Quine on Explication*

Jonas Raab

June 29, 2021

Abstract. The main goal of this paper is to work out Quine's account of explication. Quine does not provide a general account, but considers a paradigmatic example which does not fit other examples he claims to be explications. Besides working out Quine's account of explication and explaining this tension, I show how it connects to other notions such as paraphrase and ontological commitment. Furthermore, I relate Quinean explication to Carnap's conception and argue that Quinean explication is much narrower because its main purpose is to be a criterion of theory choice.
Keywords: Quine, Explication, Carnap, Paraphrase

Contents

---

*As I'd often like to know this myself, I thought I provide a short history of this paper. I first wrote it in May 2020 and first submitted it July 2020. It was rejected by 7 journals (Phil Studies (desk rejection), Canadian Journal of Philosophy (desk rejection), Pacific Philosophical Quarterly (unclear; submitted August, rejected December, yet no comments of any sort; my asking has been ignored), PPR (no comments), Noûs (no comments), Philosophical Quarterly (desk rejection), and Australasian Journal of Philosophy (desk rejection with a nice explanation as to why and that this paper is a borderline case)) before Inquiry accepted it after an R&R.

1 Introduction

In recent years, philosophy has, once again, become conscious of its own methods. One prominent question is whether our conceptual apparatus is adequate or needs improvement. This investigation is often called *conceptual engineering* (e.g. by Cappelen 2018) or *conceptual ethics* (e.g. by Burgess & Plunkett 2013). Carnap's notion of *explication* is regularly cited as one root for such endeavours (e.g. Cappelen & Plunkett 2020, 6). A lot of research has focused on Carnap's approach. But Quine also proposes an account—which is the focus of this paper.

Given the amount of literature dedicated to Carnapian explication, there is surprisingly little literature on Quine's conception (a notable exception is Gustafsson 2006, 2014). Usually, it is merely mentioned that Quine's conception is different/not different from Carnap's, but not investigated in any detail. Yet, some philosophers explicitly refer to Quine's account (e.g. Haslanger 2012, 367, n. 1) which necessitates a better understanding.

This paper is structured as follows. Section 2 provides a general understanding of explication. Section 2.1 sketches Carnap's account and distinguishes between two notions of explication. Section 2.2 considers how Quine relates his account to Carnap's, introduces examples of explication, and suggests that Quine's examples don't fit his general account. The rest of this paper also works towards resolving this puzzle.

Sections 3–5 contain Quine's account of explication. Section 3 introduces Quine's paradigmatic account, Section 4 discusses Quine's claim that explication is elimination, and Section 5 relates explication to paraphrase and provides adequacy criteria.

Sections 6 and 7 consider examples of explication and examples which aren't explications. This confirms our account, and suggests an answer to the puzzle (Section 7.3).

Section 8 places explication within Quine's argument for a set theory.

Section 9 concludes the paper by briefly comparing Quine's and Carnap's accounts.

2 Preliminaries on Explication

Let us start by introducing explication more generally. In Section 2.1, we take a brief look at Carnap's notions of explication, introduce some terminology, and consider in what respects these notions differ from one another. In Section 2.2, we consider how Quine relates his to Carnap's notion.

2.1 Carnap on Explication

Carnap's understanding of 'explication' evolved over his career and can be understood as development of his 'rational reconstruction' (Beaney

3

2004, §4; cf. Carnap 1928/98, XVII). His most extensive discussion can be found in his 'Logical Foundations of Probability' (1962, first published 1950; henceforth LFoP), but we can take his replies (1963b) in the Schilpp volume (1963) as amendments to his earlier account (Brun 2016, 1213). Carnap also uses explication in 'Meaning and Necessity' (1956, first published 1947; henceforth M&N) without providing a detailed account.

Let us briefly consider Carnap's (1962) understanding of explication. The main idea is the following.

> There is a certain term [. . .] which is used in everyday language and by scientists without being exactly defined, and we try to make the use of these terms more precise or, as we shall say, to give an explication for them. (1962, 2)

As such, Carnap's account starts with *terms* which are supposed to be made more precise; he goes over to speak of *concepts*, though:

> The task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. (1962, 3)

However, his view is not bound to concepts insofar as he also accepts to explicate terms which are used to express these concepts (see 1962, 3). I continue to speak of concepts in the following.

Carnap introduces the following terminology. He calls the concept that is to be replaced *explicandum* whereas the replacing concept is called *explicatum*. Quine calls the explicatum 'explicans' (1960, 259), but I stick to 'explicatum' (without changing quotations).

Carnap's M&N understanding of explication is, so far, virtually identical to the above:

> The task of making more exact a vague or not quite exact concept used in everyday life or in an earlier stage of scientific or logical development, or rather of replacing it by a newly constructed, more exact concept, belongs among the most important tasks of logical analysis and logical construction. We call this the task of explicating, or of giving an explication for, the earlier concept[.] (1956, 7f.)

There are two main differences between the accounts in M&N and LFoP. The first is LFoP's explicit demand to provide *interpretations*:

> for a genuine explication [. . .] an interpretation is essential. (1962, 16)

Carnap distinguishes two different steps when introducing concepts: 'formalization and interpretation' (1962, 15). For our purposes, formalization is irrelevant; Carnap's example is axiomatization (1962, 15f.).

The interpretation, on the other hand, provides the formalised theory with meaning. This can be accomplished in different ways, for example by a definition (1962, 16).

Carnap's reason to insist on an interpretation is that we want to be able to use the explicatum in scientific contexts—we want to apply the explicatum to the world, i.e., we need the explicatum 'in the description of facts' (1962, 17). The example Carnap discusses is the explication of natural numbers. He points out that within a purely formal system, we can account for 'formulas like "$3 + 2 = 5$"' (1962, 17), yet we are not provided with an understanding of a numeral 'when it occurs in a factual sentence' (1962, 18) such as in the sentence 'the number of my fingers is 10'.

Carnap summarises this point as follows.

> As soon as we go over from the field of formal mathematics to that of knowledge about the facts of nature, in other words, to empirical science, which includes applied mathematics, we need more than a mere calculus or axiom system; an interpretation must be added to the system. (1962, 18)

In other words, we need our explicatum to be *applicable* to the world.

One salient reason for the absence of an analogous requirement in M&N is that Carnap provides a semantics for modal notions so that he automatically provides interpretations when explicating certain notions in semantic terms; there is simply no need in the context of M&N to insist on interpretations.

The second main difference is that Carnap (1962, 7) introduces general *adequacy criteria* for explications. The criteria are (i) similarity, (ii) exactness, (iii) fruitfulness, and (iv) simplicity. For our purposes, we don't need to go into more detail.

I don't want to speculate whether Carnap considered his M&N- and LFoP-notions to be the same. In this context, we might also wonder whether the formulation of particular conventions (e.g., Convention 2-1 in 1956, 10) which is prominent in explications in M&N is preserved in the LFoP-explication. Of course, in the M&N description Carnap does not insist on the formulation of conventions, but many of his explications happen to be accompanied by such.[1] Carnap's reason seems to be that he only wants to make use of requirements stated in the conventions (see 1956, 8), but not of any particular technical definition. Likewise, in LFoP, Carnap does not rule out the explicit formulation of conventions, but he doesn't make it a requirement either. However, even in LFoP, he

---

[1] One reason might be that Carnap underwent his semantic turn after Tarski taught him how to explicate 'truth' semantically (Carnap 1963a, §10; see also Leitgeb & Carus 2021, §6.1 and Supplement F); Carnap might have modelled M&N-explication along the lines of Tarski's explication of truth: Tarski proposes his Convention T (1933, 187f.) which any adequate definition of truth has to satisfy (see 1933, 187); Carnap also refers to Tarski's definition of 'true' (see his example 2 in 1962, 5).

formulates 'conventions of adequacy' (e.g., C53-1 at 1962, 285). Since explications are preceded by clarifying the explicanda (1962, 4), we can understand such explicit conventions to belong to the clarification process.

2.2 Quine on His and Carnap's Conceptions

Let us consider how Quine relates his to Carnap's account. After sketching it, Quine says the following.

> Philosophical analysis, explication, has not always been seen in this way. (1960, §53, p. 259)

To which he attaches the following footnote.

> By Carnap, yes; see *Meaning and Necessity*, pp. 7f. (1960, 259, n. 4)

Quine refers to the second edition of M&N. Interestingly, Quine does not mention (or list in the bibliography, see 1960, 278) Carnap's LFoP. In particular, Carnap's M&N-notion of explication is, as outlined in Section 2.1, presumably different from the one developed in LFoP. I don't want to speculate whether or not Quine knew about Carnap's treatment of explication in LFoP and, supposing that he did, whether he made a conscious decision to only refer to M&N.[2] However, Quine's conception of explication is similar to Carnap's M&N conception. Given that Quine's examples of explication don't fit his own paradigmatic account, one can speculate that he was thinking of the M&N conception.

Nevertheless, Quine, in the paragraph leading to the above quotation, explains his understanding of what an analysis or explication is, and he goes on to refer to M&N, claiming that Carnap had the same understanding. We can take from this that Quine seems to think that his and Carnap's accounts agree at least on the spirit of explication.[3]

Besides the already cited footnote, Carnap is explicitly mentioned only once more in §53, viz., in another footnote attached to a discussion of the so-called paradox of analysis. The point Quine is making is that the paradox of analysis arises from 'the reading of a synonymy claim into analysis' (1960, 259) and refers in the footnote to Carnap's M&N-discussion of it.

This is all Quine does to relate his to Carnap's account, and the first cited footnote is the reason for interpreters to understand Quine as giving the same account as Carnap. However, all we can infer from it is that

---

[2]We can assume, though, that Quine knew LFoP; he refers to its first edition (1951, 23, n. 4).

[3]Gustafsson's interpretation (of the footnote) is that Quine 'suggests that with regard to explication, he and Carnap endorse basically the same methodological viewpoint' (2014, 508).

Quine takes Carnap to have a similar attitude towards analysis and explication: the explicatum is allowed to be different from the explicandum as long as it is capable of fulfilling the particular purpose which made it worth troubling about (1960, 258f.). In Carnap's words:

> Generally speaking, it is not required that an explicatum have, as nearly as possible, the same meaning as the explicandum; it should, however, correspond to the explicandum in such a way that it can be used instead of the latter. (1956, 8)

This alone is the agreement Quine diagnoses between Carnap's and his own understanding without any suggestion of further agreement. Thus, for the moment, we can leave it open how closely matched Quine saw Carnap's M&N-account and his own.

   This brings us to the examples Quine lists as instances of explication. They are the following (1960, 260f.):

(EXAMPLE 1) the ordered pair,

(EXAMPLE 2) singular descriptions,[4]

(EXAMPLE 3) indicative conditionals,

(EXAMPLE 4) quantifiers.

The context of the discussion is Quine's claim that explication helps us to dissolve particular 'philosophical' problems. Quine does not explicitly call (EXAMPLE 2)–(EXAMPLE 4) explications, but his discussion of explication starts with (EXAMPLE 1) before he calls it 'paradigmatic of what we are most typically up to when in a philosophical spirit we offer an "analysis" or "explication"' (1960, 258), and the sentence leading to the text containing the above list of examples is the following.

> But what it illustrates as to the nature of explication applies very widely. (1960, 260)

Hence, Quine takes these to be examples of successful explication.

   However, Quine does not discuss (EXAMPLE 2)–(EXAMPLE 4); he only presents his treatment of 'ordered pair'. In particular, he calls, as I just quoted above, his construction of 'ordered pair' 'paradigmatic' (1960, 258). I take it that this means that we can and have to generalise on this particular 'construction' (1960, 258). Quine is also explicit:

---

[4]Gustafsson claims to have shown 'that at least some such instances of paraphrasing (such as Russell's way with singular descriptions) are counted as explications by Quine' (2014, 523). Even though I can't find an explicit argument, he at least recognises that Quine claims that his treatment of singular descriptions is an instance of explication. Gustafsson also recognises the tension (2014, 518) to the paradigmatic account to explicate *nouns* (2014, 522), but still claims that singular descriptions are an explication example.

> Our example is atypical in just one respect: the demands of partial
> agreement are preternaturally succinct and explicit (1960, 259)

where he sums up these demands in (what I call) a *purpose postulate*
(viz., (OpPp); see Section 3.3). Thus, we might not always expect to
get such a succinct and explicit postulate, but the treatment is otherwise
the same in other cases of explication. Moreover, Quine also tells us the
following.

> The ordered pair has had illustrative value because of the crispness
> of [its purpose postulate] and because of the multiplicity and the
> conspicuous artificiality of the explications. But what it illustrates
> as to the nature of explication applies very widely. (1960, 260)

Directly following is the above list of examples. Indeed, the examples are
supposed to show how widely applicable explication is. This confirms
that (i) the examples are indeed examples of successful explication and
(ii) the account of 'ordered pair' generalises.

What is astonishing about this is that (EXAMPLE 2)–(EXAMPLE 4)
do *not* fit the account of (EXAMPLE 1). One reason is quite simply that
Quine explicates what he calls 'defective nouns' (1960, 257), but none of
(EXAMPLE 2)–(EXAMPLE 4) is even a *noun*.

3 The Paradigmatic Account

This brings us to the details of Quine's account as developed in §53
of 'Word and Object' (henceforth W&O). In Section 3.1, I provide a
brief sketch of Quine's account so that the following discussion has a
clear target. Section 3.2 gives an account of what defective nouns—the
explicanda—are, and Section 3.3 presents the paradigmatic construction
of 'ordered pair'.

3.1 A Sketch of Quinean Explication

As presented in §53 of W&O, *Quinean explication* (or simply *explication*)
relates (particular) natural language expressions and expressions
of a theory formulated in canonical notation.[5] What the explication is
supposed to do is to find adequate surrogates for what Quine calls 'defective
nouns' (Section 3.2);[6] and such surrogates depend on the theory
into which we explicate—explication is *theory relative*. However, not every
defective noun needs to be explicated, but only those which have a

---

[5] For the rest of this paper, 'explication' without qualification means *Quinean explication*. Thanks to an anonymous referee for making me be explicit here.

[6] This contrasts Quinean and Carnapian explication. Carnapian explication is *iterative*, i.e., some explicata might, in turn, serve as explicanda. This is impossible for Quinean explication, because his explicata are formulated in canonical notation and, therefore, are neither expressions of natural language nor defective.

useful purpose. The particular purposes of a defective noun are collected in *purpose postulates* which are adequacy criteria for the explicata of the respective defective nouns. This means that explication is, just like Carnapian explication, a *two step process*. In the first step, we find purpose postulates; this step corresponds to Carnap's *clarification* step. The second step consists in finding an appropriate surrogate for a particular theory.

Therefore, a theory T successfully explicates a defective noun N with purpose postulates $PP_1[N]$, …, $PP_n[N]$ of N ($n \in \mathbb{N}$) iff there is a T-surrogate S for N (i.e., $N(x) :\Leftrightarrow S(x)$ where $S \in \mathcal{L}_T$) such that $T \vdash \exists x S(x)$ so that $T \vdash \bigwedge_{i=1}^{n} PP_i[N/S]$ (where $PP_j[N/S]$ is the result of substituting S everywhere for N in PP_j). This means that a successful explication comes with ontological commitments, but the ontological commitments are *not new* ones (cf. Gustafsson 2006, 62), but we find part of the ontology of T capable of fulfilling the purpose postulates. Thus, even though $T \vdash \exists x S(x)$, explication does not come with additional commitments but just provides a T-surrogate for N; the role of the ontological commitment is to guarantee that the purpose postulates are *non-trivially* fulfilled.

3.2 Defective Nouns, Identity, and Ontological Commitment

As sketched in Section 3.1, the starting point of an explication is a *defective noun*. The purpose of this section is to clarify what defective nouns are, and, in particular, why they are defective.

The opening lines of W&O's §53 are as follows.

> A pattern repeatedly illustrated in recent sections is that of the defective noun that proves undeserving of objects and is dismissed as an irreferential fragment of a few containing phrases. But sometimes a defective noun fares oppositely: its utility is found to turn on the admission of denoted objects as values of the variables of quantification. In such a case our job is to devise interpretations for it in the term positions where, in its defectiveness, it had not used to occur. (1960, 257)

The sections leading to §53 are entitled '*Entia non grata*' (§50), 'Limit myths' (§51), and 'Geometrical objects' (§52), and are what Quine refers to with 'recent sections'. In these sections, he dispenses with several nouns; we come back this.

The quotation indicates a reason to dispense with these defective nouns, viz., they are 'undeserving of objects' and 'irreferential'. What Quine is getting at is that these nouns appear as if they come with ontological commitment, but not in every occurrence. Indeed, the phrase 'values of the variables of quantification' builds a strong connection to *ontological commitment*.

This also points us into the direction of an adequate understanding of why certain nouns are defective. One example is Quine's paradigmatic example, i.e., *ordered pair*; it is the first example of a *useful* defective noun. The previous sections, on the other hand, contain examples of defective nouns not worth explicating. Quine explicitly calls some nouns *defective* in §50:

> Units of measure turn out somewhat like sakes and behalves. 'Mile', 'minute', 'degree Fahrenheit', and the like resemble 'sake' and 'behalf' in being *defective* nouns: they are normally used only in a limited selection of the usual term positions. (1960, 244, his emphasis)

Thus, 'mile', 'minute', 'degree Fahrenheit', 'sake', and 'behalf' are defective nouns; further examples are 'possible concrete objects' and 'unactualized possibles' (1960, 245). Quine also provides a reason for their defectiveness: 'there is perplexity over identity' (1960, 245), i.e., defective nouns don't have proper identity conditions or, as Quine also puts it, the source of their defectiveness is 'the lack of standard of identity' (1960, 244). This can easily be seen by exposing them 'in absurd interrogation' (1960, 244) by asking, for example, whether different miles are alike or how many sakes there are.

The main problem is that they sometimes occur in 'the usual term positions', i.e., they sometimes appear as though they had entities corresponding to them. Indeed, it is the term position that gives rise to *ontological commitment*. However, Quine demands of any entity that it has identity conditions; in slogan form, this becomes his infamous 'no entity without identity' (1957/58, 23). In particular, then, such defective nouns sometimes behave as if they were terms, i.e., ontologically committing, but sometimes they don't. Quine's only explanation of their defectiveness is that 'they are normally used only in a limited selection of the usual term positions'. His general strategy to deal with such terms is to *paraphrase* them (see Section 5).

As this is important in the following, let us briefly consider Quinean ontological commitment by just stating relevant features. Most importantly, it is *exclusively* sentences formulated in *canonical notation*, i.e., a first-order language (with identity),[7] which give rise to ontological commitments (e.g., 1948, 31f.), and the ontologically committing mechanism is quantification (e.g., 1953, 105, 1968, 97). In particular, it is the bound variables which have to take values, and the values over which these variables range is the universe consisting of the objects which have to figure as values. Indeed, Quine explicitly says that it is 'the object designated by [. . .] a term that counts as a value of the variable' (1960, §40, p. 192, n. 1). As it is the quantifiers that bind variables, and it is the existential quantifier that positively asserts the existence of something, a theory T

---

[7]See Hylton (2007, ch. 10, §I) for a more detailed description.

is ontologically committed to φs iff $T \vdash \exists x \varphi$, i.e., iff T implies that φs exist.[8] This understanding suffices for our purposes.

3.3 The Ordered Pair

Let us consider, then, why some of the defective nouns can (or even: need to) be explicated.

As mentioned before, the example Quine discusses is 'ordered pair' (EXAMPLE 1).[9] As the beginning of §53 (first quotation of Section 3.2) makes clear, some nouns are useful—indeed, they are so useful as to admit corresponding objects. As we have just seen, that means that these nouns have to be connected to ontological commitment so that it becomes necessary to admit corresponding objects.

Before following up on ontological commitment, let us consider why some defective nouns are useful. Why is 'ordered pair' a useful notion? Quine identifies its utility as follows.

> A typical use of the device is in assimilating relations to classes,
> by taking them as classes of ordered pairs. (1960, 257)

That is, we can use ordered pairs as device to 'assimilate relations to classes', i.e., in a sense, we can reduce our treatment of relations to that of classes, because an ordered pair is a 'single object doing the work of two' (1960, 258). As such, Quine (1960, 258) formulates the following postulate:

(OPPP) If $\langle x, y \rangle = \langle z, w \rangle$, then $x = z$ and $y = w$.

Here, '$\langle x, y \rangle$' is the ordered pair of 'x' and 'y'. As one entity, '$\langle x, y \rangle$' is doing the job of two because it's doing the jobs of both x and y as well as figuring as entity itself. The postulate (OPPP) ensures this.

I call postulates like (OPPP) *purpose postulates*. In the case of ordered pair, one postulate suffices; in the general case, there can be any number of purpose postulates. The purpose postulates capture the useful part of the respective defective nouns. Formulating the purpose postulates corresponds to Carnap's clarification process which precedes the construction of an explicatum.

The ordered pair is distinct from an *un*ordered pair insofar as we can distinguish its elements via their order. For example, '$\langle x, y \rangle$' has as its first element 'x' and as its second 'y'. They are only then not distinct if they are the same entity, i.e., $\langle x, y \rangle = \langle y, x \rangle$ iff $x = y$; otherwise, order makes a difference.

---

[8]See Chihara (1968, 1973, §III.3, 1974) for discussion.

[9]An anonymous referee objected that 'ordered pair' is not defective. However, Quine *does* classify 'ordered pair' as defective noun; see the beginning of §53 of W&O where he briefly discusses Pierce's account. From today's perspective, this is difficult to understand as we have settled on what here turns out to be the *explicatum*, i.e., our current understanding is the end-product of an explication (not necessarily a Quinean, though).

Different ways to account for ordered pairs have been proposed. Quine (1960, 259) discusses Wiener's original proposal to identify the ordered pair $\langle x, y \rangle$ with the set $\{\{x\}, \{y, \emptyset\}\}$ where '$\emptyset$' denotes the *empty set*, i.e., the set which has no members.[10] Kuratowski, on the other hand, identifies the ordered pair $\langle x, y \rangle$ with a different set: $\{\{x\}, \{x, y\}\}$.[11]

Given that we have different options, we can wonder whether one explication is better than another and, in particular, whether there is one that is correct. Before engaging with this question, let us consider what Quine says about explication more generally. Following the explication of 'ordered pair', Quine explains his understanding of 'explication' as 'supplying lacks':

> We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those functions. (1960, 258f.)

Indeed, as long as the functions are preserved, the explication can be taken to be successful. Quine emphasises that it is important that there is no synonymy claim (1960, 258): explicandum and explicatum need not be synonymous (cf. Carnap 1956, 8). Connecting this explicitly to paraphrase (W&O §53: 259 refers to §38), Quine says:

> A paraphrase into a canonical notation is good insofar as it tends to meet needs for which the original might be wanted. (1960, §38, p. 182)

Quine goes on to point out that a paraphrase is allowed to close truth-value gaps (1960, 182). He calls such changes 'don't-cares' (1960, 182); more on that in Section 5.

As we noted that explication comes with ontological commitment, the 'don't-cares' provide us with a certain freedom as to with what entities we choose to 'fill those functions'.[12] Both Wiener's and Kuratowski's explications are presupposing *sets*, and both guarantee (OpPp) to be non-vacuously true due to the commitment to sets. These explications are both possible as long as sets are 'terms to our liking'. In particular, then, the theory into which we explicate was already ontologically committed to these entities playing the role of the explicatum. As Quine makes explicit in the case of 'ordered pair':

> The problem of suitably eking out the use of these defective nouns can be solved once for all by systematically fixing upon *some suitable already-recognized object*, for each x and y, with which to *identify* $\langle x, y \rangle$. (1960, 258, my emphases)

---

[10]Note that, as Gustafsson (2006, 61, n. 2) points out, this account is not entirely accurate as Wiener was working within Russellian type theory.

[11]The proposals diverge if $x \neq \emptyset$ since $\{\{x\}, \{x, y\}\} = \{\{x\}, \{y, \emptyset\}\}$ iff $x = \emptyset$.

[12]An example of a 'don't-care' for Wiener's proposal is that $\emptyset \in \bigcup \langle x, y \rangle \; (= \{x, y, \emptyset\})$ for any ordered pair $\langle x, y \rangle$ whereas this is not the case for Kuratowski's proposal.

To continue the above examples, we might have already been ontologically committed to sets before we explicated ordered pairs as particular sets (cf. Section 8). In any case, it is necessary to be ontologically committed to something for the explication to be successful because these entities are supposed to guarantee the purpose postulates to be *non-vacuously* satisfied.[13]

This also shows that explication is a theory-relative notion. The reason is simply that we need a theory in order to make sense of even having 'some suitable already-recognized object' which can serve as the explicandum's surrogate. In both Wiener's and Kuratowski's accounts of ordered pair this is certainly true since we presuppose a set theory in the background which is then capable of proving (OpPp); in particular, the ontological commitment to sets means that (OpPp) is not vacuously true.

4   Explication is Elimination

As argued above, an explication actually comes with ontological commitments. However, Quine also claims that 'explication is elimination' (1960, §53, p. 260), though 'not all elimination is explication' (1960, 261). As Gustafsson (2006, 2014) argues, Quine takes explications to eliminate ontological commitments.[14] We need an understanding of elimination that resolves the tension between the ontological commitment of an explication that is necessary for the explication to be adequate, and the idea that we eliminate ontological commitment; the discussion shows that Gustafsson's interpretation isn't entirely correct.

The obvious suggestion is that any explication of 'ordered pair' eliminates *ordered pairs*. Quine puts it like this:

> In the case of the ordered pair the initial philosophical problem, summed up in the question 'What is an ordered pair?', is dissolved by showing how we can *dispense with ordered pairs* in any problematic sense in favor of certain clearer notions. (1960, §53, p. 260, my emphasis)

However, since we explicate defective nouns, and defective nouns are defective because of their problematic identity conditions (Section 3.2; cf. Gustafsson 2006, 63f.), this cannot lead to the elimination of *ontological commitments*: ontological commitment *only* arises in canonical notation

---

[13]Cf. what Quine says regarding ideal objects (see Section 7.1): 'The appeal to ideal objects in mechanics occurs regularly through universal conditionals: thus (x)(if x is a mass point then ...). The nonexistence of ideal objects consequently does not falsify mechanics; it leaves such sentences vacuously true for lack of counter-instances' (1960, §51, p. 249). Quine goes on to reject this option.

[14]To be more precise, Gustafsson (2006, 58, 2014, 520) argues that we eliminate ontological commitment because the explication puts us into a position to adopt an equivalent theory not ontologically committed to the explicandum.

which cannot be defective. Thus, we just don't have entities or ontological commitment that we could eliminate by explicating. Therefore, we need a different account of elimination.

Let us consider in what way Quine draws the distinction between elimination and explication:

> Explication is elimination but not all elimination is explication. Showing how the useful purposes of some perplexing expression can be accomplished through new channels would seem to count as explication just in case the new channels parallel the old ones sufficiently for there to be a striking if partial parallelism of function between the old troublesome form of expression and some form of expression figuring in the new method. In this case we are likely to view the latter form of expression as an explicans of the old, and, if it is longer, even abbreviate it by the old word. (1960, §53, p. 261)

This suggests that there has to be a similarity constraint in place in order for something to be an explication and not just to be an elimination. Without such similarity, i.e., if the 'new channels' are not 'parallel' to the old ones, we have a case of elimination without explication. The passage continues as follows:

> If there was a question of objects, and the partial parallelism which we are now picturing obtains, the corresponding objects of the new scheme will tend to be looked upon as the old mysterious objects minus the mystery. Clearly this is merely a way of phrasing matters, and wrong only as it threatens the immunity of the don't-cares and suggests that one of two divergent explicantia must be wrong. (1960, 261)

This passage gets us closer to the question in what cases an explication is 'correct', which we consider in Section 5. For the current discussion, the remark about the 'don't-cares' is more relevant.

Besides a similarity requirement, the passages do not tell us much about why certain cases are only cases of elimination and not also of explication. In particular, Quine's 'don't-cares' seem to pull into another direction as the last quotation also suggests. His requirement for successful explication is simply that the explication preserves the functions singled out as purposeful. This means, in particular, that the explication has to lead to ontological commitments as we do need entities to fulfil the functions. Anything beyond that is declared a 'don't-care'.

To get clearer about what is going on, let us reconsider how Quine introduces the topic of explication. As we have seen in Section 3.2, Quine distinguishes between useful and useless defective nouns, and it is the useful which are the target of explications.

> The contrast drawn at the beginning of this section, between the defective noun whose objects we dispense with and the defective

> noun whose defectiveness we are at pains to eke out *so as to keep the objects*, can then be put more simply: it is just a matter of whether the ostensible objects of the defective noun played roles that still want *playing by some sort of object.* (1960, §53, pp. 261f., my emphases)

The most interesting part is that Quine tells us explicitly that the explication wants to 'keep the objects'.[15] What this shows is that Quine's claim that explication is elimination does *not* get rid of part of the ontology. What it does is to eliminate a 'troublesome form of expression'. Explication, for Quine, is not about ontology *per se* (though see Section 8). The problem is that certain nouns fulfil functions that need to be fulfilled. Insofar as the nouns are defective they sometimes, but not always, occur in places in which they enforce corresponding objects. And in such places they appear to be ontologically committing. However, they don't have clear identity criteria because they also occur in different places—and in such places they might be pictured as 'mysterious objects'. Quine's point is that if we were to admit these nouns into our regimented theories, they would come with ontological commitment to mysterious entities. And it is these entities we can be said to eliminate. Nonetheless, this is just 'a way of phrasing matters'. As his slogan 'no entity without identity' suggests, there just aren't such entities. So, in particular, there is no way to eliminate them. What we eliminate are expressions which are troublesome insofar as they would give rise to such non-entities.

What we rather want is to provide surrogates in our already regimented theory—and it is only in the context of a regimented theory that talk of ontological commitment (and an ontology) makes even sense. Which particular objects we single out for the job of the defective nouns is what Quine calls a 'don't-care'; it does not matter as long as the job is done. And the guarantee that the job is done is by showing that the theory into which we explicate (non-trivially) proves the respective purpose postulates. What we were interested in, in the first place, was just the purpose of the defective noun. Explication is Quine's way of showing that such purposes are fulfilled—and part of that is to formalise the natural language expressions, i.e., to paraphrase them.

5 Paraphrase, Explication, and Adequacy Criteria

The main goal of this section is to work out Quine's adequacy criteria for explication. In order to do so, we have to consider explication in the light of paraphrase; I argue that explication is a particular kind of paraphrase.

(Gustafsson understands Quine's treatment of, e.g., singular descriptions as paraphrases and observes that 'it is somewhat difficult to say

---

[15]This point is commonly missed; e.g., it is completely absent in Hylton's exposition (2007, ch. 9, §III).

to what extent Quine counts such paraphrasing as a matter of expli-cation' (2014, 518). He realises the tension between Quine's paradigm (Example 1) and, e.g., (Example 2) (2014, 518), but still claims that (Example 2) *is a Quinean explication*—with which I disagree. Even though there is much to agree with in Gustafsson 2014, I think he misses some key features of Quinean explication. This section helps resolving this issue.)

As is well-known, Quine is a champion of paraphrase strategies. How-ever, paraphrase strategies are poorly understood (von Solodkoff 2014, 571) and, to make things worse, there are several uses of 'paraphrase' even if we exclusively consider W&O; see Table 1 (p. 17).[16]

Of course, some of these uses might amount to the same, but some are clearly distinct. I won't discuss them in any detail, but it is important to appreciate how diverse the uses of 'paraphrase' in Quine are. Similarly, Kuhn has been criticised by Margaret Masterman (1970, 61) that he uses 'paradigm' in 21 different senses in 'The Structure of Scientific Revolu-tions' (2012, first published 1962). I think it is clear, but I won't spell it out, that Quine should, likewise, be criticised for his uses of 'paraphrase'.

As mentioned in Section 3.3, Quine relates explication to paraphrase:

> Beyond those conditions of partial agreement, dictated by our in-terests and purposes, any traits of the explicans come under the head of 'don't-cares' (§38). (1960, §53, p. 259)

The title of §38 is 'Conciliatory remarks. Elimination of singular terms' which is part of chapter V called 'Regimentation'. As its title suggests, §38 contains Quine's treatment of singular terms. In particular, Quine proposes a procedure for the 'reparsing of names as general terms' (1960, 182) which might well lead to the 'closing of truth-value gaps' (1960, 182). Let me fully quote what Quine goes on to say.

> But this was a purpose of the reparsing. It would have been wrong if paraphrase carried a synonymy claim; but it does not (§33). A paraphrase into a canonical notation is good insofar as it tends to meet needs for which the original might be wanted. If the form of paraphrase happens incidentally to produce sense where the original suffered a truth-value gap and so was wanted for no purpose, we may just let the added cases turn out as they will. [...] Such waste cases, what the computing-machine engineers call don't-cares, are a frequent feature of good paraphrases, as we shall have further occasion to note. (1960, 182)

---

[16]As noted by Szabó, Quine 'is often rather elusive on what he means by paraphrase' (2003, 21, n. 20). Szabó claims to find six uses in Quine's (1948). However, the word 'paraphrase' only occurs twice in Quine (1948) so that we have to look elsewhere to substantiate Szabó's observation.

Furthermore, Gibson (1982, §3.5.1.2) comes close to a similar point as mine when he distinguishes between explication and what he calls 'simple paraphrase'.

(Use 1) translating (§16: 73, §49: 243)

(Use 2) avoiding plurals (§19: 90, §24: 118, §36: 174ff.)

(Use 3) saving communication from failing (§26: 125)

      (i) resolving ambiguity (§27: 129, §33: 157, 159)

     (ii) disambiguating scope ambiguities (§29: 138)

(Use 4) banishing/dispensing with locutions (§30: 144, §40: 191, §45: 221, §50: 248)

(Use 5) exposing two-role subjects (§32: 153)

(Use 6) explaining functions/uses of opaque verbs (§32: 154, 155)

(Use 7) exposing grammatical form (§32: 156)

(Use 8) accounting for non-referential position/bringing out the distinction between referential and non-referential positions (§32: 156)

(Use 9) exposing structure/gaining clarity of structure (§32: 156, §40: 191)

(Use 10) formalizing into canonical notation (§33: 159, 160f., §36: 175, §38: 182, §46: 225, §49: 242)

(Use 11) simplifying/regimenting theory (§34: 162, 163, 165, §40: 191, §46: 224, §51: 250)

(Use 12) eliminating singular terms/singular descriptions (§37: 179, §38: 182f., 184)

(Use 13) philosophical analysis (§43: 206, 208, §53: 258)

(Use 14) reducing to a single term (§47: 232)

(Use 15) making ontic content explicit (§49: 242)

(Use 16) definition (§51: 250)

(Use 17) coordinating uses of diverse theories (§51: 250)

(Use 18) eliminating reference (§52: 254)

(Use 19) explication (§52: 253f., §53: 258, §54: 263f.)

Table 1: Uses of 'Paraphrase' in W&O

Before focusing on the adequacy criteria suggested in this quotation, let us consider the 'don't-cares'. The quotation is the first of six occurrences of this expression. Four of those six occurrences are contained in §53; the remaining one occurs in §47 referring to §38.

In the quotation, Quine tells us what he calls 'don't-cares', viz., 'waste cases' which don't have a purpose. As such cases don't have a purpose, Quine feels free to let them 'turn out as they will' since it doesn't make a difference. Besides that, Quine insists in §47 that they are 'a major source of simplicity of theory' (1960, 229), i.e., admitting don't-cares leads to simpler theories. For example, we can admit intuitively meaningless sentences such as Russell's 'Quadruplicity drinks procrastination' (Quine 1960, 229) and treat them as false, instead of meaningless. This simplifies our notion of sentence insofar as we do not have to distinguish between meaningful and meaningless constructions.

This also ties in with other uses of paraphrase, viz., 'paraphrasing into canonical notation' (USE 10), 'simplifying/regimenting theory' (USE 11), and 'eliminating singular terms' (USE 12). Indeed, (USE 12) is, as its title suggests, the point of §38. What the quotation above tells us is that don't-cares are 'a frequent feature of good paraphrases', because it allows us to simplify our theories. Furthermore, the quotation also tell us that paraphrased and paraphrase need not be synonymous.

More to the point, we noted that §53 refers to §38. Quine does not explicitly say that explications are paraphrases, but the similarities are striking. Neither explication nor paraphrase need to preserve synonymy, and both come with don't-cares. Moreover, explication is a purpose-driven notion since it comes with purpose postulates. The above quotation, too, speaks of the purposes of what's being paraphrased. And, crucially, paraphrase also connects to elimination (USE 12). (Section 6.2 further corroborates the claim 'explication is paraphrase'.)

Therefore, insofar as explication is reducible to paraphrase, providing adequacy criteria for paraphrase also provides adequacy criteria for explication. Of course, as every explication is a paraphrase, but not vice versa (as, e.g., (USE 2) witnesses), we are free to place additional adequacy criteria on explication; indeed, that is what we have done with the purpose postulates. Considering whether the Wiener- or the Kuratowski-definition of 'ordered pair' are correct, Quine says the following:

> Which is right? All are; all fulfill [(OPPP)], and conflict with one
> another only out among the don't-cares. (1960, §53, p. 260)

Thus, an explication is adequate as long as it guarantees the explanandum's purpose postulates. We have also seen that this enforces ontological commitment to the respective explicata in order for the purpose postulates to be non-vacuously true in the theory.

Moreover, as the potential explications conflict with respect to the don't-cares, and Quine refers that discussion to §38, i.e., back to paraphrase, we can see again that explication is a kind of paraphrase.

Therefore, we have to consider the adequacy criteria for paraphrase—or, more precisely, the adequacy criteria for the relevant use of paraphrase. The relevant use in §38 is (Use 12). The purpose of §38 is to show how to paraphrase singular terms into canonical notation (Use 10) so that variables are the only remaining singular terms (1960, 185). The displayed quotation already suggests an adequacy criterion:

> A paraphrase into a canonical notation is good insofar as it tends to meet needs for which the original might be wanted. (1960, 182)

In this respect, Quine insists that there is no 'synonymy claim' and refers to §33 entitled 'Aims and claims of regimentation'. One immediate upshot of dismissing synonymy is that Quine's dismissive attitude towards closing truth-value gaps is somewhat justified—another feature of both paraphrase (§38) and explication (§53).

The phrasing of the criterion also relates to purposes; a paraphrase is good insofar as it serves its purposes. This raises the questions (i) whose purposes we are talking about and (ii) who judges whether they are served.

To answer these questions, we have to consider what paraphrase in this sense is supposed to accomplished. In §33, Quine tells us that it is simplification:

> It is the part of strategy to keep theory simple where we can, and then, when we want to apply the theory to particular sentences of ordinary language, to transform those sentences into a 'canonical form' adapted to the theory. (1960, 158)

Quine claims that this is 'the best division of labor' between 'theoretical deduction' and 'paraphrasing ordinary language into the theory' (1960, 159). Of course, the paraphrase depends on the theory, and the theory depends on a theorist—and only the theorist is in a position to judge the adequacy of the paraphrase:

> For normally he himself is the one who has uttered, as part of some present job, the sentence of ordinary language concerned; and he can then judge outright whether his ends are served by the paraphrase. (1960, 159)

In order for the theorist to judge this, s/he has to know the purposes of what s/he uttered in ordinary language. These purposes are also not fixed once and for all, but they are allowed to change (1960, 160)—also and even in light of potential paraphrases. Crucially, this job has to be done by the theorists themselves:

> We cannot paraphrase our opponent's sentences into canonical notation for him and convict him of the consequences, for there is no synonymy; rather we must ask him what canonical sentences he is prepared to offer, consonantly with his own inadequately expressed purposes. (1960, §49, pp. 242f.)

Thus, the only one in a position to judge the adequacy is whoever does the paraphrasing.

This, then, answers our above questions: it is the theorist who asserts a sentence within ordinary language whose purposes are relevant, and it is the theorist him-/herself who judges whether the purposes have been served by the paraphrase. Insofar as we are capable of rephrasing our sentences to avoid/resolve ambiguity within ordinary language (cf. (Use 3)), we are also capable of providing paraphrases into canonical notation, because canonical notation is a special part of (semi-)ordinary language (1960, §33, p. 159). In particular, we are allowed to make changes; indeed, we also have to make changes if we want to avoid/resolve ambiguities. Given that we are in a situation in which we need to do so, if the paraphrase didn't differ from the paraphrased, then we wouldn't have avoided/resolved the ambiguity, i.e., the very purpose of paraphrasing would not be fulfilled. This is Quine's reason to dismiss synonymy.[17]

Relating this back to explication, we can see again that explication is a kind of paraphrase. As paraphrases come with purposes, explications are special only insofar as we formulate explicit purpose postulates. As argued above, explication is a theory relative notion, i.e., only after we fixed a theory can we judge the success of an explication. This is different from other kinds of paraphrase; indeed, one use of paraphrase, viz., (Use 11), is regimenting in order to get to a theory. Whether the particular paraphrase is adequate can therefore not be judged with respect to a yet to be established theory.

Furthermore, paraphrase is not always elimination. Quine is also explicit about this when he says that '[n]one of the eliminations of singular terms in §§37 and 38 eliminated objects' (1960, §40, p. 192) and, clearly, paraphrase doesn't eliminate (defective) notions. Nevertheless, the features of explication which we have worked out (Section 3) arise from the corresponding features of paraphrase.

The striking difference between paraphrase and explication is, then, the adequacy requirement of explication that enforces ontological commitment to something. Of course, one use of 'paraphrase', viz., (Use 15), is closely connected:

> To paraphrase a sentence into the canonical notation of quantification is, first and foremost, to make its ontic content explicit, quantification being a device for talking in general of objects. (1960, §49, p. 242)

Nevertheless, ontological commitment cannot be, in general, an adequacy criterion of paraphrase.

---

[17]Quine presents this argument at least twice: §33 (p. 159) to dismiss the synonymy claim for paraphrase and in §53 (p. 259) to dismiss the synonymy claim for explication. As the discussion of the text makes obvious, the argument reduces to the claim that synonymy cannot be an adequacy criterion for (Use 3) because it inhibits disambiguation.

6 Explication Examples

Let us consider some examples of explication: natural numbers (Section 6.1) and mind (Section 6.2). Besides illustrating the account, the examples also corroborate the claim that explication is (a kind of) paraphrase.

6.1 Explication Example 1: (Natural) Numbers

Quine's first explication example is 'number' (§54) which is a straightforward and clear-cut instance. To show that, we have to formulate purpose postulates and show how the relevant theory proves them.

The background theory is, as it was in the case of 'ordered pair', a set theory (we don't need to be more specific here; note that Quine calls what I call 'sets' 'classes'). Regarding the purpose postulates, Quine tells us the following:

> The condition upon all acceptable explications of number (that is, of the natural numbers 0, 1, 2, ...) can be put almost as succinctly as [(OpPp)] of §53: any progression—i.e., any infinite series each of whose members has only finitely many precursors—will do nicely. (1960, 262)

That means that we get a succinct formulation analogous to the one of 'ordered pair':

(NnPp) (Natural) numbers form a progression.

Just as in the case of 'ordered pair', several explications are possible and acceptable. Quine considers Frege's, von Neumann's, and Zermelo's definitions and notices that the entities required by the definitions are

> all present in our universe of values of variables [...], and available for selective use as convenient. (1960, 263)

The upshot is the following.

> That all are adequate as explications of natural number means that natural numbers, in any distinctive sense, do not need to be reckoned into our universe in addition. (1960, 263)

Thus, every explication comes with ontological commitment—namely, ontological commitment to sets. The explications just identify 'numbers' with certain sets in such a way that the background set theory proves the existence of a progression, i.e., accounts for (NnPp).

Furthermore, Quine is explicit that the explication of 'number' also means that we are *not* ontologically committed to numbers. As mentioned before, we are still free, though, to call the sets we used in the explication 'numbers'.

6.2 Explication Example 2: Mind and Body

This brings us to the less clear case of mind. Quine introduces the topic as follows.

> That explication is elimination, and hence conversely that elimination can often be allowed the gentler air of explication, is an observation about a philosophical activity that far transcends the philosophy of mathematics, even if the best examples are there. Before we drop the topic, we may do well to note the bearing of that observation on the philosophical issue over mind and body. Let me lead up to the matter with a defense of physicalism. (1960, 264)

The main question Quine wants to address is the following.

> Is physicalism a repudiation of mental objects after all, or a theory of them? (1960, 265)

What Quine suggests in the first quotation is that some instances of elimination might be made sense of in terms of explication. Of course, Quine is explicit that not every elimination is an explication, but it is preferable to attempt an explication where possible. Indeed, an explication puts us into a position in which we can make sense of such notions which otherwise would just be eliminated.

As Quine attempts an explication, he needs a background theory. In the case at hand he identifies 'physicalism'. The challenge is, then, to account for mental events and mental states in terms of 'certain correlative physiological states and events' (1960, 264). In order to do so, we need to find a purpose for mental events and states that is worth preserving (1960, 264). If we cannot find a purpose worth preserving, we can just eliminate the mental events and states and do with the non-mental: 'The bodily states exist anyway; why add the others?' (1960, 264).

That this does not constitute a clear-cut case of explication, though, is explained by the absence of an adequate *paraphrase*. Quine compares the case to the Fregean explication of number which works because Frege

> paraphrases the standard contexts of numerical expressions into antecedently significant contexts of the corresponding expressions for classes[.] (1960, 266)

Quine contrasts this with the case of mind as follows.

> But when we explain mental states as bodily states, or eliminate them in favor of bodily states, in the easy fashion here envisaged, we *do not paraphrase* the standard contexts of the mental terms into independently explained contexts of physical terms. (1960, 266, my emphasis)

22

The difference between these two cases is the *availability of a paraphrase*. In the case of the clear-cut explication, there is a paraphrase available. Quine chooses to just discuss Frege's definition, but it is clear that this carries over to the other explications of number.

In case of mental terms, on the other hand, we are not in possession of an explicit paraphrase. This does, of course, *not* mean that *there is no paraphrase.* On the contrary, the adoption of physicalism implies that we can account for everything in physical terms, i.e., that there is such a paraphrase (otherwise physicalism would have counterexamples):

> What may primarily be said in characterization of physicalism [...] is that it declares no unbridgeable differences in kind between the mental and the physical. (1960, 265)

The conclusion Quine draws is that

> the distinction between an eliminative and an explicative physicalism is unreal. (1960, 265)

It is unreal because an explicative physicalism is a physicalism which explicates the mental—and thereby eliminates it. And, if we have a systematic way of eliminating it, we can, likewise, understand the elimination procedure as explicative because it would be in form of a paraphrase.[18] However, we do not have such a paraphrase available, so that sentences in mentalistic terms

> merely come to be thought of as taking physicalistic rather than mentalistic complements (1960, 266)

without providing an explicit set of instructions how to do so.

7 Non-Explication Examples

This brings us to examples which are *not* explications.[19] Besides substantiating my claim that (Example 2)–(Example 4) are *not* explications (Section 7.3) we also are now in a position to diagnose *why* they might be mistaken for explications.

7.1 Non-Explication Example 1: Limits

In W&O's §51, Quine considers and dismisses several defective nouns. The first example is infinitesimals, and the strategy to get rid of them is applied to the other cases as 'limit myths'.

---

[18]Gustafsson (2006) misses this because he misses the connection to *paraphrase*. I also disagree with his assessment that Quine's points are 'applicable only to forms of dualism that do not assume disembodied minds' (2006, 65; cf. 2014, 521).

[19]Gustafsson (2014, 522) agrees that the following examples (Sections 7.1–7.2) are not explications.

Quine explains that the invention of the differential calculus came along with the introduction of infinitesimals, even though 'the idea of infinitesimals was absurd' (1960, 248). Nevertheless, such entities were necessary for the differential calculus. Only the work of Cauchy and, especially, Weierstraß made it possible to dispense with them:

> The conflict was resolved by Weierstrass, who showed by his theory of limits how the sentences of the differential calculus could be systematically reconstrued so as to draw only on proper numbers as values of the variables, without impairing the utility of the calculus. (1960, 248)

Put differently, Weierstraß showed how to *paraphrase* these sentences. Nevertheless, this is *not* a case of *explication*, because the paraphrase also shows that the purpose of infinitesimals have been served already; we don't need something like a purpose postulate that needs to be proven within the theory of limits. Indeed, we do not need to isolate anything in this theory's ontology to play the role of infinitesimals.

Quine's strategy to dispense with other notions such as ideal objects like mass points, frictionless surfaces, and isolated systems (1960, 249), is the same as in the infinitesimal case, viz., he provides (Weierstraßian) paraphrases. Quine explains their upshot as follows.

> When we paraphrase our talk of ideal objects in the Weierstrassian spirit [. . .], we are merely switching from a theory that is conveniently simple in a short view and complex in a long view to a theory of opposite character. (1960, 250)

What these paraphrases provide us with is simplicity in the long view compared to other theories. The more complex theory 'gets the inferior rating of convenient myth' (1960, 250), but, since such theories are simple in the short view, paraphrases show us how to benefit from them.

7.2 Non-Explication Example 2: Geometrical Objects

Quine also dispenses with geometrical notions such as points, curves, surfaces, and solids (§52). Quine considers to use the same strategy as 'for the ideal objects of mechanics' (1960, 252), i.e., a Weierstraßian paraphrase, but notes that

> it ill fits the existential statements of geometry (1960, 252)

which wouldn't be properly paraphrased.

Moreover, contrasting this case with the former case of mass points, Quine notes that '[n]o sense has been made of [the mass points'] date and location' (1960, 252), and, for date in particular, there 'supervenes a perplexity of identity' (1960, 252). This means, then, that in the above cases, we at least have defective notions, i.e., potential explicanda.

On the other hand, geometrical objects raise no such evident problems of position or identity; they are positions outright. (1960, 252)

Thus, geometrical objects are *not* defective nouns and, therefore, cannot be explicated in the first place.

The upshot is that this case has to be handled differently. Indeed, Quine points out that the issue turns on

what theory will best systematize the data of physics. Thus we may fairly say that the question of the nature of the geometrical objects is, like the question of the nature of the elementary particles of physics, a question of physical theory. (1960, 253)

No explication is necessary.

7.3 The Puzzle Cases

Let us finally consider why (EXAMPLE 2)–(EXAMPLE 4) are *not* instances of explications and, more importantly, why they could be mistaken for explications.

Recall that the examples are singular descriptions (EXAMPLE 2), indicative conditions (EXAMPLE 3), and quantifiers (EXAMPLE 4). Clearly, none of them are *nouns*, so, in particular, none of them are *defective nouns*. Furthermore, none of them come with *ontological commitments* because we don't need to find anything within our theory's ontology to identify them with. Lastly, there is no theory into which we explicate them in the first place so that, even if we came up with purpose postulates, they couldn't be proven in the theory.

However, Quine does treat all of them by *paraphrasing* them. Different uses of 'paraphrase' are at work, though. Furthermore, as part of natural language, they can be taken to be *defective*—but, of course, in a different sense of 'defective' than defective nouns are defective.

We can also note that these examples are cases of explication after all—but of Carnapian explication. Since he takes his attitude towards explication to be the same as Carnap's (see Section 2.2), Quine might have been thinking of Carnapian explication when taking (EXAMPLE 2)–(EXAMPLE 4) to be explications. However, it should be clear that Quine's notion of explication is much narrower than Carnap's;[20] see Section 9.

8 Explication and Classes

As a last example of non-explication, let us consider classes. (I call them here 'classes' and not 'sets'.) Quine considers them explicitly in W&O's

---

[20]Contrary to what Gustafsson claims, viz., that Carnap's and Quine's conceptions are both 'broad and loose enough to fit many of the same particular cases' (2014, 509).

§55 after his discussion of explication (§53) and explication examples (§54).

As we noted in Sections 3.3 and 6.1, Quine assumes a theory of classes in the background when explicating 'ordered pair' and 'natural number'. Indeed, his strategy to dispense with all sorts of notions and entities is by paraphrasing them in terms of classes. This means, in particular, that we cannot take classes as the outcome of an explication. However, the case of classes illustrates an important use of explication, even though Quine is not explicit about it.

The use is this. Explication or, rather, *explicative power*, is a *criterion for theory choice*: a correct theory of the world must be able to explicate every useful noun. Crucially, here, Quinean explication is *not* iterative, because the explicandum has to be defective, but no explicatum could be defective as part of a theory couched in canonical notation (cf. footnote 6). This means that we cannot try to *further* explicate those notions which a theory is not capable of accounting for.

The explication test is quite simple. Consider a theory T. To test whether T is a candidate for being a correct theory, take a useful notion and see whether T explicates it. If T does *not* for all such notions, it is not a correct theory.

Of course, being able to explicate all useful notions might not be enough to be a correct theory, so that other considerations need to enter the picture. Nevertheless, we can use the lack of explicative power of theories to dismiss them.[21]

The above test is part of Quine's argument for the adoption of a theory of classes: it explicates all useful notions:

> Classes can do the work of ordered pairs and hence also of relations (§53), and they can do the work of natural numbers (§54). They can do the work of the richer sorts of numbers too—rational, real, complex; for these can be variously explicated on the basis of natural numbers by suitable constructions of classes and relations. Numerical functions, in turn, can be explicated as certain relations of numbers. All in all the universe of classes leaves no further objects to be desired for the whole of classical mathematics. (1960, 266f.)

Indeed, Quine gets close to explicitly stating the suggested test:

> Intensional objects aside, the abstract objects that it is useful to admit to the universe of discourse at all seem to be adequately explicable in terms of a universe comprising just physical objects and all classes of the objects in the universe (hence classes of physical objects, classes of such classes, etc.). (1960, 267)

---

[21]That is also what Quine (1998b, 430f.) suggests: if physics is not capable of explicating 'telepathic effects', assuming they have been properly established, then we reject that version of physics and need to work on a new one.

Thus, if a pure theory of classes does not provide us with everything we need, we also need to add physical objects in order to explicate everything of interest. (However, Quine 1976 seems to reduce even physical objects to classes.) Also, Quine does not accept the usual theory of classes because it overshoots its target, but endorses Gödel's axiom of constructibility (1992, §40, 1998a, 400). In order to see whether this theory of classes is adequate, we have to see whether it is capable of carrying out the explications:

> Whether classes continue to do all the services claimed for them, e.g. in foregoing pages and chapters, has then to be checked up with an eye to the specific restricted theory adopted. (1960, 268)

When successful, we can keep the theory.

9 Conclusion

Let me conclude with a comparison. Both Quinean and Carnapian explication are *two step processes*. In the first step, we clarify the explicandum, and, in the second step, we propose an explicatum. Carnapian M&N-explication often formulates conventions in the first step, whereas Quinean explication comes with purpose postulates.

Quinean explication is theory-relative: given a theory T, an explication finds T-surrogates in such way that T non-trivially proves the purpose postulates. The 'non-triviality' requirement is implemented by T ontologically committing to the explicata. This might seem anti-Carnapian, but I don't think it is; Carnap (1950, 214, n. 3) recognises the role of the variable and sees Quine and himself in agreement albeit taking umbrage at Quine's terminology (1950, 215, n. 5). Indeed, Carnap's LFoP requirement of an interpretation likewise prohibits the trivialization of explications.

Nevertheless, Quinean is much narrower than Carnapian explication. It exclusively relates expressions of natural language and expressions of canonical notation. This also connects Quinean explication and paraphrase: explication is a particular kind of paraphrase. The contrast between the number and the mind explications made this especially clear. Nevertheless, Quinean explication plays no direct role in regimentation—that's the business of paraphrase.[22] Only after the paraphrasing is done and we have a theory Quinean explication enters: it is a criterion of theory choice. If a theory does *not* explicate all useful nouns, it cannot be adequate.[23]

This contrasts with Carnapian explication: Quinean explication is not *iterative*, i.e., an explicatum can *never* serve as explicandum—but this is

---

[22]Explication is also not part of W&O's ch. V entitled 'Regimentation'.

[23]This is why explication is part of W&O's ch. VII entitled 'Ontic Decision'; for one's overall ontic decision to be adequate, it has to explicate everything useful.

not the case for Carnapian explication. Carnapian explication is a much broader notion—which Quine seems to use occasionally and mistake for his own. One reason for this confusion is that Quinean explication is paraphrase, and, as Table 1 makes obvious, the uses of 'paraphrase' are vast. Distinguishing here makes the contrast between Carnapian and Quinean explication clear, though: they serve different purposes. Carnap wants to progress science, Quine wants to test the adequacy of theories.

Acknowledgements

References

Awodey, S., and C. Klein (eds.) 2004. *Carnap Brought Home. The View from Jena.* La Salle: Open Court.

Beaney, M. 2004. "Carnap's Conception of Explication: From Frege to Husserl?" In: Awodey & Klein (2004), pp. 117–150.

Brun, G. 2016. "Explication as a Method of Conceptual Re-Engineering." *Erkenntnis* 81 (6): 1211–1241.

Burgess, A., H. Cappelen, and D. Plunkett (eds.) 2020. *Conceptual Engineering and Conceptual Ethics.* Oxford: Oxford University Press.

Burgess, A., and D. Plunkett 2013. "Conceptual Ethics I." *Philosophy Compass* 8 (12): 1091–1101.

Cappelen, H. 2018. *Fixing Language. An Essay on Conceptual Engineering.* Oxford: Oxford University Press.

Cappelen, H., and D. Plunkett 2020. "Introduction. A Guided Tour of Conceptual Engineering and Conceptual Ethics." In: Burgess & Cappelen & Plunkett (2020), pp. 1–34.

Carnap, R. 1928/98. *Der logische Aufbau der Welt.* Hamburg: Felix Meiner Verlag 1998.

Carnap, R. 1950. "Empiricism, Semantics, and Ontology." In: Carnap (1956), pp. 205–221. First published in *Revue Internationale de Philosophie* 4 (11): 20–40.

Carnap, R. 1956. *Meaning and Necessity. A Study in Semantics and Modal Logic*. Second Edition. Chicago: The University of Chicago Press.

Carnap, R. 1962. *Logical Foundations of Probability*. Second Edition. Chicago: The University of Chicago Press.

Carnap, R. 1963a. "Intellectual Autobiography." In: Schilpp (1963), pp. 1–84.

Carnap, R. 1963b. "Replies and Systematic Exposition." In: Schilpp (1963), pp. 859–1013.

Chihara, C. S. 1968. "Our Ontological Commitment to Universals." *Noûs* 2 (1): 25–46.

Chihara, C. S. 1973. *Ontology and the Vicious-Circle Principle*. Ithaca: Cornell University Press.

Chihara, C. S. 1974. "On Criteria of Ontological Commitment." In: Severens (1974), pp. 69–87.

Cohen, R. S., P. K. Feyerabend, and M. W. Wartofsky, Marx W. (eds.) 1976. *Essays in Memory of Imre Lakatos*. Boston Studies in the Philosophy of Science, Vol. XXXIX. Dordrecht: D. Reidel Publishing Company.

Gibson, R. F. Jr. 1982. *The Philosophy of W.V. Quine. An Expository Essay*. With a Foreword by W.V. Quine. Tampa: University of Florida Press.

Gustafsson, M. 2006. "Quine on Explication and Elimination." *Canadian Journal of Philosophy* 36 (1): 57–70.

Gustafsson, M. 2014. "Quine's Conception of Explication – and Why It Isn't Carnap's." In: Harman & Lepore (2014), pp. 508–525.

Hahn, L. E., and P. A. Schilpp (eds.) 1998. *The Philosophy of W. V. Quine*. Expanded Edition. La Salle: Open Court.

Harman, G., and E. Lepore (eds.) 2014. *A Companion to W.V.O. Quine*. Malden: Wiley Blackwell.

Haslanger, S. 2012. *Resisting Reality. Social Construction and Social Critique*. Oxford: Oxford University Press.

Hylton, P. 2007. *Quine*. New York and London: Routledge.

Kuhn, T. S. 2012. *The Structure of Scientific Revolutions*. Fourth Edition. With an Introductory Essay by Ian Hacking. Chicago: The University of Chicago Press.

Lakatos, I., and A. Musgrave (eds.) 1970. *Criticism and the Growth of Knowledge*. Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, Volume 4. Cambridge: Cambridge University Press.

Leitgeb, H., and A. Carus 2021. "Rudolf Carnap." *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/sum2021/entries/carnap/>.

Loux, M. J., and D. W. Zimmerman (eds.) 2003. *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press.

Masterman, M. 1970. "The Nature of a Paradigm." In: Lakatos & Musgrave (1970), pp. 59–89.

Quine, W. V. O. 1948. "On What There Is." *The Review of Metaphysics* 2 (5): 21–38.

Quine, W. V. O. 1951. "Two Dogmas of Empiricism." In: Quine (1980), pp. 20–46. Originally published in *The Philosophical Review* 60 (1): 20–43.

Quine, W. V. O. 1953. "Logic and the Reification of Universals." In: Quine (1980), pp. 102–129.

Quine, W. V. O. 1957/58. "Speaking of Objects." In: Quine (1969), pp. 1–25. Originally published in *Proceedings and Addresses of the American Philosophical Association* 31 (1957/58): 5–22.

Quine, W. V. O. 1960. *Word and Object*. Cambridge: The MIT Press.

Quine, W. V. O. 1968. "Existence and Quantification." In: Quine (1969), pp. 91–113. Originally published in *l'Age de la Science* 1: 151–164.

Quine, W. V. O. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Quine, W. V. O. 1976. "Whither Physical Objects?." In: Cohen & Feyerabend & Wartofsky (1976), pp. 497–504.

Quine, W. V. O. 1980. *From a Logical Point of View*. Second Edition, Revised. With a new Foreword by the Author. New York: Harper Torchbooks.

Quine, W. V. O. 1992. *Pursuit of Truth*. Revised Edition. Cambridge: Harvard University Press.

Quine, W. V. O. 1998a. "Reply to Charles Parsons." In: Hahn & Schilpp (1998), pp. 396–403.

Quine, W. V. O. 1998b. "Reply to Hilary Putnam." In: Hahn & Schilpp (1998), pp. 427–431.

Schilpp, P. A. (ed.) 1963. *The Philosophy of Rudolf Carnap*. La Salle: Open Court.

Severens, R. H. (ed.) 1974. *Ontological Commitment*. Athens: University of Georgia Press.

Szabó, Z. G. 2003. "Nominalism." In: Loux & Zimmerman (2003), pp. 11–45.

Tarski, A. 1933. "The Concept of Truth in Formalized Languages." In: Tarski (1983), pp. 152–278.

Tarski, A. 1983. *Logic, Semantics, and Metamathematics. Papers from 1923 to 1938*. Second Edition. Translated by J. H. Woodger. Second Edition edited and introduced by John Corcoran. Indianapolis: Hackett Publishing Company, Inc.

von Solodkoff, T. 2014. "Paraphrase Strategies in Metaphysics." *Philosophy Compass* 9 (8): 570–582.