

Constructing the World and Locating Oneself

Peter Pagin¹ 

Published online: 31 August 2017

© The Author(s) 2017. This article is an open access publication

Abstract In *Our Knowledge of the Internal World*, Robert Stalnaker describes two opposed perspectives on the relation between the internal and the external. According to one, the internal world is taken as given and the external world as problematic, and according to the other, the external world is taken as given and the internal world as problematic. Analytic philosophy moved from the former to the latter, from problems of world-construction to problems of self-locating beliefs. I argue in this paper that these problems are equivalent: both arise because experience and objective, external facts jointly underdetermine their relation. Both can be seen as a problem of expressive completeness; of the internal language in the former case, and of the non-indexical language in the second.

1 Introduction

In the beginning of *Our Knowledge of the Internal World*, Robert Stalnaker characterizes the difference between two perspectives on reality, the *internalist* and the *externalist*. Stalnaker writes:

The Cartesian internalist begins with the contents of his mind—with what he finds by introspecting and reflecting. This is what is unproblematic; these are the things and the facts we know directly. The internalist’s problem is, how do we move beyond these to form a conception of an external world, and how are we able to know that the world beyond us answers to the conceptions that we form. The externalist, in contrast, proposes that we begin with the world we find

✉ Peter Pagin
peter.pagin@philosophy.su.se

¹ Department of Philosophy, Stockholm University, Stockholm, Sweden

ourselves in, and with what either common sense or our best scientific theories tell us about it. Among the things we find are human beings—ourselves—who are things that (it seems) can know about the world, can experience it, have a point of view on it. Our problem is to explain how our objective conception of the world can be a conception of a world that contains things like us who are able to think about and experience it in the way that we do (Stalnaker 2008, pp. 2–3).

In short: the internalist takes subjective internal facts as basic/given and objective external facts as problematic, while the externalist takes objective external facts as basic/given and subjective internal facts as problematic.¹

Stalnaker takes Hume and Kripke as examples of an internalist and an externalist, respectively. Interestingly, we can view the history of the past hundred years of (relevant parts of) analytic philosophy as moving from the internalist perspective to the externalist perspective, from regarding the physical world as problematic to regarding the internal world as problematic.

It is natural to regard the two grand world-constructing projects of the early 20th century as representing the internalist perspective. These are Russell's *Our Knowledge of the External World* (Russell 1914) and Carnap's *Der Logische Aufbau der Welt* (Carnap 1928). Both Russell and Carnap attempted to *construct* the physical world on the basis of what is immediately given in the sense of *reducing* statements that appear to be about the physical world to statements that are not. The reduction in both Russell and Carnap takes place by means of the formal apparatus of Russell's and Whitehead's *Principia Mathematica* (Whitehead and Russell 1910).

In the respect of being internalists, there are substantial differences between Russell and Carnap. Russell appealed to *sense-data*, that which is immediately given to the senses (Russell 1914, p. 35). The constructing subject uses his/her own sense data. The sense-data are not considered as internal to a mind, only as those parts of reality that are directly given to a mind. Russell further appealed to *possible sense-data*, which become actual when experienced but can remain unactualized.

Compared with Russell, Carnap was a more thorough-going internalist. The empirical basis for Carnap is explicitly mental ("auto-psychological", experienced by the constructing subject). In fact, Carnap uses only a single basic concept, the relational concept of a *recollection of similarity* (*Ähnlichkeitserinnerung*) between elementary experiences (Carnap 1928, p. vii). Both Russell and Carnap had the idea that the basic language in which the construction is carried out is *expressively complete*. By definition, every proposition (if it can be expressed at all) can be expressed in an expressively complete language. Some propositions may require very complicated sentences to express, but clever definitions can allow us to express them in a concise way.²

¹There may also be *objective internal* facts, e.g. facts described in third-person vocabulary concerning mental states (*John is thinking about vacation*), as well as *subjective external* facts, e.g. physical facts described from a first-person perspective (*my pants are on fire*).

²It should be noted that the question of expressive completeness is not formally equivalent with the question of reduction. For it may be that everything expressible in L is expressible in L' , although there is

Not long after the publication of the *Aufbau*, however, Carnap was convinced by Otto Neurath that the requirement of the intersubjectivity of science could not be met by a construction from a mental basis.³ He made the move, in ‘Die Physicalische Sprache als Universalsprache der Wissenschaft’, of switching from the psychological to the *physicalistic* language as the basic language. The basic language was then conceived of as *universal* in the sense that every (meaningful) sentence can be translated into a sentence of that language (Carnap 1932, p. 437). The most elementary sentences of the physicalistic language were the *protocol sentences* (Carnap 1932, pp. 438–39). Carnap offers three alternative types of protocol sentence, without himself wanting to choose between them. The alternatives were

- (1) a. ‘Now joy’ (‘Jetzt Freude’); ‘Now here blue, there red’ (‘Jetzt hier Blau, dort Rot’.)
- b. ‘Now red circle’ (‘Jetzt roter Kreis’).
- c. ‘A red die is lying on the table’ (‘Auf dem Tisch liegt ein roter Würfel’).

The choice between these types of protocol sentence had been a matter of debate between the members of the Vienna Circle, but these differences need not concern us here. Two things are to be noted, however.

The first is the universal character of the language: that any (meaningful) sentence can be translated into the physicalistic language has the consequence that the physicalistic language is *expressively complete*: all possible states of affairs can be expressed in it. In fact, Carnap argues that psychological sentences can be translated into physicalistic sentences by means of empirical correlations of states. The sentence

- (2) *S*’s body is now red-seeing

which is an abbreviated description of a state of *S*’s body, is correlated with the sentences

- (3) a. *S* is now seeing red
- b. Now red

where (3a) and (3b) are protocol sentences (Carnap 1932, p. 457). The correlation consists in the fact that the physical state can be identified by certain stimulus-response dispositions, including the disposition to utter a sentence like (3b) upon being shown something red.

The appearance that (2) has a different meaning from (3a) is dismissed as a mistake generated from the misleading *material form* of speech. In the formal mode of speech, there is nothing more to sameness of meaning than *interderivability*. The interderivability of (2) and (3a) is established by means of the empirical correlation

no *recursively definable* translation function from *L* to *L'*. Cf. next section. In the cases under consideration, however, it is because of the failure of expressive completeness that adequate recursive translation functions don’t exist.

³Carnap was in fact aware of the problem already in the *Aufbau*, and in the end tried to eliminate the one basic auto-psychological concept by means of existential generalization over relations, in order that nothing purely subjective should remain in the system of definitions. Cf. (Carnap 1928, pp. 234–38).

itself (Carnap 1932, p. 458). Thanks to this correlation, the protocol language of a person S becomes a part of the physicalistic language.

We can see here that Carnap himself switched over from an internalist to an externalist perspective around 1930. The problem of reducing sentences of a physicalistic language to an experience language is replaced by the problem of incorporating the protocol language into the physicalistic language by means of translation.

The second thing to note is that Carnap explicitly allows the use of indexicals in protocol sentences. There is no hint in Carnap (1932) that it is possible, or desirable, to replace indexicals by constant terms. In hindsight, we can see a clear connection between the reduction project and the elimination of indexicals. That indexicals are eliminable means that any proposition that can be expressed in language L with indexicals can be expressed in a language L' free of indexical context dependence. In particular, this holds if L' is expressively complete.

Carnap (1936) presents a new reason why the reduction project can not succeed: theoretical sentences cannot be reduced to protocol sentences (observation sentences) in the sense that for each theoretical sentence there is an equivalent sentence of the protocol language.⁴ Although a theoretical sentence entails sentences of the protocol language, it is not entailed by any finite set of protocol sentences. Carnap illustrates this with the example of solubility:

- (4) a. x is soluble in water.
 b. Whenever x is put into water, x dissolves.

Carnap (1936, p. 440). Carnap notes that (4b) does not mean the same as (4a), for (4b) is automatically true of an object that is never put into water, while (4a) isn't.⁵ Carnap instead proposes the method of so-called *reduction sentences*, which only *conditionally* equate theoretical sentences and protocol sentences. This is a weaker requirement than that of reduction in the sense of the *Aufbau*.

In 'Two dogmas of empiricism', Quine proposes another reason why the *Aufbau* reduction project could not succeed. He claims that in a sentence about physical space-time of the form

- (5) Quality q is at x ; y ; z ; t

the expression 'is at' remains undefinable from the basic language. We will return to this claim in the next section. It is intimately connected to the question of the definability of 'here'.

In (1936, p. 464), Carnap does introduce non-indexical reference to times:

- (6) On May 6, 1935, at 4 P.M., there is a round table in my room.

⁴I assume here that a reduction of a physicalistic language to an auto-psychological language would proceed via the reduction of theoretical sentences to protocol sentence. On this assumption the failure of the latter reduction has an indirect bearing on the *Aufbau* reduction project.

⁵It seems that we can get around the problem of non-equivalence by using counterfactuals: "If x were to be put into water, it *would* dissolve". The problem with this solution is that the part of the truth-conditions of the counterfactual that go beyond the truth-conditions of (4b) are not *testable*. Any testing that is actually known is testing in the actual world, so you only test the counterfactual when you also test (4b). Thus, if you require testability, as Carnap did in 1936, the counterfactual does not add anything to (4b).

but still allows an indexical reference to locations. The discussion of the possibility of reducing such sentences to sentences about possible perceptions concerns the possibility of eliminating the term ‘table’. Carnap also treats ‘my room’ as subject to elimination, but without stressing that the use of the indexical is problematic over and above the use of the ordinary language noun ‘room’. Presumably, the possibility of eliminating indexicals from the physicalistic language was seen as desirable but unproblematic.

I don’t know when the elimination of indexicals was first explicitly raised in the literature. It might have been by Nelson Goodman (1951), in the last chapter of *The Structure of Appearance*. Goodman says that for “formal discourse” the context-dependence (“being not freely repeatable”) of sentences with tense or indexicals (what Goodman calls “indicator words”) is awkward, even though of great practical utility (Goodman 1951, p. 268). He suggests translation as a way out. Thus, (7a) is translated into (7b).

- (7) a. Randy is running now
 b. Randy runs [tenseless] on October 17, 1948, at noon E.S.T.

He considers an objection:

Against such translations, it is sometimes urged that they do not really convey the content of the originals. A spoken [(7a)] tells us that the action takes place at the very moment of speaking, while a [spoken (7b)] does not tell us that the action takes place simultaneously with either utterance unless we know in addition that the time of the utterance is October 17, 1948 at 10 P.M., E.S.T. Since—the argument runs—we recognize the tenseless sentence as a translation of the tensed one only in the light of outside knowledge, we have here no genuine translation at all. But this seems to me no more cogent than would the parallel argument that “L’Angleterre” is not a genuine translation of “England” because we recognize it as a translation only if we know that L’Angleterre is England.

Goodman’s requirements on acceptable translations is similar to Carnap’s in 1932. I interpret Goodman here as a) taking the correlation between utterances of sentences like (7a) and sentences like (7b) as a sufficient basis for acceptable translation, and as b) taking the translatability as indicating that nothing is semantically lost in moving from (a)-sentences to (b)-sentences.

Neil Wilson (1955) also proposes a sentence format with explicit time reference and a tenseless copula, although the topic of the context here is not expressive completeness, but a discussion in metaphysics.⁶

In *Word and Object* (Quine 1960), by contrast, the elimination of tense and temporal context dependence does concern expressive power. Quine argues that we can

⁶Wilson wants to introduce *substance*-talk where properties are ascribed to individuals relative to times, to replace the ascribing of properties to space-time points, as had been proposed earlier by Quine (1950).

replace the tensed copula by a tenseless copula by introducing explicit reference to time (Quine 1960, p. 170). (8a) is transformed into (8b):

- (8) a. I will not do it again.
b. I do not do it after now.

Here ‘do’ is supposed to be tenseless. He follows up by arguing for the elimination of indexicals (“indicator words”) in order to arrive at so-called “eternal sentences”:

An eternal sentence may be expected to be free of indicator words, but there is no bar to its containing names, however parsed (§37), or other ostensibly learned terms. Terms present may well have been learned with help of indicator words (Quine 1960, p. 193).

Quine goes on to say that

I find no good reason not to regard every proposition as nameable by applying brackets to one or another eternal sentence (Quine 1960, p. 193).

In effect, Quine here suggests that a language without temporal context dependence can be expressively complete.⁷

However, already the year before, Prior (1959) argues to the contrary. Prior explicitly objects to Wilson’s (1955) proposal to make the copula tenseless by means of explicit time reference. He claims that some tensed sentences simply are not synonymous with any tensed ones:

One says, e.g. “Thank goodness that’s over!”, and not only is this, when said, quite clear without any date appended, but it says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn’t mean the same as, e.g. “Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954”, even if it be said then. (Nor, for that matter, does it mean “Thank goodness the conclusion of that thing is contemporaneous with this utterance”. Why should anyone thank goodness for that?) (Prior 1959, p. 17)

Prior makes two contrasts, the first between retaining or eliminating temporal context dependence ((9a) and (9b)) and the second between tense and explicit token-reflectiveness ((9a) and (9c)):

- (9) a. Thank goodness that’s over.
b. Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954.
c. Thank goodness the conclusion of that thing is contemporaneous with this utterance.

⁷This holds strictly only for scientifically acceptable language. When moving beyond this to belief reports, there are no context independent standards of correctness, according to Quine. In such cases, not even eternalizing the sentence in the belief context is sufficient for eliminating context dependence. Cf. (Quine 1960, p. 218).

It seems clear, by today's standards, that the first contrast is the more important one: doing without context dependence leads to a loss of expressive power. Tense, or indexical temporal expressions, are related to present *perspective*, which non-context dependent expressions lack.⁸

Prior's idea is related to the theme in a series of papers in the mid-1960s by Hector-Neri Castañeda, including Castañeda (1966, 1968). He emphasized the difference between knowledge about someone who is in fact oneself, and genuine self-knowledge. He focused on the logical form of the attribution of self-knowledge to others. There is a difference between

- (10) a. The Editor of *Soul* knows that he is a millionaire.
 b. The Editor of *Soul* knows that he (himself) is a millianaire.

Castañeda (1968, p. 440). (10a) can be true if the editor of *Soul* knows that the winner of the state lottery is a millionaire, and he is himself the winner of the state lottery, although he does not know that. (10b), by contrast, as it is interpreted by default, can only be true if the Editor of *Soul* has self-knowledge. This is knowledge that he would express by means of the first-person pronoun:

- (11) I am a millionaire.

The first-person pronoun is connected with a subjective, first-person perspective on oneself, which sets it apart from any other way referring to oneself.

This theme from Castañeda was taken up by John Perry (especially Perry 1977, 1979), and indeed, with the additions from David Lewis (1979b), taken to its logical conclusion. In his famous thought experiment (Perry 1977, p. 492), he presents the case of the amnesiac Rudolf Lingens, who is lost in the Stanford Library, where he reads, among other things, a biography about himself. Perry stresses that Lingens's knowledge about himself is not self-knowledge. Only when he can say "*This is aisle five, floor six, of Main Library, Stanford. I am Rudolf Lingens*", does he express self-knowledge.

Perry in fact makes the substantial assumption that a non-indexical language can be expressively complete. Every proposition can be expressed by means of a non-indexical sentence. In a sense, Lingens does know that he is in the Stanford library, for he knows that *Lingens* is in the Stanford library, and the content of that knowledge is the same as the content of the thought that *he* is in the Stanford library.

This is also David Kaplan's (1989) position. For instance, on Kaplan's account, the *content* of an utterance by Kaplan of (12a) in Portland on 26 March 1977 is about

⁸Basically, Prior's objection is the same as the one Goodman dismisses, aside from the evaluative component. We might psychologize Prior's statement into

- (i) Prior is happy that that's over.

From Quine's point of view, tense or context dependence might not be eliminable here, but this would be an instance of psychological talk where that cannot be required.

the same as that of an utterance by anyone at anytime of (12b) (depending only the size of the *here* and the *now*).

- (12) a. I am here now.
b. David Kaplan is in Portland on 26 March 1977.

Kaplan (1989, pp. 508–9). Thus, at the level of *content*, non-indexical language is expressively complete. Of course, this does not hold at the level of *character*, but propositions are contents, not characters.

If non-indexical language is expressively complete, then, in principle, the truth value of any proposition can be learned by reading up on the world in a comprehensive library. The resulting problem, which Perry, Lewis and Kaplan drew attention to, is that there seems to be something left out, some possible knowledge that Lingens lacks. So, it seems on the one hand that he knows everything, and on the other that there is something that he could know but doesn't.

Perry called beliefs about who or where one is “locating belief” (1979, p. 5), and they have come to be known, after Lewis (1979a, p. 522) as “self-locating” beliefs. The problem of the nature of self-locating beliefs was discussed early on by Perry, by Kaplan (1989), Lewis (1979a), and Stalnaker (1981), and later by many others. Different solutions were proposed.

I am not here going to propose a solution, either to the problem of self-locating beliefs, or to the problem of world construction. On the contrary, my view is that there is no solution to either: Auto-psychological language as well as non-indexical language is expressively *incomplete*. What I shall argue in the following sections is that that the two problems are virtually equivalent. To put it in a nutshell: first-person facts and third-person facts jointly *underdetermine* the *empirical correlations* between facts of the two kinds.⁹

2 World-Construction

I shall here consider a partly simplified and partly idealized version of the world construction project, neither identical to Russell's nor to Carnap's. It will differ from Russell's in having an explicit first-person experience basis, rather than a basis of possible sense-data. It will differ from Carnap's in accepting phenomenal concepts as given, instead of defining them by means of quasi-analysis, and, more importantly, by providing an explicit adequacy condition, again in partial contrast with the method of quasi-analysis.

We assume given an *internal* language L_i and an *external* language L_e . The external language is assumed to be a language for talk about ordinary objects in physical space. The internal language is assumed to be a language describing experiences

⁹David Chalmers's recent contribution to the world construction area (Chalmers 2012) falls outside the present considerations. The reason is that Chalmers includes in the basis of his world construction what he calls the *PQTI* facts: physics, qualia, that's all-facts, and indexical facts (first mentioned on p. 22.). Hence, Chalmers's basis already includes types of fact that will here be argued to be not reducible to each other. See also note 11.

only. We shall frame the task of *reduction* of L_e to L_i as the existence of a recursively specifiable *translation function* τ from L_e to L_i that satisfies a certain adequacy condition. We shall first specify the languages.

(L_i) Let the vocabulary of L_i consist of

- i) demonstratives and proper names (referring to individual experiences)
- ii) n -place atomic predicates (for arbitrary n) (expressing phenomenal properties of and relations between experiences)
- iii) logical connectives
- iv) Quantifiers \forall and \exists
- v) Variables x_1, x_2, \dots ranging over individual experiences, and y_1, y_2, \dots ranging over sets
- vi) the identity sign ('=') and the element sign (' \in ')

The syntax of L_i consists of the ordinary formation rules of languages of first-order logic. In addition, we assume that new expressions can be introduced by means of definitions. Over and above explicit definitions, contextual definitions and implicit definitions are allowed for. We can also allow for ostensive definitions of proper names.¹⁰ In the case of implicit definitions, the semantic values of implicitly defined expressions belong to the domain of experiences, properties of experience, sets of experience or properties, sets of sets etc. We let the description operator ' γ ' as well as the set abstraction operator ' $\{ \cdot : \dots \}$ ' be contextually definable in the usual ways.

As for semantics, we assume a subject S having experiences, and a semantic function μ_i that assigns individual experiences of S as referents to proper names, as well as to pairs of demonstratives and acts of focusing by S on an experience. μ_i also assigns phenomenal properties or relations to atomic predicates. Where d is a demonstrative and r of L_i , Q is a two-place atomic predicate of L_i , j an act of focusing by S , we have

$$(13) \quad \mu_i(Q(r, d), j) = 1 \quad \text{iff} \quad \mu_i(Q)(\mu_i(r), \mu_i(d, j))$$

That is: the sentence of L_i ' $Q(r, d)$ ' is true, with respect to the act of focusing j iff $\mu_i(r)$ and what S focuses on in j stand in the phenomenal relation $\mu_i(Q)$.

The reason for requiring the values of atomic predicates to be properties rather than merely extensions has to do with assertion. Properties are projectible to new cases. Since we want the subject S to be able to assert sentences on the basis of understanding what they express, we cannot rely on mere extensions. An assertion is made at a time, and at no time t does S know extension of atomic predicates at times later than t .

With respect to atomic sentences of L_i , as well as to propositional logical compounds of atomic sentences, assertibility and truth coincide:

(TA) For any sentence $s \in L_i$ that is atomic or a logical compound of atomic sentences of L_i , and for any time t , s is true at t iff s is assertible by S at t .

¹⁰Strictly speaking, we should then need to take extensions of L_i into account, where new names are added. I shall pass over this complication, and assume that we can talk of the same language throughout.

The reason for the coincidence of truth and assertibility for L_i is that S 's evidence for the truth of a non-quantified sentence s is also the truth-maker for s . There is no appearance-reality difference for experiences. We assume that S has perfect memory but does not know about the future except by means of inductive inference. The coincidence of truth and assertibility does not extend in general to quantified sentences, of course.

Every event occurs at a time. But the subject S does not have a primitive reference to time instants or intervals. Rather, such reference must be defined. We assume that the relations or *earlier than* and *partly simultaneous with* are accessible to S . We then assume that S uses Russell's method (1914) of defining times as maximal sets of (partly) simultaneous events. Since L_i is equipped to express quantification over sets of events, it is thereby equipped to quantify over times.

Now for the external language, L_e .

L_e Let the vocabulary of L_i consist of

- i) demonstratives and proper names (as referring to physical individuals and events)
- ii) n -place atomic predicates (for arbitrary n) (expressing properties of and relations between physical individuals and events)
- iii) logical connectives
- iv) Quantifiers \forall and \exists
- v) Variables x_1, x_2, \dots ranging over physical individuals and events, and y_1, y_2, \dots ranging over sets
- vi) the identity sign ('=') and the element sign (' \in ')

We assume that L_e contains prepositions to express spatio-temporal relations, and that it is equipped with a well-defined notation for specifying spatio-temporal locations. In particular, we assume L_e to contain the sentence-form

(14) $@(P, x, y, z, t)$

whose instances express propositions that the property P is instantiated at spatial location x, y, z at time t .

We make no further assumption about the semantics of L_e . We assume that S has an intuitive understanding of L_e , and has dispositions to assert its sentences based on evidence.

We turn now to the question of the *reduction* of L_e to L_i . We say that L_e is *reducible* to L_i iff there exists a recursively specifiable translation function τ from L_e to L_i that satisfies the *adequacy condition* (AQ(τ)) stated below. That τ is recursively specifiable simply means here that τ is specified by a (computable) clause for each atomic expression of L_e , and a (computable) clause for each syntactic construction of L_e , allowing a computation of the translation $\tau(s)$ in L_i for each sentence s of L_e .

We are now ready to state the *adequacy* condition on τ :

(AQ(τ)) Necessarily, for any sentence s of L_e , s is *assertible* by S iff $\tau(s)$ is assertible by S .

This condition, of necessary co-assertibility, does require some motivation. Historically, it parallels the goals of the logicist program in the philosophy of mathematics. The goal of logicism was to a) define mathematical concepts from logical concepts, and b) prove the mathematical theorems from logical axioms and rules of inference. The b) part meant, of course, that *accepted* mathematical theorems, proved within accepted mathematics, would be provable from the logical basis. A strong form of logicism (cf. Tennant 2014) maintains that all mathematical truths are logical truths, not only mathematical theorems. However, as an *adequacy condition* on a proposed reduction, only known theorems can be used as a check.¹¹

Similarly, we clearly cannot require that any sentence of an external language that is true according to some realist semantics would be translated into a true sentence of the internal language. In particular, if a sentence is true for which there is no observable evidence, or even no observed evidence, no collection of experiences could be a truth maker. The whole point of the reduction project is the prospect that sentences of the external language are true only to the extent that they have truth makers made up of experiences and abstract operations. Thus, if a sentence s of the external language is assertible, based on evidence known by the speaker, then, according to the reduction project, the evidence is ultimately experiential cum abstract evidence. s should therefore be translated into an assertible sentence of the internal language. We cannot of course require that any sentence of the internal language should be translatable to a sentence of the external language, for not every collection of experiences correspond to an external state of affairs.

Therefore, for a given translation function τ' and a sentence s of L_e , if s is assertible by S while $\tau'(s)$ is not, or vice versa, τ' is not an adequate translation function. If there is *no* adequate translation function from L_e to L_i , the reduction project fails.

The adequacy condition is modally strengthened, however. The strengthening should be understood as requiring that in any possible world where the languages L_i and L_e have the same meaning for S as in the actual world, assertibility between any sentence s of L_e and its translation $\tau(s)$ coincide.¹² The idea behind the modal

¹¹Should we perhaps consider alternative ways of strengthening co-assertibility, with appeal to *analyticity*, or a *priority*, as discussed in chapter 1 of Chalmers (2012)? Analyticity is immediately ruled out unless it holds that an analysis of the *meanings* of the expressions of the external language already shows that they are expressible in terms of the internal language, which is completely implausible. And to the extent that analyticity entails necessity, if the necessity strengthening fails, so does anyway the analyticity strengthening.

The a priority strengthening works in the mathematical case, but only because a) the Peano axioms are derivable from the logicist base, and b), a mathematical sentence s is a logical consequence of mathematical sentences s_1, \dots, s_n iff the logicist translation of s is a consequence of the logicist translations of s_1, \dots, s_n . This has no counterpart in the empirical case we are concerned with here.

Note that Chalmers's strategy in his (2012) is to keep that reduction base rich enough to allow what he calls *a priori scrutability* (Chalmers 2012, p. 6), which is very different from selecting it by the independent criteria employed by Russell or Carnap. Note for instance that Chalmers takes Quine's 'is at' problem to be solved by "allowing spatiotemporal expressions into the basic vocabulary (Chalmers 2012, p. 9)".

¹²Of course, the translation function itself does not vary between worlds. Rather, the translation function has to satisfy a modal condition.

strengthening is simple: if the sentences of L_e are *semantically* reducible to the sentences of L_i , then it cannot depend on contingent non-semantic facts whether assertibility is preserved by the translation. The non-semantic facts that vary between worlds are facts about what experiences S has, and the distribution of physical objects and properties. If we have a sentence s of L_e that coincides in assertibility with $\tau(s)$ in world w_1 , but not in a world w_2 , then the evidence for asserting s in w_1 is *not* exhausted by the truth-maker for $\tau(s)$ in w_1 either: there is gap in assertibility conditions between the sentences that is bridged by contingent facts in w_1 but not in w_2 . As we shall see, this is exactly what happens.

Historically, the general adequacy conditions on reduction have not played a prominent role. Russell speaks at the end of Lecture III, in Russell (1914, p. 44), about accommodating the data by hypothetical construction:

Our hypothetical construction [...] shows that the account of the world given by common sense and physical science can be interpreted in a way which is logically unobjectionable, and finds a place for all the data, both hard and soft. It is this hypothetical construction, with its reconciliation of psychology and physics, which is the chief outcome of our discussion.

The distinction between hard and soft data (Russell 1914, p. 34) distinguishes the hard data which will “resist the solvent influence of critical reflection” from the soft data “which, under the operation of this process, become to our minds more or less doubtful”. Everyday beliefs about the physical world would then contribute to what counts as soft data. The quote above can then be understood as imposing the requirement that such beliefs come out justified by way of the construction. This in turn can be recast as a requirement on the translation function as currently conceived.

In addition, Russell also appeals to empirical beliefs, namely the accepted physical laws, for selecting those classes of *aspects* which constitute *things*:

Things are those series of aspects which obey the laws of physics. That such series exist is an empirical fact, which constitutes the verifiability of physics (Russell 1914, p. 49).

Again in terms of the current framework: we select a translation function τ from sentences of the external language that contain reference to things, to the sentences of the internal language, such that the translations of sentences expressing accepted laws of physics are assertible on the basis of experience/sense data.

The issue of the adequacy condition in connection with Carnap’s project is rather tricky. In his explanation of what a construction is, Carnap says

According to the explanation given above, if an object a is reducible to objects b, c , then all statements about a can be transformed into statements about b and c . To reduce a to b, c or to construct a out of b, c means to produce a general rule that indicates for each individual case how a statement about a must be transformed in order to yield a statement about b, c . This rule of translation we call a construction rule or *constructional definition* (it has the form of a definition; cf. §38). (Carnap 1928, p. 6)

This passage by itself does not impose any adequacy condition on the general rule (corresponding to our translation function) itself. However, immediately before this passage Carnap provides the example of reducing *fractions* to pairs of natural number by reducing statements about the former to statements about the latter. In the example, (15a) reduces to (15b):

- (15) a. $3/7 > 2/5$
 b. For any natural numbers x and y , if $7x = 5y$, then $3x > 2y$.

The general rule exemplified here is of course adequate: it does transform true/assertible inequality statements about fractions to true/assertible quantified statements about pairs of natural numbers.

When discussing the relation between original expressions and their constructional definitions in §50 (Carnap 1928, pp. 83–4), Carnap claims that the constructed expression (translation) frequently has a different *meaning*, but that the *extension* (in the case of sentences, truth value), is preserved. Carnap says that the constructional method is concerned solely with the *logical value* of a sentence, not with its *epistemic* value. This suggests opting for an adequacy condition in terms of truth value, leaving assertibility out. However, in the preface to the second edition of the *Aufbau*, Carnap states that coextensiveness was too weak a requirement (Carnap 1928, p. ix), for the coextensiveness must not be *accidental*, but *necessary*. Then assertibility, and even necessary co-assertibility, again becomes relevant (which is not to say that this is the notion Carnap had in mind).

The situation is complicated, however, by Carnap's idea of *quasi-analysis*. Ordinary analysis specify the constituents of complex objects (Carnap 1928, p. 112). Quasi-analysis, by contrast, is applied to *unanalyzable units*, and construct "quasi-characteristics" or "quasi-constituents" that have no independent existence (Carnap 1928, p. 114). It seems therefore that while ordinary analysis can be *correct* or *incorrect*, this does not really hold of quasi-analysis. In the case of quasi-analysis, it would seem, there is no independent standard of correctness (except for consistency).

This is not, however, Carnap's real standpoint. He does distinguish between desired and undesired results:

We have mentioned here and earlier (§72) that the application of the method of quasi analysis leads to the desired result only if special "unfavorable conditions" do not obtain. These unfavorable conditions could consist, for example, in the fact that certain *P* qualities^P always or frequently occur together with certain others. This would lead to irregularities in the derivation of *c* quality classes^c and later on in the division into *c* sensory classes^c and in the *c* Sim-order^c within the sensory classes. However, a more detailed investigation, which we have to omit for lack of space, shows that these interferences in the concept formation through quasi analysis can occur only if circumstances are present under which the real process of cognition, namely, the intuitive quasi analysis which is carried out in real life, would also not lead to normal results (Carnap 1928, p. 133).

Thus, not only can quasi analysis lead to undesired results. Carnap also assumes that it is legitimate to regard our ordinary real-world conceptions and beliefs as resulting

from a quasi-analysis that is tacitly carried out “in real life”. Hence, whatever properties we have identified in ordinary cognitive practice, and whatever justified beliefs we have arrived at in it, it should be possible to arrive at the same results through an explicitly reconstructed quasi-analysis.¹³

All in all, I think (AQ(τ)) is in itself justified, and well motivated from the historical point of view.¹⁴

3 The World-Construction Problem

I shall consider two kinds of problem for the world construction project, Type 1 and Type 2 problems. The former will be discussed rather briefly, and set aside as less important. We shall then focus on Type 2 problems.

Type 1 problems have to do generally with the relation between theory and observation. Some are exemplified by Carnap (1936, pp. 464–66). He notes that there is no finite set of protocol sentences that would correspond to a simple sentence about a physical thing such as (6):

(6) On May 6, 1935, at 4 P.M., there is a round table in my room.

The reduction of (6) would include (as a first step) sentences of the kind

(16) If on May ... somebody is in my room and looks in such and such direction, he has a visual impression of such and such a kind.

The first problem Carnap draws attention to is that no finite set of such sentences (which would also have to take account of other sense modalities) is equivalent with (6). Although incontestably true, this is only an example of the general theory-observation relation. The theory entails the observation sentences, but no finite collection of observation sentences jointly entail the theory. This holds for theories stated in an external language in relation to observation sentences (or other non-theoretical atomic sentences) in the external language itself, and likewise for theories stated in an internal language in relation to atomic sentences of the same language.

¹³I therefore completely agree with Hannes Leitgeb (2007, p. 184) when he writes “While similarity is prior to properties or qualities on the level of the phenomenalist constitution system, properties or qualities are prior to similarity on the level of Carnap’s extrasystematic description of the phenomenalist constitution system. When Carnap discusses the intended results of quasianalysis (cf. §81, 72) and distinguishes them from unintended ones that might occur under unfavourable circumstances, he takes the extrasystematic standpoint.”

¹⁴Leitgeb (2011, p. 268), in his new *Aufbau*-like system, requires that the translation (from a scientific sentence to the sentence constructed from logical vocabulary on terms for experiences) preserves *empirical content*, although not necessarily *extension*. This seems related in spirit to the present proposal of preserving assertibility. Since Leitgeb leaves the notion of empirical content undefined, a precise comparison between his notion of adequacy and the present one cannot be made.

However, in view of the underdetermination problem to be set out in the next section, we can say the following: if assertibility tracks empirical content, and if Leitgeb’s new *Aufbau* system succeeds in *preserving* empirical content by the proposed translation, then the truth conditions of sentences assigning colours to space-time points (of the form (15a)), are *not* reducible to their empirical content.

It has nothing in particular to do with the reduction problem. It is exemplified in the problem of induction.

Carnap moves on to point out “a more serious objection”: sentences of the (16) kind come out true if there is neither an observer nor a chair in his room. Again, this is incontestable, but a result of only considering actual observations. Even Quine, who did not in general accept intensional idioms, talked about “all possible observations” in discussing underdetermination (Quine 1975, p. 313).

The theme of underdetermination of theory by observation connects with Carnap’s first problem. The problem of underdetermination, as Quine (1975) presents it, is that two theories can entail exactly the same set of observation conditionals (a conditional where both the antecedent and the consequent are observation sentences; Quine 1975, p. 318). All possible observations together determine which observation conditionals are true. Conversely, the truth values of all observation conditionals fixes what the possible observations are (given that every possible observation can be expressed by an observation sentence). Hence, underdetermination by the set of all observation conditionals is underdetermination by all possible observations.

This leads over to Type 2 problems. Type 2 problems specifically concern the correlation between experience and external facts, translated into the correlation between the internal and the external language. It does not seem implausible that the totality of experience *underdetermines* physical reality, even when we restrict ourselves to *observable* physical reality. The observable physical reality is that part of physical reality about which we are justified to make assertions based only on definitions and observation. Thus, we are not taking into consideration the possibility of underdetermination of physical theory by observations storable in the external language itself. We can then imagine the following.

(UD1) There are two semantically independent sentences s_1 and s_2 of L_e that report the total sequence of facts observed by S . Any translation function τ from L_e to L_i that satisfies the semi-adequacy condition (AQ+) is such that $\tau(s_1) = \tau(s_2)$.

Here (AQ+) is the *only if*-part of (AQ(τ)):

(AQ+) Necessarily, for any sentence s of L_e , if $\tau(s)$ is *assertible* by S , then s is assertible by S .

Thus, the assumption that S has made a total sequence of observations described by $\tau(s_1)$ entails that S justified in asserting $\tau(s_1)$. By (AQ+), he is then justified in asserting s_1 . But the same holds of s_2 . If S has made a total sequence of observations described by $\tau(s_2)$, he is justified in asserting $\tau(s_2)$. By (AQ+), he is then justified in asserting s_2 . Since $\tau(s_1) = \tau(s_2)$, it follows that the same sequences of experiences/observations that make S is justified in asserting s_1 also make S justified in asserting s_2 .

But this means that (AQ(τ)) is not satisfied. For s_1 and s_2 are semantically independent. Hence being justified in asserting s_1 does not entail being justified in asserting s_2 , and vice versa. But if S would be justified in asserting s_1 , S would, by (AQ(τ)), be justified in asserting $\tau(s_1)$, and therefore by (UD1), justified in asserting

$\tau(s_2)$, and then by (AQ+), justified in asserting s_2 , contradicting the independence assumption.

What would be examples of a sentence-pair such as s_1 and s_2 ? In ‘Two Dogmas’, Quine makes the following well-known comment about Carnap’s execution of the *Aufbau* project:

[...] it provides no indication, not even the sketchiest, of how a statement of the form ‘Quality q is at $x; y; z; t$ ’ could ever be translated into Carnap’s initial language of sense data and logic. The connective ‘is at’ remains an added undefined connective; the canons counsel us in its use but not in its elimination (1951, p. 40).

Judging from the *Aufbau* commentary by Alan Richardson (2008), the observation is original with Quine, and also correct. It is one thing, however, to point out that Carnap slurs over the reduction of the ‘is at’ idiom and the spatio-temporal location vocabulary, and another to show that the reduction is impossible. Quine does the former, not the latter. Is there any reason to think that it is impossible?

We find one, I think, if we make use of an idea suggested by Max Black (1952, p. 156) and later also made use of Peter Strawson (1959, p. 20). The idea is that of a possible symmetric universe, with two parts that are qualitatively identical. We can think of it as two similar spheres, as did Black, or as one sphere with two similar halves. I shall here assume the latter, a sphere with two halves, A and B, that are qualitatively identical. In such a universe, a perceiving subject would get exactly the same sequence of experiences whether he starts out in A or in the exactly similar corresponding part in B. The argument, then, is based on the idea that any sequence of experiences *underdetermines* the physical location, since whatever the location of the experiencing subject is, he *could have had* those very experiences (or experiences indistinguishable from them) in the exactly similar but opposite location.

We assume, for *reductio*, that a plausible definition of ‘is at’, and of space coordinates, can be made by S in L_i . We need not know exactly how, but we can assume it as done along the lines suggested by Russell (1914, pp. 41–2). Russell suggested constructing physical space as a system of *perspectives*. Each perspective is a view of the universe, whether actually perceived or not (a perceived perspective is a “private world”). Perspectives can be ordered by similarity, and a point in constructed space is associated with a perspective. Two points are close in constructed space just in case their associated perspectives are *similar*. He suggests a definition of “here”:

We may define “here” as the place, in perspective space, which is occupied by our private world. Thus we can now understand what is meant by speaking of a thing as near to or far from “here”. A thing is near to “here” if the place where it is is near to my private world. We can also understand what is meant by saying that our private world is inside our head; for our private world is a place in perspective space, and may be part of the place where our head is Russell (1914, p. 42).

Applying Russell’s method straight off in a symmetric universe would have the effect of collapsing the universe into half of it. For every perspective induced from a location in one of the halves would be qualitatively identical to the perspective induced

from the corresponding location on the opposite half, and the similarity relations to the respective nearby perspectives on each half would again match each other perfectly. In terms of the similarity of perspectives, then, the duplication of qualitatively similar locations would not be registered. However, we could block this result by adding kinaesthetic experiences, recording change of orientation as well as movement. We would then be able to distinguish between moving across the boundary from A into the opposite half, B, from turning back towards the center of A.

Carnap, in §126 the *Aufbau* (1928, pp. 195–7), somewhat similar to Russell, presents desiderata on the assignment of colours to world-points. For instance, “The color of a visual sensation is assigned to a world point of the corresponding line of view” (p. 196). Carnap does not, however, propose a definition. He comments on this in the preface to the second edition of the *Aufbau* by saying that his procedure is related to the method of introducing concepts by means of *postulates* (I take this to amount to the idea of implicit definition; see note 17 on Leitgeb’s method).

Suppose then that our subject *S*, applying the kinaesthetically modified Russellian method, or any similar method, defines a coordinate system in L_i . We assume then that sentences of the form

$$(14) \quad @(P, x, y, z, t)$$

are definable in L_i (by assumption they occur in L_e), as short for “Quality q is at space-time location (x, y, z, t) ”.¹⁵ And I shall assume a translation function τ that maps any sentence of L_e of the form (14) on the syntactically *same* sentence of L_i . This means that any numeral in the x argument place is mapped on the same numeral in the x argument place, and so on. I shall assume two possible worlds w_A and w_B . In w_A , *S* is initially located at coordinates x_A, y_A, z_A in sector A. In world w_B , by contrast, *S* is initially located in the *corresponding*, qualitatively identical location in sector B, with coordinates x_B, y_B, z_B . Of course, the sentence is an abbreviation of *some* sentence in the primitive vocabulary and syntax of L_i .

Assuming that *S* has adequate perceptual abilities, correctly observing the instantiation of qualities. Then, in w_A , co-assertibility is perfectly upheld, as far as sentences of the form (14) are concerned. The L_e sentence

$$(14) \quad @(blue, x_1, y_1, z_1, t_1)$$

is assertible by *S* on the basis of observation. It is mapped by τ on the very same sentence (syntactically) of L_i , which is likewise assertible by *S* on the basis of experience.

In world w_B , however, τ does *not* preserve assertibility. When the sentence (17), as a sentence of L_i becomes assertible by *S*, it is not the *same* sentence in L_e that becomes assertible, but a sentence

$$(17) \quad @(blue, x_2, y_2, z_2, t_1)$$

¹⁵Since we are now dealing with the coordinates on the surface of a sphere we could make do with two space coordinates, but I’ll stick with three.

where the space coordinates specify the location in sector B that is qualitatively identical to the location in sector A specified by the space coordinates in (17). Note that the only difference between worlds w_A and w_B is the physical location of S . The vocabulary in L_i is defined in exactly the same way in relation to the experiences that S undergoes, and these are phenomenally identical in w_A and w_B . So the semantics of L_i is the same in the two worlds. Note further that the semantics of L_e is the same as well. The coordinates specify external locations, and the same locations are specified in L_e in w_B as in w_A . Therefore, since the co-assertibility of (17) and $\tau(17)$ is preserved in w_A but not in w_B , τ does not meet the (AQ(τ)) condition.

Does any other translation function meet the (AQ(τ)) condition? The alternative translation function τ' , which reflects the initial position of S in w_B and maps (17) in L_e on (18) in L_i does preserve assertibility in w_B . But for the parallel reasons τ' fails to preserve assertibility in w_A . And no *other* translation function would preserve assertibility even *within* the two respective worlds. Hence, no translation function satisfies (AQ(τ)), and thus the reduction fails.

It must be observed that on the assumption of symmetry between sectors A and B, sentence (17) is true iff sentence (18) is true. The assertibility of (17) (by S) will not coincide with the assertibility of (18) (by S), because S will have evidence either for the one or for the other, depending in S 's location. So, in this respect the argument depends on a variation in assertibility that does not match a variation in truth value.

However, on the assumption that S is a physical subject, for instance a human being, S occupies some space and will make a difference to what is observable by S . This will break the symmetry. For instance, should (17) be a sentence reporting a quality that depends on the presence of S himself at the location, the assertibility of (17) as a sentence of L_e will neither be assertible nor *true* in w_B , for the blueness is reported by (17) to obtain in sector A, while in w_B it only obtains in sector B.

We can, nevertheless, restore complete symmetry by assuming the existence of a physically similar *second* subject, S' , who starts out (perhaps having experiences) in the opposite sector, in sector B in world w_A and in sector A in world w_B , moving around in exactly the analogous way.¹⁶

This duplication of observers blocks truth value variation of sentences of the form (14) between w_A and w_B . The distribution of qualities is the same in both worlds. Nonetheless, we can add *self-reflective* sentences that will vary in truth value between the two worlds. We need a form of sentence that is different from (14):

(19) @($a, (x, y, z, t)$)

saying that an *object* a is at the spatial location (x, y, z) at time t . In the external language, this makes reference to a physical object. In the internal language, a is a construction of sense data / experience. As a special case of (19), we have

20 @($I, (x, y, z, t)$)

¹⁶This makes the analogy with self-location problem even better.

where ‘I’ is the first-person pronoun. Can the first-person pronoun be defined?

Russell (1914, p. 42) claims that there are two places associated with every aspect of a thing, and continues:

Every aspect of a thing is a member of two different classes of aspects, namely: (1) the various aspects of the thing, of which at most one appears in any given perspective; (2) the perspective of which the given aspect is a member, *i.e.* that in which the thing has the given aspect. The physicist naturally classifies aspects in the first way, the psychologist in the second. The two places associated with a single aspect correspond to the two ways of classifying it. We may distinguish the two places as that *at* which, and that *from* which, the aspect appears.

And just before this, Russell says that “[. . .] our private world is a place in perspective space, and may be part of the place where our head is”. Thus, the subject can define a term referring to his own head or body as the object occupying the place from which aspects of things appear.

Similarly, Carnap (1928, pp. 199–200) suggests the the expression “my body” can be defined as that thing satisfying certain conditions, including constant visual proximity and correlation between experiences of the visual and the tactile sense modalities. And the *self* (or “my mind”) can be defined as the class of “autopsychological states” (Carnap 1928, p. 205).

So let us assume that the bodily self is constructible, and that ‘I’ can be defined to refer to it. Let us also assume that ‘I’ exists in the *external* language as a self-referring expression. Assume further that τ maps ‘I’ on itself (syntactically). Also, assume that the ‘@’ predicate is defined for such referring expressions, not just quality terms. Then consider

$$(21) \quad @(I, x_1, y_1, z_1, t_1)$$

In w_A , both (21) of L_e and its translation $\tau(21)$ in L_i , syntactically identical to (21), are assertible by S and true. In w_B , however, (21) in L_i is still assertible and true, while (21) as a sentence of L_e , the external language, is not only not assertible, but also false. It is false, because S at time t_1 is not at (x_1, y_1, z_1) in sector A but at (x_2, y_2, z_2) in sector B. Hence, the reduction fails, also with respect to truth value underdetermination.

Again, at bottom, the failure depends on the fact that, in a symmetric universe, the totality of experiences underdetermines the physical facts, in particular the fact of the location of the subject S . Note that adding the general physical features of the universe, except for the location of S , still underdetermines S ’s location. So knowing these external facts about the universe will not help. Note further that underdetermination still obtains after adding knowledge of the *qualitative* correlation between the environment and the experiences (an environment of such-and-such qualities causes such-and-such experiences). In addition, if we keep the assumption of a second experiencing subject S' , even if S knows that S is in sector A while S' is in sector B, this

is not enough unless S also knows that *he* is S rather than S' (which highlights the self-location aspect).¹⁷

4 The Self-Location Problem

In his seminal paper *Frege on Demonstratives* (Perry 1977), John Perry presents the case of the amnesiacs Lingens, who finds himself in a big library with no recollection of his past, but with access to written information about the world, including information about an amnesiac named ‘Lingens’ who is lost in the Stanford library. If we could add that he has access to a true and complete–non-indexical–theory of the world, Lingens could infer that *he* is Lingens, by first-person knowledge of his memory loss and that he is in a big library, together with the library knowledge that there is exactly one person in the world that is in such a situation.

This possibility was eliminated in David Lewis’s (1979a, p. 535) extension of the thought experiment, where he adds a second amnesiac, in a qualitatively identical situation, but located at the Harvard library. That is Lauben.¹⁸ The libraries are excellent, in fact in a certain respect complete. Every objective proposition that has been made true at the time can be learned from the books available in these libraries.

By ‘objective proposition’ is meant propositions that are not essentially *indexical*. That is, every objective proposition can be *expressed* in wholly non-indexical vocabulary. When David Kaplan utters ‘I am here now’ in Portland on 26 March 1977, if what he expresses is an objective proposition, then that proposition can equally well be expressed by ‘David Kaplan is in Portland on 26 March 1977’. The former expression is an indexical way of presenting a proposition, and the latter expression a

¹⁷ I shall make some brief remarks about the failure result. Firstly, from Carnap’s own point of view, it could be claimed that the reduction will indeed not work in a symmetric universe, because then circumstances are “unfavourable” (see the quote above from Carnap (1928), §§, 70, 72, 81), but will work in an asymmetric universe, and this is good enough. It is indeed good enough if we lower the requirements on what we regard as proper translation, i.e. as a meaning-preserving mapping. See the remarks on Carnap (1932) and on Goodman (1951) above.

Secondly, Thomas Mormann (2004) proposes a way (similar in spirit to Russell’s method) of answering Quine’s objection to Carnap by defining a geometric structure directly over experiences. This allows a definition of ‘is at’. It does not eliminate the underdetermination, however.

Thirdly, Leitgeb (2011, pp. 289–96), proposes to treat ‘*col*’, used as a name of the function that assigns colours to space-time points, as implicitly defined by the desiderata provided by Carnap in Section 126 of the *Aufbau*. ‘*col*’ corresponds to the relational predicate ‘@’ used here. Leitgeb suggests formalizing Carnap’s desiderata as axioms, collectively abbreviated ‘*A*’. Thus ‘*A[col]*’ implicitly defines ‘*col*’ (Leitgeb 2011, pp. 290–1). Sentences that use ‘*col*’, of the form ‘*B[col]*’ can then be interpreted by means of this definition. Leitgeb suggests three ways, where the first is by means of a Ramsey sentence: ‘*B[col]*’ is interpreted as ‘ $\exists x(A[x] \wedge B[x])$ ’. The second uses Lewis’s uniqueness condition, and the third Hilbert’s ϵ operator.

Of course, the reduction project here requires the semantic values of ‘*col*’ (witnesses of the Ramsey sentence) to be auto-psychological properties, not functions that actually map colours on space-time points (which would be question-begging). But this again makes the definition open to the underdetermination problem.

¹⁸Lewis calls him ‘The Widener amnesiac’. Who first started using ‘Lauben’ as the name, I have not found out.

non-indexical way of presenting the very same proposition. Both Lingens and Lauben have access to records of every objective proposition that has been made true at the time, but only presented in non-indexical ways. This means, among other things, that Lingens will not find a sign saying ‘This is Stanford library’, giving him the indexical information that he is in Stanford library.

Because of knowing all objective propositions true at the time, both Lingens and Lauben know that there are two amnesiacs, named ‘Lingens’ and ‘Lauben’ in the Stanford and Harvard library, respectively, that read the books and learn about themselves.¹⁹ The epistemic problem for Lingens (and Lauben) is that this information is not enough for determining where he is, in Stanford or in Harvard. And since he knows that Lingens is in Stanford and Lauben in Harvard, he knows that he is Lingens if he is in Stanford and Lauben if he is in Harvard, he also does not know which of them he is. He can know, by combining knowledge from reading and knowledge from experience (perception and introspection), and applying Leibniz’s law, that he is distinct from every one other than Lingens and Lauben, but not which of the two he is. All his encyclopedia knowledge and all his experience put together is compatible with both alternatives.

Now there is a question in which sense there is something that Lingens does not know, and that question is also the main topic of Perry’s paper. If all propositions are objective propositions then there is no true *proposition* that Lingens does not know is true. He considers the sentence

(22) I am Lingens

and comes to the conclusion that he does not know its truth value. But if there are only objective propositions, then what is expressed by (22) is a proposition whose truth value he already does know. For instance, as is considered by Perry, if the contribution of ‘I’ to the proposition expressed is just the speaker of the sentence on the occasion, and if the contribution of the name ‘Lingens’ is just the referent, Lingens, then (22) expresses the same proposition as

(23) Lingens is Lingens

which, of course, is known to be true by both Lingens and Lauben. So there is a question in what sense Lingens does not know that he is Lingens. Similarly, there is a question in which sense he can *believe* that *he* is Lingens, or Lauben, or in Stanford, or in Harvard. This is the problem of self-locating beliefs.²⁰

¹⁹We should assume that they can’t check predictions like ‘At midnight Lingens jumps up and down while Lauben scratches his head’ and then determine who they are by a combination of decision and self-observation.

²⁰Cappelen and Dever (2013, pp. 100–2) argue that there is nothing special about the *de se* over and above ordinary Fregean problems of coreference. They complain that the problem depends on the coarseness of grain of possible-worlds account of content. This is not correct, however. As is pointed out by Perry (1979, p. 7), whatever α is, a subject might fail to realize the *he* is α . This does not depend on grain. We can let Lingens know all objective propositions with a finer grain than that of possible-worlds propositions; the problem with the apparent lack of indexical knowledge arises anyway. Indeed, as Frege himself pointed out, “everyone is presented to himself in a special and primitive way, in which he is presented to no one else” (Frege 1918, p. 333).

Various solutions have been offered (belief states, centered-worlds propositions, reflexive propositions), but our task here is to characterize the problem in a semantic way. Thus, we can assume two languages, L_s (the subjective language) and L_o (the objective language). The vocabulary of the two languages are as follows:

L_s L_s contains

- i) indexicals: *I, you, here, now* . . . , and demonstratives
- ii) nouns, verbs, and adjectives that can be applied on the basis of observation
- iii) prepositions
- vi) logical connectives, quantifiers, variables

In addition, we allow L_s to be *extended* by the introduction of new proper names, defined either from the original vocabulary or by means of ostensive definitions.

L_o L_o contains

- i) Proper names, non-indexical terms for locations and times
- ii) nouns, verbs, and adjectives that do not take contextual arguments
- iii) prepositions
- vi) logical connectives, quantifiers, variables

Both languages have standard grammatical constructions to form simple and complex sentences.

By the thought experiment, Lingens's experiences, together with all (objective) facts underdetermine his location, i.e. underdetermine at which location the experiences are taking place. The assumption that Lingens *knows* all objective propositions but does not know where he is, is a dramatic way of bringing this out. We could also dispense with the appeal to amnesia.²¹

That there is a sense in which Lingens does not know where he is, comes out in the fact that, as a sentence of ordinary English,

(24) I am in Stanford.

is not assertible by Lingens, although true. We can recast this in terms of considering the possibility of a mapping ν from L_s to L_o , analogous to τ . This is the proper mapping direction, since we are considering the question whether L_o is expressively complete: are there sentences of L_s that cannot adequately be translated into L_o ?

We then need an adequacy condition on ν corresponding to $(AQ(\tau))$, that ν preserves assertibility:

$AQ(\nu)$ Necessarily, for any sentence s of L_s , s is assertible by Lingens iff $\nu(s)$ in L_o is assertible by Lingens.

²¹Lewis's story about the two gods in Lewis (1979b) does not appeal to amnesia, but as many have pointed out, the thought experiment itself is problematic, and perhaps not coherent. Kaplan (1989, p. 531) presents the analogous case of Castor and Pollux, two identical twins, raised under qualitatively identical conditions, getting qualitatively identical stimuli (analogous to S and S' in the symmetric universe).

We now consider two worlds, w_S , the actual world of the story, where Lingens is in Stanford, and w_H , the non-actual world where Lingens is in Harvard and Lauben is in Stanford, and the other assumptions are changed as needed, *mutatis mutandis*. We can now contrast

- (25) a. I am here.
b. Lingens is in Stanford.

(25a) is a sentence of L_S and (25b) a sentence of L_O . Clearly, (25a) is assertible by Lingens in both w_S and w_H , while (25b) is assertible by Lingens in w_S but not in w_H . Hence, a translation function ν that maps ‘I’ on ‘Lingens’ and ‘here’ on ‘Stanford’ does not satisfy (AQ(ν)).

This is of course unsurprising in light of the standard difference between indexical and non-indexical sentences, but corresponds well to the situation of the Carnapian subject S with respect to the sentence (21). We might even consider a sentence-pair like

- (26) a. Blue spot at position z in this library.
b. Blue spot at position z in Stanford library.

(26a) is a sentence of L_S and (26b) a sentence of L_O . (26a) is assertible by Lingens on the basis of observation, in both w_S and w_H . If the spot emerges at a time after the period covered by the books in the library, then (26b) is not assertible in w_H , even if true (since Lingens has no evidence for its truth). This corresponds well to the situation of S with respect to (17) and (18). Of course, if the blue spot *is* described by the library books, then both sentences are assertible by Lingens in both worlds (for there must also be a recorded blue spot in the Harvard library.)

These differences between L_S and L_O depend crucially on the difference between indexical and non-indexical language. It may seem, however, that we could bridge this difference by allowing Lingens access to *new* proper names that he can define by ostension. Thus, let Lingens introduce the names ‘Lingens*’ and ‘Stanford*’ as follows:

- (27) a. Let ‘Lingens*’ be a name of me.
b. Let ‘Stanford*’ be a name of this place (here).

In virtue of these stipulations, the L_S sentence

- (28) Lingens* is in Stanford*

is assertible in w_S , just as (25a) is. Since the semantic effect will be that ‘Lingens*’ refers to Lingens and ‘Stanford*’ to Stanford, (28), as defined in w_S , will have the same truth value as (25b), both in w_S and in w_H .²² On the other hand, in w_H , the stipulations are carried out under the conditions that obtain in *that* world. This means,

²²We assume here that proper names are rigid designators, or directly referring, to keep the discussion in line with Perry’s.

that in w_H , ‘Lingens*’ refers to Lingens, but ‘Stanford*’, as defined in w_H , refers to Harvard. So defined, (28) is true in w_H and false in w_S . Nevertheless, (28) is *assertible* by Lingens in both w_S and w_H , just like (25a).

Although it might seem that the introduction of new proper names in L_S bridges the difference between the two languages, a moment’s reflection shows that this is not really the case. For there is no mapping ν from L_S to L_O that preserves meaning and co-assertibility across the two worlds. The translation ν that maps ‘Lingens*’ on ‘Lingens’ and ‘Stanford*’ on ‘Stanford’ yields both co-assertibility and co-extension between (28) and (25b) in w_S , but it does not preserve extension in w_H . For in w_H , ‘Stanford*’ is made to refer to Harvard, while ν (‘Stanford*’) = ‘Stanford’, still refers to Stanford. Likewise, the mapping ν' that maps ‘Stanford*’ on ‘Harvard’, preserves extension in w_H but not in w_S . So, the (AQ(ν)) condition is not met by any mapping from L_S to L_O .

The result in this case is not that there are mappings between L_S and L_O w_S and w_H that preserve meaning but none that also preserves co-assertibility across worlds. Rather in this case, there is no mapping that even preserves meaning across the two worlds. This results from the assumption that there is a direct option of introducing proper names in L_S by stipulation. There is no counterpart in the world-construction case to this option.

Still, we can make use in a slightly different way of this option. It is easy to see that there is a difference in the contribution to assertibility for Lingens of the two names ‘Lingens’ and ‘Lingens*’. For if they were the same, the two names would be intersubstitutable in all extensional use contexts, and the name ‘Lingens*’ could be *added* the *objective* language, L_O without changing the expressive power. However, the sentence

(29) Lingens* is in Stanford.

is clearly *not* assertible in w_S by Lingens, in contrast to (25b). If it were, given how the name has been defined, Lingens *would* know that he is in Stanford. This outcome again depends on the fact that it is underdetermined by the book facts and the experiences where the experiences are taking place.

The upshot is that the objective, non-indexical language, is expressively *incomplete*, since there is no adequate translation from the indexical to the non-indexical language.²³

5 Conclusion

The main claim in this paper is that world-construction problem that is generated by failure of reducing space-time location vocabulary is at bottom the same as the

²³More precisely, the objective language is incomplete by the translation co-assertibility condition. By Kaplan’s semantics, and by Perry’s assumptions, as indicated above, the objective language is semantically complete in the sense of able to express the same propositional contents as the indexical language. I leave to another occasion to argue for the condition and against Kaplan’s semantics.

problem of self-locating beliefs. We can regard the two problems as problems concerning expressive completeness, in case of the world-construction problem of the *internal language*, and in case of the self-location problem, of the *objective language*.

The reason why completeness fails in both cases is the same, the underdetermination by experiences and objective facts of the correlation between them. More specifically, the total record of experiences together with the total record of objective facts jointly underdetermine the correlation of experiences with objective facts about the location of the observer. Experience facts can be kept constant between two worlds while the objective location varies. That is why there is no adequate, and hence world-independent, translation function from the internal to the external language in the world-construction case, and why here is no adequate translation function from the objective language to the indexical language in the self-location case.

It therefore also turns out that there is a close parallel between on the one hand the difference between internal experience-language and external language and on the other and the difference between indexical language and non-indexical language. In both cases the problem can be set out in terms of the failure of translations that preserve co-assertibility. The common factor is that both experience language and external *indexical* language can be applied to what is immediately observed, without relying on any background knowledge of the external world.²⁴

The two reductions fail because of the same underdetermination. However, this is not enough to show that the two reductions are equivalent in a *stronger* sense: that there is an adequate translation function in the one case iff there is an adequate translation function in the other case. For one reduction could fail for *another* reason that might *not* also make the other reduction fail. This possibility remains to be investigated. The upshot of the present paper is only that there is one basic underdetermination that make both fail, and that they are equivalent in at least this weaker sense.

Acknowledgments Earlier versions of this paper have been presented to audiences at Stockholm University, the University of Milan, and at the 2nd PLM conference at CEU, Budapest, all in 2013. I am grateful for questions and comments on these occasions. The paper has improved thanks to comments from anonymous referees, and as always I have benefitted much from discussing with Kathrin Glüer-Pagin. I would also like to thank the special editors, who have been exceptionally helpful.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement no. 675415.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

²⁴The option of introducing new proper names by ostensive definition in the self-location case does not correspond to anything in the world-construction case. It alters the format of the failure to preserve co-assertibility.

References

- Beaney, Michael (ed.) 1997. *The Frege Reader*. Oxford: Blackwell Publishing.
- Black, Max. 1952. The Identity of Indiscernibles. In: *Mind* 61, 153–164.
- Cappelen, Herman, and Josh Dever. 2013. *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.
- Carnap, Rudolf. 1928. *Der Logische Aufbau der Welt*. Berlin: Bernary. Translated and published with an additional essay as *The Logical Structure of the World; Pseudoproblems in Philosophy*, La Salle: Open Court 2005. Page references to the translation.
- Carnap, Rudolf. 1932. Die Physicalische Sprache als Universalsprache der Wissenschaft. In: *Erkenntnis* 2, 432–465. Translated by Max Black as *The Unity of Science*, Routledge and Kegan Paul, London 1934.
- Carnap, Rudolf. 1936. Testability and Meaning. In: *Philosophy of Science* 3, 419–470.
- Castañeda, Hector-Neri. 1966. He: A Study in the Logic of Self-Consciousness. In: *Ratio* 8, 130–157.
- Castañeda, Hector-Neri. 1968. On the Logic of Attributions of Self-Knowledge to Others. In: *Journal of Philosophy* 65, 439–456.
- Chalmers, David. 2012. *Constructing the World*. Oxford: Oxford University Press.
- Frege, Gottlob. 1918. Der Gedanke. In: *Beiträge zur Philosophie des Deutschen Idealismus* 100, 25–50 Reprinted in Frege 1980. Translated and reprinted in Beaney 1997, pp. 325–345. Page references to the translation.
- Frege, Gottlob. 1980. *Funktion, Begriff, Bedeutung*. Göttingen: Vandenhoeck & Ruprecht.
- Goodman, Nelson. 1951. *The Structure of Appearance*. Cambridge: Harvard University Press.
- Kaplan, David. 1989. Demonstratives. In: *Themes from Kaplan*, eds. J. Almog, J. Perry and H. Wettstein, 481–563 Oxford: Oxford University Press.
- Leitgeb, Hannes. 2007. A New Analysis of Quasianalysis. In: *Journal of Philosophical Logic* 36, 181–226.
- Leitgeb, Hannes. 2011. New life for Carnap's *Aufbau*? In: *Synthese* 180, 265–299.
- Lewis, David. 1979a. Attitudes de dicto and de se. In: *The Philosophical Review* 88, 513–543. Reprinted in Lewis 1983, 133–156.
- Lewis, David. 1979b. Scorekeeping in a Language Game. In: *Journal of Philosophical Logic* 8, 339–359. Reprinted in Lewis 1983, 233–249.
- Lewis, David. 1983. *Philosophical Papers* Vol. I. Oxford: Oxford University Press.
- Mormann, Thomas. 2004. A Quasi-Analytical Constitution of Physical Space. In: *Carnap Brought Home. The View from Jena*, eds. Steve Awodey, and Carsten Klein. Chicago and La Salle, Ill.: Open Court.
- Perry, John. 1977. Frege on Demonstratives. In: *The Philosophical Review* 86, 474–497.
- Perry, John. 1979. The problem of the Essential Indexical. In: *Noûs* 13, 3–21.
- Prior, Arthur N. 1959. Thank Goodness That's Over. In: *Philosophy* 34, 12–17.
- Quine, Willard Van Orman. 1950. Identity, Ostension, and Hypostasis. In: *Journal of Philosophy* 47, 621–633.
- Quine, Willard Van Orman. 1951. Two Dogmas of Empiricism. In: *The Philosophical Review* 60, 20–43. Reprinted in Quine, *From a Logical Point of View*, third edition, 20–46, Harvard University Press, Cambridge, Mass. 1980. Page references to the reprint.
- Quine, Willard Van Orman. 1960. *Word and Object*. Cambridge: MIT Press.
- Quine, Willard Van Orman. 1975. On Empirically Equivalent Systems of the World. In: *Erkenntnis*, 313–328.
- Richardson, Alan W. 2008. *Carnap's Construction of The World. The Aufbau and the Emergence of Logical Empiricism*. Cambridge: Cambridge University Press.
- Russell, Bertrand. 1914. *Our Knowledge of the External World*. Chicago, Ill.: The Open Court Publishing Company. Republished 1993, by Routledge, London.
- Stalnaker, Robert. 1981. Indexical Belief. In: *Synthese* 49, 129–151.
- Stalnaker, Robert. 2008. *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- Strawson, Peter F. 1959. *Individuals. An Essay in Descriptive Metaphysics*. London: Methuen.
- Tennant, Neil. 2014. Logicism and Neologicism. In: *Stanford Encyclopedia of Philosophy*. Stanford University. <http://plato.stanford.edu/archives/fall2014/entries/logicism>.
- Whitehead, Alfred North, and Bertrand Russell. 1910. *Principia Mathematica*. Cambridge: Cambridge University Press.
- Wilson, Neil L. 1955. Space, Time, and Individuals. In: *Journal of Philosophy* 52, 589–598.