Philosophy and Model Theory

Tim Button and Sean Walsh
with a historical appendix by Wilfrid Hodges

3

Ramsey sentences and Newman's objection

Whilst the main focus of this book is the philosophy of mathematics, Chapter 2 puts us in an excellent position to consider a particular topic in the philosophy of science. In particular, in this chapter, we will consider some recent discussions of Ramsey sentences and Newman's objection.

In 1928, Newman presented a now famous objection to Russell's (1927) theory of causation. The simplest form of his objection is just an application of the Push-Through Construction of §2.1. In this simple form, Newman's objection has a rather specific target. However, as subsequent commentators have noted, Newman's objection can be directed at a wider range of positions. The aim of this chapter is to develop several versions of the Newman objection, and to relate them to different notions of *conservation*. Our main observation is that the best version of Newman's objection is a slightly spruced-up application of the Push-Through Construction, and that the dialectic surrounding this objection should precisely parallel the dialectic surrounding Putnam's permutation argument.

3.1 The o/t dichotomy

The focus of this chapter is a certain kind of *structuralism* within the philosophy of science. So we must start by motivating that version of structuralism. It begins with the idea is that our scientific theorising is split into two parts. The *okay* part is in good standing, and poses no particular philosophical puzzles. Unfortunately, there is also a *troublesome* part.

Such a dichotomy might arise by treating the *observational* as okay but the *theoretical* as troubling. Consider, in particular, the following view. Some statements are observation statements. Their role is simply to make claims about observations, and nothing beyond that. There is no mystery (on this view) concerning how observation statements work: to determine whether they are true, one simply makes a relevant observation, and thereby checks whether the world is as the observation statement claims. An example of an observation statement might be 'a cat is on the mat', and to check whether this is true, you simply *look and see*. But, continuing with this view, some statements are theoretical statements. An example might be 'carbon has fifteen isotopes, only two of which are stable'. And there may be some initial mystery concerning how we could so much as *check* whether the world is as the theoretical statement claims.

For those who want to effect such a dichotomy between the observational and the theoretical, the dichotomy shines through in the *vocabulary*. The predicate '... is a cat' is an observation-predicate, but the predicate '... is a carbon isotope' is not. To keep track of this, in what follows we let $\mathscr{V}$ be the signature which we use to frame our scientific theories. (Recall from Definition 1.1 that a signature is a formal, regimented *vocabulary*.) Then $\mathscr{V}_o$ is a privileged sub-signature of $\mathscr{V}$, which is treated as $\mathscr{V}$'s *observational* vocabulary, and $\mathscr{V}_t$, i.e. $\mathscr{V} \setminus \mathscr{V}_o$, is its *theoretical* vocabulary.

Now, we have motivated an okay/troublesome dichotomy—more briefly, an o/t dichotomy—by considering a dichotomy between the observational and theoretical. We do not want to suggest that this is the only way to generate an o/t dichotomy.[1] Equally, we do not want to suggest that there *is* an important o/t dichotomy (we revisit this in §3.4). We just want to use model theory to explore what happens *if* one effects some o/t dichotomy.

3.2 Ramsey sentences

Suppose, then, that we effect an o/t dichotomy. Our next question is how we should handle this within our philosophising. This is the context in which Ramsey sentences present themselves.

Ramsey sentences are sentences of second-order logic. We outlined second-order logic in §§1.9–1.11; here are the main points. Syntactically, second-order logic allows us to quantify into 'predicate position'. Semantically, matters are more subtle; but we can afford to relegate the subtleties to footnotes, since all of the results of this chapter go through with either of the two most common semantics for second-order logic, namely, *full* semantics and *faithful Henkin* semantics. (For more, see §§1.9–1.11, and footnotes 17, 23, 24, and 30 of this chapter.)

Ramsey sentences are defined *from* theories. So we start by making some assumptions about the theory, T, whose Ramsey sentence we want to define. (These assumptions hold good throughout this chapter, but no further.)

First, we assume that our theory, T, is given in the signature $\mathscr{V}$, which has an appropriate o/t dichotomy. So, $\mathscr{V}_o$ and $\mathscr{V}_t$ are disjoint, but together exhaust $\mathscr{V}$.

Second, we assume that $\mathscr{V}$ is relational, i.e. that it contains only predicates. This is no loss, as we can simulate constants and functions using relations and identity.

Third, we assume that T is a *finite* set of sentences. This is a bigger assumption.

---

[1] Lewis (1970), for example, considered the *old/theoretical* distinction. And in the philosophy of mathematics, there is Hilbert's (1925: 179) analogous real/ideal distinction. In §4.7, we consider the idea that *infinitesimals* are ideal.

However, once we have made it, we can harmlessly treat T as a *single* sentence, by conjoining all of its finitely many members. Moreover, since T is finite, only finitely many $\mathscr{V}_o$-symbols occur in T. And this allows us to construct T's Ramsey sentence. Intuitively, we do this by treating all of T's theoretical predicates as variables—i.e. we treat the $\mathscr{V}_t$ relation symbols as relation-variables—and then bind them with existential quantifiers:

Definition 3.1: *Let T be a finite $\mathscr{V}$-theory, and let $R_1, \ldots, R_n$ be the only $\mathscr{V}_t$-predicates appearing in T. Where $\hat{T}$ is the conjunction of the sentences in T, and where $\hat{T}[\overline{X}/\overline{R}]$ is the result of replacing each instance of the predicate R_i in $\hat{T}$ with the variable X_i, we say that T's* Ramsey sentence *is*

$$\text{Ramsey}(T) := \exists X_1 \ldots \exists X_n \hat{T}[\overline{X}/\overline{R}]$$

Consequently, Ramsey(T) is a second-order $\mathscr{V}_o$-sentence; that is, the only symbols occuring in Ramsey(T) are logical expressions and $\mathscr{V}_o$-symbols. (And note that T's Ramsey sentence is only *definable* when T is finite.)

3.3 The promise of Ramsey sentences

Ramsey sentences were first formulated in Ramsey's posthumously published 'Theories'.[2] Carnap brought them to prominence in the late 1950s and 1960s, relating them to his long-running concerns about the meaning of theoretical terms.[3] In this section, we will outline a contemporary motivation for such concerns.

As noted in §2.3, it is plausible that our (causal) acquaintance with cats helps to explain why our word 'cat' refers to cats rather than to cherries. In this case, however, we are likely to treat '…is a cat' as an observational predicate. It is much less clear what could pin down the reference of *theoretical* predicates. Crudely put, we cannot point at electrons in the same way as we point at cats. And so it may seem that our knowledge of theoretical predicates will have to come via description—in terms of their impact on our more immediate observations—rather than by acquaintance.[4]

If we buy this line of thought, then Ramsey(T) may start to recommend itself to us. After all, Ramsey(T) is a $\mathscr{V}_o$-sentence, and so it contains no theoretical vocabulary. As such, it is not obviously vulnerable to the concerns just mentioned.[5] Moreover, and crucially, not much is lost by considering Ramsey(T) rather than T

[2] Ramsey (1931).

[3] Carnap (1958: 245, 1959: 160–5, 1966: 248, 252, 265ff). For an account of the history, see Psillos (1999: 46–9).

[4] See Maxwell (1971), Zahar and Worrall (2001: 239, 243), and Cruse (2005: 562–3), and the discussion of moderate objects-platonism in Chapter 2.

[5] See Carnap (1958: 242, 245, 1966: 251, 269).

itself, for it is easy to show that the two are *observationally equivalent* in the following sense (for the proof, see Proposition 3.5, below):[6]

Proposition 3.2: *Let T be a finite $\mathscr{V}$-theory. Then $T \vDash \varphi$ iff* $\text{Ramsey}(T) \vDash \varphi$*, for all $\mathscr{V}_o$-sentences φ.*

This result led Carnap to claim that $\text{Ramsey}(T)$ expresses T's 'factual content'.[7]

Suppose we follow Carnap's initial motivations: there are difficulties with the $\mathscr{V}_o$-vocabulary; but $\text{Ramsey}(T)$ does not share those difficulties; and indeed $\text{Ramsey}(T)$ has the same observational content as T, in the sense of Proposition 3.2. But suppose we *also* maintain that $\text{Ramsey}(T)$ expresses some additional content, beyond its observational consequences. That additional content is existential: there *are* some theoretical objects and relations, whose behaviour is specified by $\text{Ramsey}(T)$. It is worth noting that, for those who accept an o/t dichotomy, $\text{Ramsey}(T)$ will seem to provide the fullest thesis that one could plausibly hope to maintain concerning a *troublesome* realm within which determinate reference is bound to be problematic.[8] And this line of thought culminates in what we call *ramsified realism*.[9] This holds that, for a physical theory T,

(a) $\text{Ramsey}(T)$ expresses T's genuine content, but
(b) the content expressed by $\text{Ramsey}(T)$ goes significantly beyond T's mere observational consequences, since $\text{Ramsey}(T)$ makes substantial—and hopefully true—claims about the theoretical.

3.4 A caveat on the o/t dichotomy

In what follows, we will use model-theoretic results to raise problems for ramsified realism. First, though, we should emphasise that there are plenty of problems with ramsified realism which have nothing to do with model theory.

[6] Carnap (1958: 245, 1959: 162–4, 1963: 965, 1966: 252), Psillos (1999: 292 n.6), Worrall (242–3 2007: 150), and Zahar and Worrall (2001).

[7] Carnap (1959: 164–5). Carnap's own interest in Ramsey sentences was sharpened by a further observation: $\text{Ramsey}(T) \wedge (\text{Ramsey}(T) \rightarrow \hat{T}) \vDash \hat{T}$ and $T \vDash \text{Ramsey}(T) \wedge (\text{Ramsey}(T) \rightarrow \hat{T})$. Consequently, Carnap maintained that physical theories can be exclusively and exhaustively factored into a purely *factual* part, $\text{Ramsey}(T)$, and a purely *analytic* part, the conditional $(\text{Ramsey}(T) \rightarrow \hat{T})$, which records the theory's 'meaning postulates' (1958: 246, 1959: 163–4, 1963: 965, 1966: 270–2). So Carnap's own interest in Ramsey sentences is tied to his long-standing interest in analyticity. Admittedly, most contemporary philosophers are less interested in analyticity than Carnap was; so the conditional $(\text{Ramsey}(T) \rightarrow \hat{T})$, which is now called *T's Carnap sentence*, will not detain us.

[8] See Worrall (2007: 148) and Demopoulos (2011: 200).

[9] Psillos (1999: ch.3) seems to suggest that Carnap himself was a ramsified realist. For a contrasting view, see Demopoulos (2003: 384–90, 2011: 195–7). We cannot pursue this in detail, but we note that Carnap (esp. 1966: 256) may well have simply *embraced* the conclusions of the various Newman-style objections against ramsified realism.

Crucially, ramsified realism only makes sense given that one has accepted some o/t dichotomy. And it is worth emphasising that this really is a *dichotomy*. The o/t distinction is *exhaustive*, in that $\mathscr{V}_o \cup \mathscr{V}_t = \mathscr{V}$. It is *exclusive*, in that $\mathscr{V}_o \cap \mathscr{V}_t = \varnothing$. Then all and only the troublesome vocabulary is 'ramsified away' (i.e. replaced with an existentially bound relation-variable).

It is far from clear that natural language—whether mundane or scientific—really displays such a dichotomy. To take an example from Cruse, consider the predicate '*x* is bigger than *y*'.[10] Some instances of this are *observational*: Big Ben is clearly bigger than the tourist next to it. However, some instances seem straightforwardly *theoretical*: protons are bigger than electrons. But some instances seem to be *mixed*: Big Ben is bigger than a proton. So the predicate '*x* is bigger than *y*' resists categorisation as (dichotomously) either observational or theoretical.

If we insist on an o/t dichotomy, then the only way to handle this natural-language predicate will be to split it in half. We will use a $\mathscr{V}_o$-predicate, B_1, to formalise the observational instances, and a $\mathscr{V}_t$-predicate, B_2, to formalise the theoretical instances. When it comes to the mixed instances, we must make a choice as to whether to formalise them with B_1 or B_2. However, it is likely that we will use B_2, since we cannot directly observe the size of the proton, so that these judgments of comparative size must be (somewhat) troublesome. In this manner, we will retain the o/t dichotomy and 'ramsify away' B_2 but not B_1. How (un)satisfying this is will, of course, depend upon whether we think that there really are two dichotomously different relations here.

This simple example illustrates a fundamental point. It is one thing to believe that there is a *distinction* between the observational and the theoretical; it is another to believe that there is a *dichotomy*, in the sense endorsed by ramsified realism. For our part, we believe that any distinction we should draw between between the observational and the theoretical is likely to be, not a once-and-for-all dichotomy, but rough and ready, porous, and context-sensitive.[11] So, *we* are not ramsified realists.

However, to pursue this any further would take us deep into issues at the intersection of philosophy of science and philosophy of perception. We raise the point simply to highlight that the very idea of an o/t dichotomy is quite contentious, and for reasons which have nothing to do with model theory.

3.5 Newman's criticism of Russell

With this caveat out of the way, we will consider various model-theoretic results which raise problems for ramsified realists (who *do* embrace an o/t dichotomy).

[10] Cruse (2005: 561).

[11] This is the line pushed in Button (2013: 50–1), and it draws in various ways on Maxwell (1962: 7–8, 14–15), Putnam (1987: 1, 20–1, 26–40, 1994: 502ff), and Okasha (2002: 316–9).

These problems came to contemporary prominence when Demopoulos and Friedman related Ramsey sentences to Newman's criticism of Russell.[12]

As Newman understood Russell's position in *The Analysis of Matter*, Russell was committed to the doctrine that we could only know the *structure* of the external world. Newman's objection to this view was straightforward. For any structure $\mathcal{W}$,

> Any collection of things can be organised so as to have the structure $\mathcal{W}$, provided there are the right number of them. Hence the doctrine that *only* structure is known involves the doctrine that *nothing* can be known that is not logically deducible from the mere fact of existence, except ('theoretically') the number of constituting objects.[13]

Note that Newman's initial claim is *just* a restatement of the Push-Through Construction from §2.1. Indeed Newman gestures at the Construction himself.[14]

Newman's objection raises serious problems for anyone who thinks that we can know only the structure of the external world. But this seems to be an extremely restricted target.[15] For example, our ramsified realist accepts an o/t dichotomy, and glosses it in terms of the observational and the theoretical. Given the troublesome nature of the theoretical, our ramsified realist might concede that our knowledge of the external *theoretical* realm is limited to its structure. But she will surely claim that we can know more about the external *observational* realm than its mere structure. So the obvious question to ask is whether Newman's objection against Russell can be turned into an objection against ramsified realism.

Over the next few sections, we will try to formulate the strongest possible Newman-style objection against ramsified realism. It arises by combining two simple thoughts: one relating to the Push-Through Construction; the other relating to various model-theoretic notions of *conservation*.

3.6 The Newman-conservation-objection

We start by developing the notion of *conservation*, and proving that *Ramsey sentences are object-language statements of conservation*. This is essentially our Proposition 3.5, below. We then use this to present an interesting, but ultimately ineffectual, Newman-style objection against ramsified realism.

Suppose that S is a theory which has been formulated in a purely observational language, and that T is a wider theory of physics. Suppose, too, that we have a guarantee that no more observational consequences follow from T than follow from S alone. Obviously, the notions of 'guarantee' and 'following from' need to be made

[12] See Demopoulos and M. Friedman (1985), Newman (1928), and Russell (1927).

[13] Newman (1928: 144), with a slight change to typography.

[14] Newman (1928: 145–6).

[15] See Zahar and Worrall (2001: 238–9, 244–5), Worrall (2007: 150), and Ainsworth (2009: 143–4).

more precise. But if they are suitably rich, this guarantee will allow us to become instrumentalists (or fictionalists) about everything in T which goes beyond S. That is: we could in good conscience employ the full power of T, without having to insist that its distinctively theoretical claims are *true*; they would simply be an expedient way to allow us to reason our way around observables. (Note, in passing, that this line of thought seems to pull us away from ramsified realism, since ramsified realists wanted to incur some *substantive* theoretical commitments; this was the point of clause (b) of §3.3.)

To tidy up the line of thought that points us towards instrumentalism, we can employ the following definition:

Definition 3.3: *Let T be an $\mathscr{L}^+$-theory and S be an $\mathscr{L}$-theory, with $\mathscr{L}^+ \supseteq \mathscr{L}$. T is* consequence-conservative *over S iff: if $T \vDash \varphi$ then $S \vDash \varphi$, for all $\mathscr{L}$-sentences φ.*

The relation $T \vDash \varphi$ in this definition is the model-theoretic consequence relation from §1.12, defined by: if $\mathcal{M} \vDash T$ then $\mathcal{M} \vDash \varphi$, for all structures $\mathcal{M}$. The immediate interest in consequence-conservation is as follows. Suppose that S is the set of all *true* observation sentences, expressed in some suitably rich vocabulary. To say that T is consequence-conservative over S is, then, to say that T is perfectly 'observationally reliable'. After all, any observation sentence entailed by T is entailed by S, and hence true by assumption.

Here, though, is a second way to think about 'observational reliability'. Imagine that we can turn any model of S into a model of T, just by interpreting a few more symbols in T's vocabulary. Then any configuration of observational matters that makes S true is *compatible* with the truth of T. Intuitively, then, T cannot make any false judgements about observational matters; so, again, T is 'observationally reliable'.

Formalising these intuitive ideas yields a second notion of conservation. We already formalised the idea of interpreting new symbols, when introducing Robinsonian semantics (see Definition 1.4). So we can offer the following:

Definition 3.4: *Let T be an $\mathscr{L}^+$-theory and S be an $\mathscr{L}$-theory, with $\mathscr{L}^+ \supseteq \mathscr{L}$. T is* expansion-conservative *over S iff: for any $\mathscr{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \vDash S$, there is an $\mathscr{L}^+$-structure $\mathcal{N}$ which is a signature expansion of $\mathcal{M}$ and such that $\mathcal{N} \vDash T$.*

It is natural to ask how these two notions of conservation relate to each other. In fact, the Ramsey sentence allows us to demonstrate that they align perfectly:[16]

[16] We have not found Proposition 3.5 stated in the literature on Ramsey sentences, though we note three near misses. First: as just noted, Carnap's Proposition 3.2 is an immediate corollary of Proposition 3.5. Second: Worrall (2007: 151–2) suggests that Demopoulos and M. Friedman (1985) might want to invoke something like Proposition 3.5. Third: Ketland (2009: 42) obtains a restricted corollary of Proposition 3.5.

Proposition 3.5: *Let T be any finite $\mathscr{V}$-theory, and let S be any $\mathscr{V}_o$-theory. The following are equivalent:*

(1) T is consequence-conservative over S
(2) T is expansion-conservative over S
(3) $S \vDash \text{Ramsey}(T)$

Proof. *(3) ⇒ (2).* Suppose $S \vDash \text{Ramsey}(T)$. Let $\mathcal{M} \vDash S$. Since $\mathcal{M} \vDash \text{Ramsey}(T)$, we can choose suitable witnesses for $\text{Ramsey}(T)$'s initial existential quantifiers, and then take these witnesses as the interpretations of the $\mathscr{V}_t$-predicates, giving us a $\mathscr{V}$-structure $\mathcal{N} \vDash T$ which is a signature expansion of $\mathcal{M}$.

(2) ⇒ (1). Let φ be any $\mathscr{V}_o$-sentence such that $T \vDash \varphi$. Suppose $\mathcal{M} \vDash S$. As T is expansion-conservative over S, there is a $\mathscr{V}$-structure $\mathcal{N} \vDash T$ which is a signature expansion of $\mathcal{M}$. Since $\mathcal{N} \vDash T$, we have $\mathcal{N} \vDash \varphi$, and hence $\mathcal{M} \vDash \varphi$ since φ is a $\mathscr{V}_o$-sentence and $\mathcal{N}$ and $\mathcal{M}$ agree on the interpretation of the $\mathscr{V}_o$-vocabulary. Since $\mathcal{M}$ was arbitrary, $S \vDash \varphi$.

(1) ⇒ (3). $T \vDash \text{Ramsey}(T)$,[17] and $\text{Ramsey}(T)$ is a $\mathscr{V}_o$-sentence. So if T is consequence-conservative over S, then $S \vDash \text{Ramsey}(T)$. □

Informally glossed, this Proposition states that Ramsey sentences are object-language statements of conservation (in both senses). Additionally, Proposition 3.2—which told us that T and $\text{Ramsey}(T)$ are observationally equivalent—is an easy corollary of Proposition 3.5.[18] However, it should be emphasised that this Proposition crucially relies upon the use of *second-order* logic: not only does condition (3) involve Ramsey sentences, which are by definition second-order, but the equivalence between (1) and (2) *fails* for first-order theories (see §3.B).

This Proposition now allows us to formulate a neat (if ultimately ineffectual) Newman-style objection against ramsified realism:

Newman-conservation-objection. *Instrumentalists and realists can agree that their favourite theories should be $\mathscr{V}_o$-sound, i.e. that all of their $\mathscr{V}_o$-consequences should be true.*[19] *Where S is the set of all true $\mathscr{V}_o$-sentences, to say that T is $\mathscr{V}_o$-sound is to say that T is consequence-conservative over S. By Proposition 3.5, this is equivalent to the claim that $S \vDash$* $\text{Ramsey}(T)$*. So* $\text{Ramsey}(T)$ *expresses no truth-evaluable content beyond T's observational consequences. So clause (b) of §3.3—which states precisely that the content expressed by* $\text{Ramsey}(T)$ *goes* beyond *T's mere observational consequences—is false, and ramsified realism fails.*

[17] Assuming only that our structures satisfy a few instances of the Comprehension Schema (see §1.11).

[18] For the left-to-right direction of Proposition 3.2, just take $S = \text{Ramsey}(T)$, so that condition (3) of Proposition 3.5 is trivially satisfied, and then apply the equivalence between condition (3) and condition (1) guaranteed by Proposition 3.5. The right-to-left direction of Proposition 3.2 follows since $T \vDash \text{Ramsey}(T)$, as we had occasion to note in the proof of Proposition 3.5.

[19] Ketland (2009: Definition D) calls $\mathscr{V}_o$-soundness 'weak O-adequacy'.

3.7 Observation vocabulary versus observable objects

The ramsified realist has only one possible response to the Newman-conservation-objection: she must deny that instrumentalists are committed to the $\mathscr{V}_o$-soundness of physical theories.

As Worrall explains, this is much less strange than it initially sounds.[20] Suppose T entails the existence of an object which falls under *no* $\mathscr{V}_o$-predicate. If we believe that our $\mathscr{V}_o$-vocabulary is adequate to deal with any observable object we encounter, then we can characterise this by saying that T entails a $\mathscr{V}_o$-sentence which says, in effect, 'there is an unobservable object'. So, in order for T to be $\mathscr{V}_o$-sound, there must be unobservable objects. But instrumentalists need not be committed to the existence of any unobservables. So instrumentalists should not be regarded as being committed to the $\mathscr{V}_o$-soundness of physical theories after all.

To deal with this point, in what follows we will assume that our observational vocabulary $\mathscr{V}_o$ includes a primitive one-place predicate, O, which is intuitively to be read as '... is an observable object'.[21] In this new setting, we should not count just *any* $\mathscr{V}_o$-sentence as an *observation* sentence. Rather, the observation sentences will be just those sentences which, intuitively, are restricted to telling us *about* the observable objects. Specifically, the observation sentences are the $\mathscr{V}_o$-sentences whose quantifiers are restricted to O. Call these the $\mathscr{V}_o^O$-sentences. We define these precisely as follows.

Definition 3.6: *Where φ is any $\mathscr{V}_o$-formula, we recursively define:*

$$\begin{aligned}
\varphi^O &:= \varphi, \text{ if } \varphi \text{ is atomic}\\
(\varphi \wedge \psi)^O &:= (\varphi^O \wedge \psi^O)\\
(\neg\varphi)^O &:= \neg(\varphi^O)\\
(\exists x\varphi)^O &:= \exists x(O(x) \wedge \varphi^O)\\
(\exists X^n\varphi)^O &:= \exists X^n\left(\forall\bar{v}\,[X^n(\bar{v}) \to (O(v_1) \wedge \ldots \wedge O(v_n))] \wedge \varphi^O\right)
\end{aligned}$$

The set of $\mathscr{V}_o^O$-formulas is then $\{\varphi^O : \varphi$ is a $\mathscr{V}_o$-formula$\}$.

We can now easily accommodate Worrall's point. The sentence 'there are unobservable objects', i.e. $\exists x\neg O(x)$, uses an unrestricted existential quantifier, and so is a $\mathscr{V}_o$-sentence but not a $\mathscr{V}_o^O$-sentence. Moreover, in these terms, instrumentalists and realists will agree only on the fact that T should entail true $\mathscr{V}_o^O$-sentences, i.e. that theories should be $\mathscr{V}_o^O$-*sound*.[22] But it is clear that $\mathscr{V}_o^O$-soundness is strictly

[20] Worrall (2007: 152).

[21] This predicate allows us to simulate Ketland's (2004: 289ff, 2009: 38ff) use of a two-sorted language.

[22] This is essentially Ketland's (2004: Definition D) notion of 'weak empirical adequacy'.

weaker than $\mathscr{V}_o$-soundness, and it is only obvious that *realists* have any motivation for insisting on the latter. So the Newman-conservation-objection fails.

3.8 The Newman-cardinality-objection

Our aim, now, is to present a version of Newman's objection which can handle the fact that the quantifiers range over observables and unobservables alike. The essential idea behind the objection is quite simple. In concessive spirit, we will grant to the ramsified realist that there are no problems whatsoever with regard to the $\mathscr{V}_o^O$-claims. In particular, they have an *intended model*. But we will then ask the ramsified-realist what more it takes for her 'troublesome' claims to be true. Model theory threatens that she must answer, only, that there be *enough* unobservables.

To implement this strategy, let C be the *correct* model of the $\mathscr{V}_o^O$-sentences. Now, in allowing that there *is* such a model, we are simply waiving any Putnam-esque concerns about the reference of $\mathscr{V}_o$-expressions, such as those from §2.3. To do this, though, is just to allow the ramsified realist (for the sake of argument) that this is all *okay*.

Since C is the intended model of the $\mathscr{V}_o^O$*-sentences*, it is a $\mathscr{V}_o$-structure whose domain $C = O^C$. Since C is *correct*, C comprises all and only the actual observable entities in the physical universe, and any given entities stand in a certain observation relation to each other iff C represents them as so doing.

We now turn to the question of what it takes for T, or Ramsey(T), to be *true*. Since we have granted the ramsified realist the existence of an *intended model* of the $\mathscr{V}_o^O$-sentences, at the very least T or Ramsey(T) must be *compatible* with C. Somewhat more precisely, there must be a model of T whose 'observable part' is just C itself. But, to make this fully precise, we must explain the idea of a 'part' of a structure. We do this by introducing the notion of a substructure:

Definition 3.7: *Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathscr{L}$-structures, with $M \subseteq N$. Then we say that $\mathcal{M}$ is a* substructure *of $\mathcal{N}$ iff: for any $\mathscr{L}$-constant symbol c, any n-place $\mathscr{L}$-relation symbol R, and any n-place $\mathscr{L}$-function symbol f:*[23]

$$c^{\mathcal{M}} = c^{\mathcal{N}}$$
$$R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^n$$
$$f^{\mathcal{M}} = f^{\mathcal{N}}|_{M^n}$$

[23] The notation $g|_X$ indicates g's restriction to X, implemented set-theoretically as $g|_X := \{(\bar{x}, g(\bar{x})) \in g : \bar{x} \in X\}$. When we consider Henkin $\mathscr{L}$-structures, as in §1.6, the definition of *substructure* requires two further clauses (see Shapiro 1991: 92) for each $n < \omega$: (i) if $Q \in M_n^{\text{rel}}$ then there is some $R \in N_n^{\text{rel}}$ such that $Q = R \cap M^n$; and (ii) if $g \in M_n^{\text{fun}}$ then there is some $h \in N_n^{\text{fun}}$ such that $g = h|_{M^n}$.

Now, where $\mathcal{M}$ is a $\mathscr{V}_o$-structure, we say that *$\mathcal{M}$'s observable part,* $\mathrm{Ob}(\mathcal{M})$, is the $\mathscr{V}_o$-reduct of the substructure of $\mathcal{M}$ whose domain is the observables in $\mathcal{M}$, i.e. $O^{\mathcal{M}}$.[24] Since we have assumed that $\mathscr{V}$ only contains relation symbols, this amounts to the following condition: $\mathrm{Ob}(\mathcal{M})$'s domain is $O^{\mathcal{M}}$, and $R^{\mathrm{Ob}(\mathcal{M})} = R^{\mathcal{M}} \cap (O^{\mathcal{M}})^n$ for each n-place $\mathscr{V}_o$-predicate, R.[25]

We now say that T is *$\mathcal{C}$-alright* iff there is a $\mathscr{V}$-structure $\mathcal{M} \vDash T$ such that $\mathrm{Ob}(\mathcal{M}) = \mathcal{C}$.[26] So, intuitively, if T or $\mathrm{Ramsey}(T)$ is true then they must (at least) be $\mathcal{C}$-alright.

To run a Newman-style argument in this setting, we simply need to combine two ideas that we have encountered already. The first idea comes from the Push-Through Construction of §2.1, which was also the basis for Newman's original objection to Russell in §3.5. In the present setting, the important point is that *any* object will do as an 'unobservable'; so we can push-through the 'unobservables' as much as we like, so long as we leave the 'observables' undisturbed.[27]

The second idea relates to a trivial corollary of Proposition 3.5, that T is expansion-conservative over $\mathrm{Ramsey}(T)$. This lets us move freely between models of T and models of $\mathrm{Ramsey}(T)$, by signature-expansion and reduction.

Combining these two ideas, we obtain just the result we need:[28]

Proposition 3.8: *Let T be any $\mathscr{V}$-theory, with $O \in \mathscr{V}_o$. Let $\mathcal{A}$ be any $\mathscr{V}$-structure satisfying $\forall x O(x)$. Then the following are equivalent:*

(1) There is a $\mathscr{V}$-structure $\mathcal{M}$ such that $\mathcal{M} \vDash T$ and $\mathrm{Ob}(\mathcal{M}) = \mathcal{A}$

(2) There is a cardinal κ with the following property: if U is a set of cardinality κ with $A \cap U = \varnothing$, then there is a $\mathscr{V}_o$-structure $\mathcal{P}$ with domain $A \cup U$, with $\mathrm{Ob}(\mathcal{P}) = \mathcal{A}$ and with $\mathcal{P} \vDash \mathrm{Ramsey}(T)$

---

[24] When we consider Henkin $\mathscr{L}$-structures, we also need to stipulate that $\mathrm{Ob}(\mathcal{M})^{\mathrm{rel}}_n := \{X \cap (O^{\mathcal{M}})^n : X \in M^{\mathrm{rel}}_n\}$ and $\mathrm{Ob}(\mathcal{M})^{\mathrm{fun}}_n := \{g|_{(O^{\mathcal{M}})^n} : g \in M^{\mathrm{fun}}_n \text{ and } \mathrm{range}(g|_{(O^{\mathcal{M}})^n}) \subseteq O^{\mathcal{M}}\}$.

[25] For similar notions, see Przełęckie (1973: 287), Hodges (1993: 202), and Lutz (2014: §4.3).

[26] There is a delicacy here. So far, we have made no particular assumptions about how the observable *entities* interact with the observation *relations*. Two plausible assumptions are: (a) observable objects never fall under $\mathscr{V}_t$-predicates, or (b) unobservable objects never fall under $\mathscr{V}_o$-predicates. The legitimacy and consequences of such assumptions are touched upon in the literature, in particular in connection with predicates like 'x is bigger than y' which, as discussed in §3.4, resist categorisation as (dichotomously) either observational or theoretical. (See Cruse 2005: 561ff; Ainsworth 2009: 145–6, 155–60; Ketland 2004: 292, 2009: 38–9.) However, *provided* that some o/t dichotomy has been embraced, we think that we can sidestep this debate. A philosopher who rejects (b) might insist on expanding $\mathcal{C}$, so that it includes all unobservable entities which fall under some observation relation. If we concede this point, then we will want to expand the definition of $\mathrm{Ob}(\mathcal{M})$ accordingly. (It is harder to see how rejecting (a) would force us to change anything in our setup.) But making these changes this will not affect the fundamental idea behind Proposition 3.8: we can always arbitrarily permute away the entities falling outside $\mathrm{Ob}(\mathcal{M})$'s domain.

[27] This idea is in Winnie (1967: 226–7). Putnam (1981: 218, Second Comment) mentions it in his fullest presentation of his permutation argument. See also Button (2013: 41–2).

[28] Again, as is standard, we use κ to denote a cardinal; see the end of §1.B for a brief review of cardinals.

Proof. (1) ⇒ (2). Let $\mathcal{M}$ be as in (1), so $O^{\mathcal{M}} = O^{\mathcal{A}} = A$. Let κ be the cardinality of the unobservables in $\mathcal{M}$, i.e. $\kappa = |M \setminus A|$. Let U be any set of cardinality κ with $A \cap U = \varnothing$; so there is a bijection $g : (M \setminus A) \longrightarrow U$. Define a bijection $h : M \longrightarrow (A \cup U)$ by setting $h(x) = x$ if $x \in A$, and $h(x) = g(x)$ otherwise. Now use a Push-Through Construction, applying h to $\mathcal{M}$ to define a $\mathscr{V}$-structure, $\mathcal{N}$, and simply let $\mathcal{P}$ be $\mathcal{N}$'s $\mathscr{V}_o$-reduct.

(2) ⇒ (1). Let $\mathcal{P}$ be as in (2). As T is expansion-conservative over Ramsey(T), there is a $\mathscr{V}$-structure $\mathcal{M}$ which is a signature expansion of $\mathcal{P}$ and $\mathcal{M} \vDash T$. □

The significance of this result emerges, when we substitute the intended model C of the $\mathscr{V}_o^O$-sentences for $\mathcal{A}$, to obtain a new objection to ramsified realism:

Newman-cardinality-objection. *Instrumentalists and realists can agree that, at a minimum, T must be C-alright. But Proposition 3.8 entails that T is C-alright* iff *there is some cardinal κ such that we can obtain a model of* Ramsey(T) *just by adding κ 'unobservables' to C. Otherwise put:* Ramsey(T) *is* true *provided both that T is C-alright and there are 'sufficiently many' unobservables. And so, we are back at Newman's criticism: 'the doctrine that* only *structure is known', when it comes to the theoretical, 'involves the doctrine that* nothing *can be known' about the theoretical, except 'the number of' theoretical objects.*[29]

This is the most powerful version of the Newman objection, as generalised against ramsified realism.[30] Maybe some ramsified realists can learn to embrace the idea that the purpose of physics is to get everything right at the level of observables and additionally tell us the mere *cardinality* of the theoretical realm; but we doubt that this view will find many takers.

We can summarise the problem as follows. Proposition 3.8 uses model theory to generate a 'trivialising' structure, $\mathcal{P}$, from a given structure, $\mathcal{M}$. The existence of $\mathcal{P}$ is guaranteed by the standard set-theoretic axioms which we assume whenever we do model theory. So the ramsified-realist needs to explain why $\mathcal{P}$ is somehow *unintended* (cf. §2.3). To close the chapter, we consider two attempts to do this.[31]

3.9 Mixed-predicates again: the case of causation

When we discussed the permutation argument in §2.3, we noted that someone might complain that models obtained by the Push-Through Construction can fail

[29] Newman (1928: 144).

[30] Our *Newman-cardinality-objection* is close to Ketland's preferred version of the Newman objection (See also Ainsworth 2009: 144–7). Ketland's objection falls out of his Theorem 6 (2004: 298–9), which is close to our Proposition 3.8. However, Ketland invokes *full* second-order logic, whereas our Proposition 3.8 holds with both full and faithful Henkin semantics (see §3.2, and also footnote 1 from Chapter 2).

[31] For an excellent survey of responses, see Ainsworth (2009: 148ff).

to respect causal connections between words and the world. The ramsified realist might offer a similar response, concerning our trivialising model, $\mathcal{P}$.

By construction, $\mathcal{P}$ will respect all the causally-constrained reference relations between *okay* vocabulary and *okay* entities. The only possible concern, then, must be about the relationship between *troublesome* vocabulary and *troublesome* entities. We want to argue for a conditional claim: *if* our realist took the o/t dichotomy seriously in the first place, then she cannot make any hay here.

To see why, we must revisit §3.4. There, we considered the predicate 'x is bigger than y'. and noted that, prima facie, there are *observable* instances of this relation, *theoretical* instances of this relation, and *mixed* instances of this relation. The same point applies to the predicate 'x causes y'. Waiving Humean concerns, there are observational instances of causation: the striking of the match causes the fire. There are also theoretical instances of causation: consider the microscopic events involved in a chemical reaction. And then there are *mixed* instances: the decay of a radium atom causes the Geiger counter to click.

As in §3.4, such considerations might lead us to doubt the very idea that there is a sharp o/t *dichotomy*. And if we deny that there is any such dichotomy, then nothing will prevent us from going on to insist that the observable instances (in some rough and ready sense) of causation give us sufficient handle on the general notion of causation for us to see why $\mathcal{P}$ may be unintended.[32] However, to abandon the o/t dichotomy is *precisely* to abandon the position against which the Newman-cardinality-objection was targeted.

If, instead, we continue to insist on an o/t dichotomy, then we have no option but to 'split' the 'mixed' claims about causation into the okay ones and the troublesome ones. (Compare the fate of 'x is bigger than y' in §3.4.) But once we have done *that*—and taken seriously that this is a dichotomy and not a continuous transition from more to less observable—then it is profoundly unclear what would allow us to maintain that we have sufficient handle on the 'troublesome side' of causation, even to *articulate* the idea that $\mathcal{P}$ has failed to respect the (troublesome) causal relationship between troublesome vocabulary and troublesome entities. For the 'troublesome side' of causation is *just more $\mathcal{V}_t$-vocabulary*, to be ramsified away.[33]

3.10 Natural properties and just more theory

An alternative response is for the ramsified-realist to insist that some properties are more *natural* than others.

The existence of the trivialising structure, $\mathcal{P}$, is guaranteed by the standard set-theoretic axioms which we assume whenever we do model theory. The ramsified

[32] See the references in footnote 11.

[33] This is essentially the point of Button (2013: 50–1).

realist may, then, maintain that certain properties and relations are *natural*, whilst others are *mere artefacts* of a model theory that might be suitable for pure mathematics, but which is unsuitable for the philosophy of science. If she can uphold this point, then she can reject the significance of Proposition 3.8 and so answer the Newman-cardinality-objection.[34]

The obvious question is whether ramsified realism is consistent with a belief in such natural kinds. To believe that some properties or relations are natural is to postulate a higher-order property of properties, namely, *naturalness*. Syntactically, this will be introduced via a second-order predicate, N, such that we can say of a first-order predicate, R, such as '… is an electron', something of the form $\mathrm{N}(R)$, i.e. roughly '*electronhood* is a natural property'. Since, though, our ramsified realist believes in an o/t dichotomy, we must ask her whether N falls on the o-side or the t-side.[35] Certainly N cannot belong to the *observational* vocabulary—even if *electronhood* is natural, one cannot simply *observe* that it is—so N must belong with the *theoretical* vocabulary, i.e. $\mathrm{N} \in \mathscr{V}_{\mathrm{t}}$. As such, the ramsified realist is duty bound to ramsify away the higher-order property of *naturalness*. That is, she must replace N with an existentially bound (third-order) variable Y, so that her new Ramsey sentence for T is:

$$\exists \mathrm{Y} \exists X_1 \dots \exists X_n \hat{T}[\mathrm{Y}/\mathrm{N}, \overline{X}/\overline{R}]$$

But now, given even remotely permissive comprehension principles for third-order logic, the invocation of a *ramsified* property of *naturalness* will impose no constraint whatsoever upon our structures.[36]

The shape of this problem should look familiar: in the last few paragraphs, we have essentially recapitulated the dialectic surrounding Putnam's permutation argument from §2.3. Our ramsified realist wanted to maintain that a Ramsey sentence is made true (if at all) only by the existence of a structure whose second-order entities are *natural*. This is just a version of the idea that certain referential candidates are *preferable*, as considered in §2.3 in response to Putnam's permutation argument.[37] And, as in §2.3, any defender of this notion of *preferability* must face up to Putnam's just-more-theory manoeuvre. In the present context, the allegation will be that to invoke 'naturalness' is just more *theory*.

---

34 See Psillos (1999: 64–5), Ketland (2009: 44), and Ainsworth (2009: 167–9).

35 Brian King suggested to us that the ramsified realist might decline to answer the question, on the grounds that N is not itself an expression from the sciences, but an expression from some metasemantic theory. Whilst the reply is possible in principle, it is hard to see how it could be developed. A ramsified realist who offers this reply must explain why her initial motivations for embracing the o/t dichotomy stop short of rendering *naturalness* troublesome. This is particularly difficult if those initial motivations were essentially epistemological, for our epistemological situation with regard to *naturalness* is worse than our epistemological situation with regard to *electronhood* (for example).

36 For a similar argument, see Ainsworth (2009: 160–2, 169); also Demopoulos and M. Friedman (1985: 629), Psillos (1999: 64–6), and Ketland (2009: 44).

37 See Ainsworth (2009: 162–3).

Now, in §2.3, we noted that there are positions according to which the just-more-theory manoeuvre simply looks hopelessly question-begging. (If causation fixes reference, then permuting the reference of 'causation' is besides the point.) However, we also noted that the just-more-theory manoeuvre poses real difficulties for certain other positions, such as moderate objects-platonism. (To the moderate objects-platonist, anything compatible with moderation looks like more mere theory.) The same point arises here. Belief in intrinsically natural kinds may be a defensible philosophical thesis. However, the present allegation is that the ramsified realist is bound, *by her own lights*, to treat 'naturalness' as just more theory, in the precise sense that it belongs to theoretical-vocabulary rather than the observational-vocabulary, and hence *must* be ramsified away.

At this point, though, we see that there are no *shortcuts* to the philosophical significance of the Newman-cardinality-objection against ramsified realism. Its significance can only be settled by delving deeply into the very idea of an o/t dichotomy. And *that* would take us far away from model theory, and so far beyond the scope of this book.

So, to close the chapter, we simply note the deep connections between the permutation argument and the Newman objection. First: ramsified realism itself can be motivated by permutation-style worries about the reference of theoretical vocabulary (see §3.2). Second: the Push-Through Construction is used in Putnam's permutation argument, in Newman's original argument, and in the Newman-cardinality-objection.[38] Third: similar responses are available to both arguments, and such responses must confront the just-more-theory manoeuvre head-on.[39]

3.A Newman and elementary extensions

We have presented and discussed our favourite version of the Newman objection. However, in this appendix, we consider a Newman-esque objection in a first-order context. Before starting, we should issue a caution. The appendix draws on technical results which we will first encounter in Chapter 4. As such, some readers may want to read ahead before returning to this appendix. And the material in this appendix can, indeed, be safely omitted, since the *best* version of the Newman objection is the Newman-cardinality-objection of §3.8.

In the next chapter, we prove the following:[40]

---

[38] See Demopoulos and M. Friedman (1985: 629–30), Lewis (1984: 224fn.9), Ketland (2004: 294–5, 298–9), and Hodesdon (forthcoming: §5).

[39] See Hodesdon (forthcoming: §6).

[40] Actually, in Chapter 4 we state the result in terms of deduction-conservativeness (see Definition 4.17) rather that consequence-conservativeness. However, for first-order theories, these notions are identical, by the Soundness and Completeness of first-order logic.

Proposition (4.18): *Let T be a first-order $\mathscr{L}^+$-theory and S be a first-order $\mathscr{L}$-theory, with $\mathscr{L}^+ \supseteq \mathscr{L}$. The following are equivalent:*

(1) T is consequence-conservative over S.

(2) For any $\mathscr{L}$-structure $\mathcal{M} \vDash S$, there is an $\mathscr{L}^+$-structure $\mathcal{N}$ which satisfies T and whose $\mathscr{L}$-reduct is an elementary extension of $\mathcal{M}$.

This result connects a notion of conservation (see Definition 3.3) with the existence of an elementary extension (see Definition 4.3), i.e. typically to a structure with an enlarged *ontology*. It is worth explicitly contrasting this first-order result with our second-order Proposition 3.5, which connected a notion of conservation with a structure with no new ontology, but merely a richer *ideology*.

Despite the differences, Demopoulos has recently used Proposition 4.18 to present a Newman-esque objection.[41] His objection does not consider Ramsey sentences, or target ramsified realism, *per se*. Rather, Demopoulos's target is a realist who believes that physical theories have a kind of content which goes beyond the observational claims, but which is in some sense 'merely structural'. Demopoulos criticises this realist via something very much like the Newman-conservation-objection of §3.6, but employing the first-order Proposition 4.18 rather than the second-order Proposition 3.5. As we understand it, Demopoulos' objection runs as follows:

Newman-extension-objection. *As in the Newman-conservation-objection of §3.6, instrumentalists and realists of all stripes can surely agree that their theory, T, should be consequence-conservative over the set of all true $\mathscr{V}_o$-sentences, S. Additionally, instrumentalists and realists of all stripes can surely agree that there is a model, say $\mathcal{C}$, of S. So, by (1) $\Rightarrow$ (2) of Proposition 4.18, there is a model of T (indeed, an elementary extension of $\mathcal{C}$). So the realist does not, in fact, manage to incur any substantial commitments by affirming T.*

Now, like the Newman-conservation-objection of §3.6, this Objection invokes the idea that the instrumentalist must be committed to $\mathscr{V}_o$-soundness. However, as we pointed out in §3.7, that is a mistake. If there is a one-place $\mathscr{V}_o$-predicate, O, to be read as '… is an observable object', then $\exists x \neg O(x)$ is a $\mathscr{V}_o$-sentence which is probably entailed by T, but which the instrumentalist need not accept. Consequently, instrumentalists need not be committed to the $\mathscr{V}_o$-soundness of physical theories after all. Exactly the same point applies to the Newman-extension-objection.

However, it is worth pointing out a *further* defect with the Newman-extension-objection. The Objection involves the claim that there is an intended model, $\mathcal{C}$, of the true $\mathscr{V}_o$-sentences. When we apply Proposition 4.18 to $\mathcal{C}$, we obtain a structure,

[41] Demopoulos (2011: 186–90).

$\mathcal{N}$, whose $\mathscr{V}_o$-reduct is an elementary extension of $\mathcal{C}$. Crucially, this elementary extension will (probably) *enlarge* the ontology; that is, $N \setminus C$ will be non-empty. And if the objection is to hit its target, then we should think of C as all that is *observable*, so that the entities in $N \setminus C$ are all deemed *theoretical*.[42] But Proposition 4.18 provides no guarantee that we might not find some new element $a \in N \setminus C$ such that $\mathcal{N} \vDash O(a)$. On the one hand, such an element would be *observable*, at least according to $\mathcal{N}$; on the other hand, it should be(merely) *theoretical*, since it is in $N \setminus C$. It is unclear why such a confused structure should command our attention.

We can illustrate this concern via a toy example, which employs a lovely result from the model theory of PA. The result is that PA cannot define the notion of a 'standard number' (the notation $S^n(0)$ was defined in (*numerals*) of §1.13):

Lemma 3.9 (Overspill)**:** *Let $\varphi(v)$ be any one-place formula in the signature of* PA*. Let $\mathcal{M}$ be a non-standard model of* PA*. If $\mathcal{M} \vDash \varphi(S^n(0))$ for all $n < \omega$, then there is some non-standard element b such that $\mathcal{M} \vDash \varphi(b)$.*

Proof. Suppose $\mathcal{M} \vDash \varphi(S^n(0))$ for all n in the metatheory. Let a be a non-standard element of $\mathcal{M}$. If $\mathcal{M} \vDash \varphi(a)$, then we are done. Otherwise, $\mathcal{M} \vDash \neg\varphi(a)$. Since $\mathcal{M} \vDash$ PA, it satisfies the least-number-principle,[43] so let a_0 be the least element of $\mathcal{M}$ satisfying $\neg\varphi(x)$. Since a_0 is non-standard, $a_0 \neq 0$, so there is some b such that $\mathcal{M} \vDash S(b) = a_0$. Then b is also non-standard and since $\mathcal{M} \vDash b < a_0$, we have that $\mathcal{M} \vDash \varphi(a)$. $\square$

To see the significance of this result, suppose that the observable objects happen to 'line up' nicely, so that they can be enumerated. So we can treat the (standard) natural numbers as *proxies* for the genuinely observable objects, and regard the (standard) natural number structure as a proxy for our intended model of the observables, $\mathcal{C}$. Then when $\mathcal{N}$ is a *proper* elementary extension of $\mathcal{C}$, it must add some new elements to the domain. In our toy, this will amount to adding some *nonstandard* numbers. Suppose, furthermore, that we have defined a one-place predicate 'O' in the language of PA, which we can usefully treat as a proxy for the predicate '… is an observable' in $\mathcal{C}$. Then, by the Overspill Lemma 3.9, some of these new nonstandard elements in $N \setminus C$ must satisfy the predicate 'O' in $\mathcal{N}$.

In sum, the Newman-extension-objection mishandles observable entities in at least two ways. This is why we favour the Newman-cardinality-objection.

[42] Hence Demopoulos writes: 'Let $\mathcal{C}$ be the $\mathscr{V}_o$-structure whose domain is the domain of observable events…. We must show that the domain C of $\mathcal{C}$ has an extension $\mathcal{N}$ which is the domain of a model of T, where N is the set of observable and theoretical events' (Demopoulos 2011: 186–7, notation changed to match main text).

[43] Roughly: if some elements are φs, then there is a <-least φ. See e.g. Kaye (1991: 44–5).

3.B Conservation in first-order theories

Theorem 3.5 shows that consequence-conservation and expansion-conservation align for second-order theories. In this appendix, we prove that this alignment *fails* for first-order theories. In particular, while expansion-conservation still implies consequence-conservation for first-order theories, the converse fails. This is precisely the content of Proposition 3.11, below.

As with §3.A: the topic of this appendix means that it belongs in this chapter, even though the result requires some ideas from later chapters. In particular, our proof of Proposition 3.11 uses both non-standard models of PA (see Chapter 4) and the Löwenheim–Skolem Theorem 7.2 (see Chapter 7). Still, our proof requires only basic model-theoretic tools.[44]

We say that a theory T is *complete* iff either $T \vDash \varphi$ or $T \vDash \neg\varphi$ for every sentence φ in T's signature. Then the theory of arithmetic in the signature $\{0, S\}$ is complete. More precisely, let SA, for *successor arithmetic*, be the theory whose axioms are just (Q1) and (Q2) from Definition 1.9, i.e., $\forall x\, S(x) \neq 0$ and $\forall x \forall y (S(x) = S(y) \rightarrow x = y)$. Now:

Proposition 3.10: SA *is complete.*

Proof. Where $\mathcal{M}$ is a model of SA, say that a *chain* in $\mathcal{M}$ is a subset $Z \subseteq M$ such that for every $a, b \in Z$ there is some natural number n in the metatheory such that either $(S^n(a))^{\mathcal{M}} = b$ or $(S^n(b))^{\mathcal{M}} = a$, using the (*numerals*) notation from §1.13.

Let $\mathcal{M}_0$ and $\mathcal{M}_1$ be two models of SA of a fixed uncountable cardinality, κ. Since they are of the same cardinality κ, their domains can be written as the disjoint unions $M_i = N_i \cup \bigcup_{\alpha<\kappa} Z_{i,\alpha}$ where the N_i component is isomorphic to the standard natural numbers in the signature $\{0, S\}$, and each $Z_{i,\alpha}$ component is isomorphic to the integers in the signature $\{S\}$. By mapping N_0 to N_1 and $Z_{0,\alpha}$ to $Z_{1,\alpha}$, we have that $\mathcal{M}_0 \cong \mathcal{M}_1$. Since any incompleteness in SA would have to be registered on models of size κ by the Löwenheim–Skolem Theorem 7.2, SA is complete. □

We can now use this result to offer an example of consequence-conservation without expansion-conservation.

Proposition 3.11: PA *is consequence-conservative over* SA *but not expansion-conservative over* SA.

Proof. First observe that that PA's signature strictly expands SA's.

[44] As such, our proof is much simpler than the one usually offered in the literature, which requires familiarity with formal theories of truth. See Halbach (2011: Theorem 8.12 p.80, Theorem 8.31 p.98).

Consequence-conservation. Let φ be a first-order sentence in the signature $\{0, S\}$ such that PA $\vDash \varphi$. If SA $\vDash \neg\varphi$, then also PA $\vDash \neg\varphi$, since PA has all of SA's axioms; but then, absurdly, PA would be inconsistent; so SA $\nvDash \neg\varphi$. Since SA is complete, SA $\vDash \varphi$, as required.

Failure of expansion-conservation. Let $\mathcal{M}$ be a model of SA given by a copy of the natural numbers followed by a single disjoint copy of the integers, with zero and successor interpreted in the ordinary way. So we may write $\mathcal{M}$'s domain as $M = N \cup Z$, where the N component is isomorphic to the standard natural numbers in the signature $\{0, S\}$, and the Z component is a chain.

For reductio, suppose there is a model $\mathcal{M}^*$ of PA whose $\mathscr{L}$-reduct is just $\mathcal{M}$. Let a be an element of Z. Now, PA proves that every number is even or odd. Begin by supposing that $\mathcal{M}^* \vDash b + b = a$ for some b in M. Then b too is in Z rather than N. Since a and b are part of the same chain, $\mathcal{M} \vDash S^k(b) = a$ for some $k \in \omega$. So $\mathcal{M}^* \vDash S^k(0) + b = S^k(b) = a = b + b$. Subtracting b from both sides, $\mathcal{M} \vDash S^k(0) = b$, contradicting the fact that b is in Z rather than N. And a similar contradiction arises from supposing $\mathcal{M}^* \vDash b + b + 1 = a$ for some b in M, completing the reductio. □