of a system can also change their functions over
 3478 time.

3479 FORCE OVERLAY (pp. 130–131). Complex systems typically involve
 3480 interactions between all four of the basic physical interactions (EM,
 3481 gravity, strong and weak).

3482 NON-ERGODIC/COMPLEX PHASE SPACES (pp. 131–132). We cannot
3483 predict the trajectory of a complex system over its phase space by
3484 averaging over volumes of that phase space.

3485 DRIVENNESS (pp. 132–136). A driven system is a system that does
3486 not generally converge to equilibrium because it has access to a
3487 reliable energy source.

3488 CONTEXT-DEPENDENCE (p. 137). The interface between a complex
3489 system and its environment is constantly changing, e.g., which ele-
3490 ments are part of the system vs. part of the environment, or what
3491 states the system can occupy.

3492 CHAOS (p. 137–138). Chaotic systems are unpredictable because
3493 small differences in initial conditions may lead to large differences
3494 down the road.

3495 Say that systems having all seven of these features are *fully complex*. Land-
3496 grebe and Smith's master argument is that AGI would only be possible if fully
3497 complex systems could be adequately and synoptically modelled (either the
3498 ones in the environment or the ones in the brain), but that fully complex
3499 systems cannot be adequately and synoptically modelled.

3500 I am not going to say much here about whether they are correct that fully
3501 complex systems cannot be adequately and synoptically modelled. It is in-
3502 tractable to find exact solutions to the dynamical equations for most complex
3503 systems (even ones that are not *fully complex*, like three-body gravitational
3504 problems). Approximation is thus the name of the game. The more chaos a
3505 system exhibits, the more its distribution changes over time, etc., the harder it
3506 can be to find approximations that are both tractable and accurate enough for
3507 the problem at hand. This much is beyond dispute. However, Landgrebe and
3508 Smith are arguing for something extremely ambitious: not just that suitable
3509 approximations are sometimes or even typically very costly, but that they are,
3510 in principle, unavailable for a wide range of cases and will continue to be,
3511 even with the increases in computing power that we can expect the future to
3512 bring. This is less clear. It is hard not to look at, for example, NASA's recent
3513 successes on missions like DART or OSIRIS-REx and come away with the
3514 impression that, when there is a will to find suitably accurate approximations,
3515 there is a way.

3516 For the remainder, though, I'll focus on the more philosophical questions
 3517 that arise in Landgrebe and Smith's defence of their claim that AGI would
 3518 only be possible if fully complex systems could be adequately and synoptically
 3519 modelled. They pursue two routes to this conclusion. I'll call these the
 3520 argument from coping and the argument from emulation.

3521 According to the argument from coping, there are fully complex systems in
 3522 our environments; we cope with them, and AGI is only possible if you can
 3523 achieve this coping by means of adequate and synoptic modelling.

3524 According to the argument from emulation, our neuro-cognitive systems
 3525 are fully complex systems, and AGI is only possible if you can emulate them
 3526 by adequately and synoptically modelling them.

3527 The coping argument is mainly developed in an earlier work, Landgrebe
 3528 and Smith (2021), but it serves as background for the emulation argument,
 3529 which is the focus of the present book.

3530 On coping: here, I worry that there is an equivocation. I'll grant unequivocally
 3531 that there are fully complex systems in our environments, like weather
 3532 or the stock market, but I'm not sure what it means to allow that we cope with
 3533 them. Do individuals really cope with hurricanes or stock market crashes?
 3534 Arguably, imperfect though they are, computational models of hurricanes are
 3535 the best tools we have for coping with hurricanes. Our coping abilities turn
 3536 on bounded, often flawed approximations of the chaotic world around us. If
 3537 those models are enough for coping, then clearly calling for adequate and
 3538 synoptic modelling sets the bar too high (as a necessary condition for an AI
 3539 system to count as coping). On the other hand, if these models aren't enough
 3540 for coping, then, presumably, we can't cope. Either way, the argument from
 3541 coping fails: either we cannot cope, or the computational methods we use to
 3542 cope (which fall short of adequate and synoptic modelling) suffice for coping.

3543 On emulation: here, I have a few worries. First, it isn't obvious that our
 3544 neuro-cognitive systems are fully complex. For example, it is debatable how
 3545 much chaos there is in the healthy brain, as opposed to criticality or near-
 3546 criticality (see O'Byrne and Jerbi 2022).

3547 Second, there is an equivocation lurking in the notion of 'emulation' at
 3548 issue. Is the aim of emulation to create a perfect replica of a specific token
 3549 system, e.g., to build a concrete model of a specific hurricane, accurate enough
 3550 to predict where and when that particular hurricane will make landfall? Or is
 3551 the aim simply to generate a new sample from the same distribution, a new
 3552 token of the relevant type? If we are after our own digital immortality, then

3553 maybe we must pursue the former project. In contrast, if all we want to do is
3554 build an AGI, we only need pursue the latter project.

3555 But I worry that from the claim that the human neuro-cognitive system is
3556 fully complex, taken together with the claim that it is impossible to adequately
3557 and synoptically model fully complex systems, we only get the impossibility
3558 of *token-level* emulation, leaving open the possibility of type-level emulation.

3559 Maybe we cannot build a model of an actual specific hurricane currently
3560 out at sea that will allow us to predict to the minute or square mile when and
3561 where it makes landfall. But we can build models of hurricanes that embody
3562 the profile of hurricanes in general (see Weisberg 2013 for a discussion of
3563 distinctions between kinds of predictive models). So too here: the AGI we
3564 build might not be a perfect copy of you or me, but this does not preclude
3565 that it is adequate as a type-level emulation—especially so since the type in
3566 question is *system that has human-level intelligence* and not *system that is as*
3567 *complex as humans*.

3568 Of course, the token-level question is important, too; it seems relevant
3569 to questions of uploading and digital immortality. But there is no simple
3570 refutation of the possibility of uploading or digital immortality here since
3571 numerical identity over time does not require qualitative identity over time.

3572 For the type-level question, of course, it remains to be shown that such
3573 an emulation is possible. Maybe adequate and synoptic modelling of fully
3574 complex systems would still be required, even if the thing we create is not
3575 constrained to perfectly resemble an existing intelligent being.

3576 This is one way that the coping argument fits into the dialectic of the book:
3577 if human-level coping involves harnessing our full complexity in order to ride
3578 out storms, and we can do this as well as we would if we had adequate and
3579 synoptic models of those storms, then emulating our type *in silico* presumably
3580 entails adequate and synoptic modelling of fully complex systems somewhere
3581 or other. But if, as I suggest above, we don't cope as well as that, then it does
3582 not follow that emulating us entails adequate and synoptic modelling of fully
3583 complex systems somewhere or other.

3584 Landgrebe and Smith might advocate a slightly weaker fallback claim,
3585 which is that the full (or nearly-full) complexity of our neuro-cognitive systems
3586 surely must have something to do with all of our intellectual successes, and so
3587 nothing will succeed in emulating us (at the type level) if it cannot at the very
3588 least *instantiate* the seven key features that make a system fully complex. This
3589 isn't obvious: again, the type to be emulated is not *system that is as complex*

3590 *as humans* but rather *system that has human-level intelligence*. Still, it merits
 3591 consideration.

3592 But there is no clear reason to doubt that we can build digital systems that
 3593 instantiate these features. Think of Conway's Game of Life or other complex
 3594 systems generated *in silico* via cellular automaton rules. If we assess criteria
 3595 like drivenness and force-overlay within the simulation, arguably, these are
 3596 fully complex. So, too, for deep learning systems, especially if we focus on
 3597 their dynamics during training (as opposed to inference). During training,
 3598 parameters evolve (and the sampled distribution changes). Stochastic gradi-
 3599 ent descent is non-ergodic: systems get stuck in local minima all the time.
 3600 It can also be chaotic or near-chaotic: there are trajectories that pass along
 3601 the borders of basins of attraction for local minima. Even during inference,
 3602 some complex features can be seen. For example, the whole point of atten-
 3603 tional mechanisms is to allow models to handle context during inference
 3604 (see Søgaard 2022), and functional differentiation between different neural
 3605 network layers (attention vs. feed-forward, pooling vs. convolution, etc.) exem-
 3606 plifies element-dependent interactions. Finally, let's not forget about neural
 3607 organoids, which are programmable assemblies of biological neural cells:
 3608 these certainly fit the bill if nothing *in silico* does.

3609 Thus, we have a few reasons to doubt that Landgrebe and Smith fully suc-
 3610 ceed. Even so, their arguments are important and merit further consideration.
 3611 If our uploads are guaranteed to differ from us, this problematizes the claim
 3612 that we can survive into them or that they preserve us, even if the matter is
 3613 far from settled. And I certainly agree with Landgrebe and Smith about the
 3614 limits of *current* AI systems and that the question of how and to what degree
 3615 we adaptively harness our underlying complexity is a key open question, one
 3616 which we must answer to fully understand the difference between biological
 3617 minds and AI systems. That said, I am still afraid.*

3618 Jonathan Simon

3619  0000-0001-6081-4127

3620 Université de Montréal

3621 jonathan.simon@umontreal.ca

* Thanks to David Anderson, Jaan Aru, Axel Constant, George Deane, Jordan O'Byrne, Steve Petersen, GPT-4, Claude, and the editors of *Dialectica* for comments on an earlier draft.

3622 **References**

- 3623 LANDGREBE, Jobst and SMITH, Barry. 2021. “An Argument for the Impossibility of
3624 Machine Intelligence.” ArXiv preprint, doi:10.48550/arXiv.2111.07765.
- 3625 —. 2022. *Why Machines Will Never Rule the World: Artificial Intelligence Without Fear*.
3626 London: Routledge, doi:10.4324/9781003310105.
- 3627 O’BYRNE, Jordan and JERBI, Karim. 2022. “How Critical Is Brain Criticality?” *Trends*
3628 *in Neurosciences* 45(11): 820–837, doi:10.1016/j.tins.2022.08.007.
- 3629 SØGAARD, Anders. 2022. “Understanding Models Understanding Language.” *Synthese*
3630 200(6): 443, doi:10.1007/s11229-022-03931-4.
- 3631 WEISBERG, Michael. 2013. *Simulation and Similarity: Using Models to Understand the*
3632 *World*. New York: Oxford University Press, doi:10.1093/acprof:oso/978019993366
3633 2.001.0001.

PROOF

PROOF

3634

Published by *Philosophie.ch*

3635

Verein philosophie.ch

3636

Fabrikgässli 1

3637

2502 Biel/Bienne

3638

Switzerland

3639

dialectica@philosophie.ch

3640

<https://dialectica.philosophie.ch/>

3641

ISSN 0012-2017

3642

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

3643

3644

Dialectica is supported by the [Swiss Academy of Humanities and Social Sciences](https://www.snf.ch/).

3645

3646

Abstracting and Indexing Services

3647

The journal is indexed by the Arts and Humanities Citation Index, Current Contents, Current Mathematical Publications, Dietrich's Index

3648

Philosophicus, IBZ — Internationale Bibliographie der Geistes- und

3649

Sozialwissenschaftlichen Zeitschriftenliteratur, Internationale Bibliographie der Rezensionen Geistes- und Sozialwissenschaftlicher Literatur, Linguistics

3650

and Language Behavior Abstracts, Mathematical Reviews, MathSciNet,

3651

Periodicals Contents Index, Philosopher's Index, Repertoire Bibliographique de la Philosophie, Russian Academy of Sciences Bibliographies.

3652

3653

3654

3655

Contents

3656	ADAM PATTERSON, <i>Avoid Avoiding the Wishful Thinking Problem</i>	257
3657	RAN LANZET, <i>Whittle vs. Cantor on the Size of Infinite Sets</i>	271
3658	EMILIE PAGANO, <i>Functionalism, Pluralities, and Groups</i>	283
3659	SERENA GALLI, <i>Causal Inference from Big Data?: A Reply to Pietsch (2021)</i> ..	319
3660	CIRO DE FLORIO & ALDO FRIGERIO, <i>Retro-Closure Principle and Omniscience</i>	335
3661	AGNÈS BAEHNI, <i>Review of Schofield (2021)</i>	363
3662	STEVE HUMBERT-DROZ, <i>Review of Lepine (2023)</i>	369
3663	JONATHAN SIMON, <i>Review of Landgrebe and Smith (2022)</i>	377

PROOF