

Arne Næss's Experiments in Truth

Jamin Asay

University of Hong Kong

asay@hku.hk

Forthcoming in *Erkenntnis*

Abstract

Well over half a century before the development of contemporary experimental philosophy, the Norwegian philosopher Arne Næss conducted a number of empirical investigations intended to document non-philosophers' convictions regarding a number of topics of philosophical interest. In the 1930s and 1950s, Næss collected data relevant to non-philosophers' conceptions of truth. This research attracted the attention of Alfred Tarski at the time, and has recently been re-evaluated by Robert Barnard and Joseph Ulatowski. In this paper I return to Næss's research on truth in order to better develop an account of how such empirical data does or doesn't bear on the philosophical study of truth. I examine Næss's findings from his various studies on truth, and challenge the interpretation of those studies offered by Barnard and Ulatowski.

1. Introduction

Decades before the development of contemporary experimental philosophy, the Norwegian philosopher Arne Næss was conducting experiments intended to probe non-philosophers' convictions regarding various topics of philosophical interest (1938a, 1938b, 1953a, 1953b). Those experiments, particularly the ones involving truth, attracted some notable attention at the time—Tarski (1944) comments on them in one of his seminal papers on truth—and have received renewed engagement in recent years in the work of Robert Barnard and Joseph Ulatowski (Barnard and Ulatowski 2013, 2016, 2019, and 2021; Ulatowski 2016, 2017, and 2018). Næss's research raises methodological questions about how empirical investigation can contribute to the philosophical study of truth, and how such research should be conducted. In this paper I consider what Næss's work can teach us about the nature of truth, and the relevance that empirical research bears toward it. In so doing, I present

and challenge the assessment that Barnard and Ulatowski have put forward of Næss's findings.¹ Though my comments on their interpretation are largely critical, my goal is not to show that nothing of philosophical value can be gleaned from Næss's work; quite to the contrary, I believe that Næss's work has important implications for the contemporary study of truth. Nevertheless, my aim is to highlight just how difficult it is to draw substantive conclusions from the relevant data that speak to the concerns traditionally raised by philosophers who have studied truth.

I begin with some preliminary remarks regarding the empirical study of truth, and the distinctions that need to be drawn in order to best assess what Næss's studies teach us about truth and how they can inform further empirical work. Then I consider Næss's two main studies on truth—first his expansive 1938 study into non-philosophers' varying conceptions of truth, and then his 1953 study that covered synonymy relations between truth-theoretic and non-truth-theoretic expressions. For each study, I offer my own interpretation of what philosophers should take away from the studies, and my critical evaluation of Barnard and Ulatowski's competing interpretation.

2. The empirical study of truth

The goal of this section is to identify some challenges of articulating and developing the empirical study of truth. The aim is not to undermine empirical attempts to understand truth, but rather enhance them.

The first challenge is to identify what it is we intend to study when we study truth empirically. This question is no less fraught than the question of what we intend to study when we study truth non-empirically. Theorists about truth in recent years have relied on a tripartite classification that is relevant to their object of study: they distinguish between the *property* of truth (hereafter '*truth*'), the *concept* of truth (hereafter 'TRUTH'), and alethic *vocabulary* like the words 'truth', 'is true', and 'it is true that'.² The topic of the property of truth is primarily metaphysical. Properties are borne by objects that share a particular feature; they help explain what kind of thing an object is, what powers it has, and how it relates to other objects. My belief that snow is white, the English sentence 'Paris is in France', and an assertion that $7 + 5 = 12$ all possess the property *truth*. Properties have their role to play in accounting for the nature of and resemblances between objects in the world. Concepts, by

¹ For more on how Næss's research figures into the history of analytic philosophy see Chapman 2008, chapter 6, 2011, and 2018, and Murphy 2014.

² See, e.g., Bar-On and Simmons 2007, Lynch 2009, and Asay 2013.

contrast, have their role to play in furnishing minds with the ability to think thoughts and represent the world around them. Those who possess TRUTH can think and talk about truth as such. Words, of course, are our vehicles for expressing our thoughts and ideas about the things that we use our concepts to represent.

The property *truth*, concept TRUTH, and words like ‘truth’ are all the *objects* of various philosophical study: philosophers want an account of what each of them is, what they do, and how they are related to other properties, concepts, and words. When we study truth empirically, then, we must at a minimum distinguish between studying *truth*, TRUTH, and ‘truth’ empirically. But interest in truth is not exhausted by interest in properties, concepts, and words. In addition, we have thoughts, ideas, and beliefs that are *about* those various objects. It’s one thing to possess a particular concept and deploy it within one’s cognition, and another to possess a set of beliefs or *theory* about that concept. I will refer to *conceptions* as a person’s set of beliefs or theories (whether implicitly or explicitly held) about some particular object of study. A person’s conception of φ is their view concerning it; what makes that view true or false is φ itself. Of particular importance is the *folk conception* of truth: the theory of truth that dominates in the general population. So far, then, there are at least five things we might be interested in studying when we study truth empirically: *truth*, TRUTH, ‘truth’ (and its cognates), individuals’ conception(s) of truth, and the folk conception of truth (if there is such a thing).³

The second challenge is to identify the sorts of observations and analyses that will ultimately constitute the empirical study of truth. For some of the objects of study, this question is fairly straightforward. To determine a person’s conception of truth, ask them about it. (How and what to ask, of course, are subtle and difficult questions that skilled experimentalists will wrestle with.) To identify the folk conception of truth, run statistical analyses of data that collect individuals’ conceptions of truth. The collection of linguistic corpora, meanwhile, will be indispensable to the study of ‘truth’.

Less straightforward is identifying the empirical traces of *truth* and TRUTH. For one thing, *truth* might be, as Horwich (1998: 37) argues, a “logical property”, and so not obviously amenable to naturalistic analysis. Furthermore, it’s unclear how we would understand *truth* better by hauling the things that possess it—true concrete truth-bearers, say—into the lab. The English sentence ‘Snow is

³ We could further distinguish between, say, (folk) conceptions of *truth*, TRUTH, and ‘truth’. Whether such finer-grained things exist depends on whether people are actually drawing these distinctions themselves (either implicitly or explicitly). I’ll be thinking of conceptions of truth as collections of beliefs concerning any of the various dimensions of truth.

white' possesses *truth* not because of some feature intrinsic to it, but because something else which isn't in the lab, snow, possesses an entirely different property, *being white*. As for TRUTH, many of course have long held that the study of concepts is an entirely *a priori* matter, a view challenged by many philosophers, but especially by experimental philosophers.⁴ If concepts are abstract entities, then it would seem that they cannot be directly observed. And even if they are mental entities, and minds are entirely physical, it's still not clear how to ascertain the empirical consequences relevant to a conceptual analysis.⁵ In general, it's a substantive and difficult open question to say which empirical facts are relevant to conceptual analysis.⁶

Nevertheless, there are familiar strategies for gaining empirical traction on concepts, which leads to the third challenge: identifying the valid and inductively strong inferences that may be drawn from the observations and analyses in the second challenge to the objects of study in the first. Here is an example of an *unwarranted* inference in this terrain. Ulatowski writes: "We cannot derive a theory of truth from what ordinary people say about it or how people use the term 'true' or 'truth'. That would be to confuse the descriptive data with normative theory" (2017: ix). What I read Ulatowski as saying here is that while the data we can collect about people's use of alethic language do not themselves determine what truth is (the project of normative theory), they are relevant to settling empirical questions about what *conceptions* of truth are operant in the population (the project of descriptive theory). That again raises the hard to answer question as to what we *may* infer about truth—and TRUTH in particular—on the basis of people's conceptions of truth.⁷

Here is one familiar line of thought that attempts to bridge conceptions and concepts. One's theory of a particular concept should, *ceteris paribus*, capture or otherwise resemble the prevailing folk conception of that concept. Here are Barnard and Ulatowski:

⁴ See, e.g., Weinberg, Nichols, and Stich 2001 for a canonical example of bringing experimental results to bear on conceptual analysis.

⁵ Furthermore, TRUTH is often considered a paradigm abstract concept, as in Barsalou and Prinz (1997: 289-292), which could further isolate it from empirical investigation.

⁶ See Jackson 1998 and Machery 2017 for a pair of competing views.

⁷ Note that Ulatowski here is quite cautious (and rightly so) about the inferences we may and may not draw for "normative" theory from "descriptive" theory. (Thanks to a referee for stressing this point.) For a bolder perspective on empirical data making inroads into philosophical disputes, see Greene 2007. For a perspective on the futility of using the sort of data experimental philosophers have collected to make philosophical inroads, see Kauppinen 2007.

Part of what is at stake is whether our philosophical thinking about truth has been properly calibrated to the folk notion. Either the philosopher’s intuition about the folk’s views is properly calibrated or it’s not. If it is properly calibrated, then we need not be concerned with the foundation of the philosopher’s theory of truth. The philosophers have cottoned on to the folk notion! If it is not properly calibrated, then theories of truth have been built upon an incorrect assessment of [the importance of correspondence]. In either case, further philosophical theorizing about truth would clearly benefit from a better grasp of the folk notion of truth, one perhaps informed by empirical investigation. (2013: 623)

The basic picture they paint is that most everyone in the theory of truth agrees that there is something intuitive to the idea of correspondence, something that non-philosophers also find intuitive. As a result, theories of truth should try to respect the correspondence idea, either by building it into their theories or adequately explaining it away.⁸ But the question of whether correspondence really is intuitive to philosophers and non-philosophers alike is an empirical one, and so the theory of truth ignores it at its own peril.

Hence, even if facts about conceptions don’t directly entail facts about concepts, they do impose dialectical constraints when formulating theories about those concepts. If certain philosophical theories (or the motivations for believing them) depend upon certain empirical assumptions, then undermining those assumptions with empirical data undermines the relevant theories.⁹ As I highlight below, precisely that sort of critique is what fuels Næss’s research. Whether there are more constructive theses about *truth* and TRUTH to be found in the empirical data remains to be seen. In my critique below, I challenge the philosophical conclusions about truth that Barnard and Ulatowski draw from Næss’s research. I argue that we should be far more circumspect about the inferences we make from the available data.

⁸ This attitude is evident in both substantive accounts of truth (e.g., Asay 2013: 129-137 and Rasmussen 2014: 1-2) and deflationary accounts (e.g., Horwich 1998: 104-117).

⁹ This use of empirical data—see Reuter and Brun (forthcoming) for an example involving truth—is of a piece with what has been called the “negative program” of experimental philosophy (see, e.g., Alexander, Mallon, and Weinberg 2010), which comes in for criticism by Deutsch 2015.

3. Næss's 1938 study

In the 1930s Næss undertook extensive qualitative research into the “commonsense” notion of truth (1938a, 1938b).¹⁰ His research involved extensive interviewing of 300 subjects, none of whom was a professional philosopher. In addition to being asked to give examples of truths, subjects were queried, among other things, as to whether or not they thought there was a “common characteristic” of what is true, and if truth is absolute. In many cases, the research subjects “were allowed to associate freely,” and the leaders of the research were themselves novices on the topic of truth, lest they inadvertently lead the conversation in a particular way (Næss 1938b: 19). Reading through Næss’s presentation of his results—his paper (1938a) is a streamlined report of the far more comprehensive monograph (1938b)—provides a crash course in the challenges of systematizing and regimenting hundreds of hours of qualitative interviews of research subjects speaking extemporaneously into quantitative data amenable to statistical analysis. On the whole, Næss found that all of the views regarding the theory of truth discussed by philosophers could be found among the lay. Furthermore, Næss writes, “Our material does not support the statement that everybody, or the majority, means the same or approximately the same. It is therefore nonsensical to speak of *the* common sense view of the truth-notion. Equally nonsensical it is to speak of *the* view of the man in the street, of the uneducated, of the prephilosophic mind etc.” (1938b: 85).

Næss was spurred into his research by noticing a widespread tendency among truth theorists of his day writing as if common sense speaks with a single voice regarding the nature of truth. Noting that philosophers often make claims about what non-philosophers think about truth, he wonders: “How do the philosophers *know* these things? What is the source of their knowledge? What have they done to arrive at it? [...] Perhaps some of them have asked their wives or assistants for their opinions on the truth-notion, but there is very little to prove that they actually employed such a method” (1938b: 15; cf. 1938a: 40). Non-experimental philosophers, in other words, typically rely on their own beliefs and intuitions when it comes to claims about the ordinary understanding of truth. At best, they consult with their fellow philosophers, or, if we’re lucky, their wives and assistants. (Apparently philosophers don’t have husbands.) Experimental philosophers, by contrast, consult experimental studies to inform their beliefs about the ordinary notion of truth.

¹⁰ For short but highly critical reviews of Næss 1938b see Moore 1939 and Nagel 1939.

Næss's distaste for philosophers making claims about the folk, at least in the absence of any accompanying empirical evidence, is more than evident.¹¹ As is clear, Næss thinks his study reveals that there is no empirical basis for the idea that there is a single, prevalent commonsense theory or conception of truth. Philosophers who make claims that imply that there is such a view, as a result, appear to be saying something false. It may be that Næss valued his study more for its negative implications for current philosophical practice than for its positive value in informing us about how people actually think about truth. Chapman writes that Næss's investigations, including his 1938 study, "were negative or at least open-ended in their conclusions. Naess demonstrated that philosophers had been wrong or over hasty in their pronouncements about ordinary uses of language, and argued that words were often more fluid in their meanings, and more apt to variation depending on context and circumstance, than philosophers had allowed" (2008: 127). On Næss's view, philosophical practice presupposes that natural language is far more precise and unambiguous than empirical studies demonstrate, and that this problematizes much of what philosophers attempt to argue for.

Næss's study, then, provides evidence that there are myriad conceptions of truth to be found among non-philosophers, and that none is deserving of the title 'folk conception of truth'. If that is correct—and I am not here questioning Næss's analysis that denies that any conception was the majority view—then we have evidence that there is no such thing as *the* folk conception of truth.¹² Philosophers should therefore downplay the relevance of the "ordinary notion" of truth in their theories, when that phrase is understood as referring to conceptions of truth. If there is no robust folk conception of truth, then there is no philosophical theory that gains support by conforming to it, or that incurs a theoretical debt by denying it.

What is less clear is if there is anything positive to say with respect to truth itself. In particular, Næss's findings do not support his "pluralist" conclusion that his study tells against the view that "everybody, or the majority, means the same or approximately the same" by 'truth' (1938b: 85). To claim that two people mean something different in their use of a word is to claim that they express different concepts with their competing uses. So it appears that Næss thinks his data suggest some sort of *conceptual* diversity within his subjects: the differing *conceptions* that arise from their reflections

¹¹ And stable: the same attitude is present when he returns to these results decades later in Næss 1981.

¹² A more thorough defense of Næss's claim here would require an evaluation of Næss's standards for grouping individuals' responses into the relevant categories, modifying those standards if necessary, and then assessing the frequency of the different groups. That lies beyond the scope of the present paper, but it's safe to say that Næss's own analysis revealed no close contender for a majority view.

on truth are said to demonstrate that there are different *concepts* underlying people's use of 'truth'. But the fact that people subscribe to competing theories of truth does not entail that they express different concepts by their use of the word 'truth'. If that were the case, we would have overwhelming evidence that deflationists, verificationists, correspondence theorists, and others have been failing to communicate over the last century: rather than disagreeing over the nature of a single thing, they have simply all been talking past one another. The more modest conclusion, which I am inclined to draw, is that there is a considerable amount of *theoretical* diversity about truth within the folk, but that we cannot take this to establish that there is, in addition, significant *conceptual* diversity alongside it. After all, the presence of conceptual diversity undermines the possibility of theoretical diversity.

In contrast with my own modest assessment, Barnard and Ulatowski draw fairly significant conclusions from Næss's 1938 studies. It's impossible to draw sober and defensible conclusions from Næss's work without being absolutely clear on the distinction between truth itself and theories and conceptions regarding its nature. For that reason, I urge that we avoid talking about *notions* of truth, as this way of speaking elides the difference: 'notion' can serve as a synonym for either 'concept' or 'conception'. As a result, the phrase 'ordinary notion of truth' might be interpreted as referring to TRUTH, the ordinary concept of concern to philosophers and non-philosophers for millennia, or it might be interpreted as referring to a standard, dominant *conception* of the nature of truth (whose existence Næss provides evidence against). Barnard and Ulatowski at one point present the results of Næss's study by reporting that "The data yielded a wide array of truth notions operative among ordinary people" (2013: 625). This statement could be read either as the modest claim that Næss found that various people expressed competing *theories* or *beliefs* about the nature of truth when asked, or as the far more surprising and radical claim that Næss found that different people have different concepts underwriting their use of 'truth', such that there are "multiple concepts of truth" out in the wild. The former is a straightforward extrapolation of what Næss uncovered; the latter claim is far more theoretically loaded, and not even obviously coherent. If T_1 and T_2 are distinct concepts, such that X expresses T_1 by their use of 'truth' and Y expresses T_2 by their use of 'truth', they can't both be TRUTH. If they are "versions" of something else, then maybe that something else is TRUTH, in which case neither T_1 nor T_2 is TRUTH.

Furthermore, Ulatowski contends that "Not only is there no clearly unified folk theory but also Næss' results do not even reveal a dominant folk account. Therefore, we have not discovered in Næss' work data that perspicuously identifies a unique ordinary, common-sense, everyday, average concept of truth" (2016: 82). Notice the inference to there being no unique *concept* of truth from the

fact that Næss finds a multiplicity of folk *theories* of truth. Later, Ulatowski writes that “Næss’ results fail to settle the question of what *the* ordinary concept of truth could be. Indeed, Næss’ results show that there is no single folk-concept of truth against which to test Tarski’s assumptions of the classical theory. The “folk” of Næss’ study seem to hold every view ever imagined by philosophers” (2016: 85). Indeed, they might. But so do philosophers, and we should not then infer that philosophers from competing camps all have different concepts of truth, and so are all talking past one another. (Maybe they are, but we shouldn’t infer that simply from the reality of apparent peer disagreement.) Again, Ulatowski is concluding that there isn’t a shared concept of truth within a population because its members hold competing views about the phenomenon associated with that concept. To the contrary, the shared concept is essential if we are to interpret the population as engaged in disagreement over its subject matter. On my view, then, Ulatowski (like Næss) misidentifies the nature of the very real plurality that Næss has uncovered.

My takeaway is that Næss’s study provides useful insight into the sociology of views about truth, and whether or not there is such a thing as a robust folk conception of truth. It provides evidence that there is no robustly unified ordinary theory of the nature of truth. Barnard and Ulatowski, following Næss, take Næss’s findings also to provide evidence against the claim that there is a shared concept TRUTH among people, such that there exists some degree of *conceptual* plurality about truth.¹³ Disentangling the claim that two subjects have different concepts from the claim that they share competing views regarding a single shared concept is no simple matter; what is simple is that we cannot infer the former directly from the existence of theoretical disagreement. Næss’s subjects clearly expressed different theories and ideas when asked to reflect on the word ‘truth’.¹⁴ The explanation for that fact *might* be that they actually have different concepts (and so are giving

¹³ Ulatowski (2017) motivates his own view, which he labels ‘endoxic alethic pluralism’, in part by looking to Næss’s results. It is notably different from the more familiar forms of alethic pluralism (e.g., Wright 1992 and Lynch 2009) which focus on the *metaphysical* plurality of there being multiple *properties* of truth.

¹⁴ We should bear in mind that it’s possible that some of Næss’s subjects didn’t have *any* conception of truth until asked to reflect upon the topic during the study. As Næss writes: “The reader may ask whether one may look upon the definitions [of truth offered by the test subjects] as *real opinions* of the test-persons or whether they are mere >>Einfälle>>, mere >>words>> occurring to them during examinations” (1938a: 49). Indeed, a person possessing some concept may have *no* theory about the nature of the phenomenon associated with it, but still be able to come up with something on the fly when questioned. (See also Crockett 1959: 109.) In fact, Næss points out that he has “never heard a non-philosopher state something similar to a “definition of truth” without being urged” (1938b: 18; cf. 1938a: 43). Probing subjects for their theory of truth in an experimental setting may well be bringing those theories into existence.

compatible theories of different concepts, not competing theories of the same concept), but that interpretation goes far beyond what one can claim Næss to have shown. A simpler explanation, and one which doesn't have to insist that the presumably disagreeing parties are in fact talking past one another, is that they have different theories about the phenomenon associated with a concept they all share—whether deeply ingrained or developed on the spot.

4. Næss's 1953 study

4.1. Findings

Næss returned to his empirical research on truth in the 1950s, as presented in his *An Empirical Study of the Expressions "True", "Perfectly Certain" and "Extremely Probable"* (1953a). The goal of this study is to explore when research subjects find various kinds of sentences synonymous. The cases of synonymy to be tested are ones that Næss thinks are relevant to ongoing philosophical debates. "In contemporary philosophical literature," Næss writes, "questions are raised and answered which admittedly are empirical. Why not try to test the answers by procedures used in contemporary science?" (1953a: 5). Næss proceeds by providing research subjects with questionnaires that ask them to judge various combinations of sentences as being synonymous or not. Tellingly, he titles the questionnaires "ST" for "the questionnaire concerning the semantical concept of truth" (1953a: 6). This suggests that Næss takes his research to be engaging claims made by Tarski (1944, 1956). In fact, Næss never mentions Tarski in the entire study, save for a reference to him in the bibliography. Næss leaves it as an exercise to the reader to determine how his results shed light on Tarski's various assertions and theories. His own analysis is limited to the hypotheses involving synonymy he proposed for the study, and their confirmation or refutation.¹⁵

Here is how the study took place. Næss presented the ST survey, originally in Norwegian, to 129 students at the University of Oslo. The students were asked to evaluate "whether certain sentences are expressive of the same or of different assertions" (1953a: 37). They were "invited" to use the following criteria for individuating assertions:

¹⁵ Næss was concerned with the nature of synonymy for its own sake, and its implications for philosophy. A major theme of his monograph *Interpretation and Preciseness*, which reports on the greater research project to which the 1953 study on truth belongs, is that sameness of meaning—which Næss painstakingly endeavors to study empirically—is far more elusive than philosophers imagine, whether we are concerned with interpersonal synonymy or even *intrapersonal* synonymy. (See, for instance, his study of the Soviet writer Zaslavski and his use of 'démocratie' (1953b: 300-334).) Regarding inter- and intrapersonal synonymy with respect to 'truth', see Næss 1953b: 237-238 and 258-264.

A sentence A is for you expressive of a different assertion from that of another sentence B, if and only if you can imagine possible (but perhaps not actual) circumstances (conditions, existing state of affairs) of such a kind that if they were present you would accept A as warranted but reject B, or vice versa.

A sentence A expresses for you one and the same assertion as another sentence B, if and only if you can not imagine such circumstances, that is, if you under all conditions whatsoever either would accept both A and B as warranted, or reject both A and B. (1953a: 38).

The survey instructions make clear that the relevant notion of possibility is logical, not physical. Students' imaginations are allowed to flout natural laws, but not logical laws (1953a: 38).

With these instructions in hand, students were left to evaluate the synonymy (or lack thereof) of various groups of sentences. Simplifying somewhat, students were presented with eight different clusters of four assertions that (mostly) shared a common form, and were asked to judge whether they detected any synonymy among the four assertions.¹⁶ For each of the eight clusters, sentence (A) featured no truth-theoretic terminology, (B) featured 'true', (C) featured 'perfectly certain', and (D) featured 'extremely probable'. For example, cluster (1) was presented as follows:

- (1A) There is at least one copy of the Bible in the University Library.
- (1B) It is true that there is at least one copy of the Bible in the University Library.
- (1C) It is perfectly certain that there is at least one copy of the Bible in the University Library.
- (1D) It is extremely probable that there is at least one copy of the Bible in the University Library.

Students were asked to find any pairs of these four sentences that they took to be expressive of the same assertion. Clusters 2 through 6 follow this same format, though 2 and 5 involve negation. Here are the "ground-level" (A) assertions for each of the other clusters; the other three relevant assertions can be derived by modeling them on cluster 1.

¹⁶ The simplification is due to the fact that two clusters, those featuring negation, had seven assertions. The results involving these extra assertions were lumped together with the results for the four basic ones in Næss's analysis.

- (2A) There is no copy of the Bible in the University Library.
- (3A) It will be raining on May 17.
- (4A) It will be raining tomorrow.¹⁷
- (5A) It will not be raining on May 17.
- (6A) There is an ether that oscillates in accordance with the following laws—(here you may imagine a complete formulation of “the wave theory of light” or closely related theories).

As for clusters 7 and 8, they do not conform to this format in a few significant ways:

- (7A) (Here you may imagine that the selection theory of Darwin be formulated.)
- (7B) Darwin’s selection theory is true.
- (7C) Darwin’s selection theory is perfectly certain.
- (7D) Darwin’s selection theory is extremely probable.

- (8A) S
- (8B) S is true.
- (8C) S is perfectly certain.
- (8D) S is extremely probable.

One difference Næss discusses is the fact that these clusters don’t provide explicit content at the ground level. What Næss does not acknowledge is that clusters 7 and 8 shift from using *operators* (‘It is true that’) to *predicates* (‘is true’). Furthermore, any kind of systematic substitution within cluster 8 would yield nonsensical results: A sentence needs to be *used* in (8A) to make it intelligible, but *mentioned* in the others to make them intelligible. (Whether this last detail made any impact on the research subjects is impossible to know; perhaps only a nit-picking truth theorist would notice it.)

¹⁷ Technically there is no cluster 4. Question 4 asked participants if they would change their answers regarding cluster 3 if ‘May 17’ were replaced by ‘tomorrow’. 80% said ‘no’, 14% said ‘yes’, and 6% didn’t answer (Næss 1953a: 22). Because of the different format to the question, there is no column for cluster 4 in Table 1 below.

The table below reproduces Table III from Næss’s paper (1953a: 12), which collects the results, in percentages, of the questionnaire. ‘*n*’ is the number of respondents for each question; ‘=’ stands for ‘is synonymous with’ (in the sense corresponding to Næss’s instructions).

Cluster:	1 <i>n</i> : 126	2 <i>n</i> : 129	3 <i>n</i> : 114	5 <i>n</i> : 128	6 <i>n</i> : 118	7 <i>n</i> : 112	8 <i>n</i> : 118
(A) = (B)	87	75	81	69	85	44	64
(A) = (C)	84	7	78	5	80	41	60
(A) = (D)	10	4	21	4	17	16	20
(B) = (C)	91	7	91	8	87	87	89
(B) = (D)	11	5	17	5	12	9	15
(C) = (D)	9	26	17	33	12	8	14
(A) = (B) = (C)	83	6	77	5	77	40	59
(A) = (B) = (D)	10	4	16	4	12	8	14
(A) = (C) = (D)	9	4	16	3	12	7	13
(B) = (C) = (D)	9	4	17	5	11	8	13
(A) = (B) = (C) = (D)	9	4	16	3	11	7	12

Table 1: overall results

The (A) versions, recall, are non-truth-theoretic, while the (B) versions appended ‘It is true that’ (or ‘is true’ in clusters 7 and 8), the (C) versions appended ‘It is perfectly certain that’, and the (D) versions ‘It is extremely probable that’. Hence, for example, 87% of respondents judged that ‘There is at least one copy of the Bible in the University Library’ expresses the same assertion as ‘It is true that there is at least one copy of the Bible in the University Library’, whereas only 10% judged the former to express the same thing as ‘It is extremely probable that there is at least one copy of the Bible in the University Library’.

4.2. Analysis

I first present what I take to be significant from Næss’s findings, and what they do and don’t show for the theory of truth. Then I turn to a critical review of the interpretations offered by Barnard

and Ulatowski. Næss himself, recall, doesn't comment on the relevance of his results for the philosophical study of truth.

The first thing to notice is the kind of synonymy at work in the study, often expressed in terms of “expressing the same assertion”.¹⁸ Philosophers will be quick to notice that Næss's instructions, on a straightforward interpretation, capture something closer to co-assertibility rather than synonymy. Any two logically equivalent sentences ('7 + 5 = 12' and 'Red is a color', say) are “synonymous” in the sense of the study.¹⁹ Participants are asked to find pairs where *warrant* coincides, and that means they are on the lookout for a relation weaker than synonymy. This is not an objection to Næss's study. Recall that the broader research project here is concerned with the vagaries of the notion of synonymy itself, and how presuppositions concerning it impact philosophical methodology. How one should regiment the notion of synonymy for purposes of empirical investigation is a question to which Næss is quite sensitive, and one he tackles in detail in his work (see particularly his 1953b and 1956/1958). The important point to bear in mind is that Næss's study is not obviously a tool for detecting synonymy in the sense perhaps most familiar to contemporary philosophy.

Next consider the consistently high levels of synonymy between (A), (B), and (C) sentences in clusters 1, 3, and 6.

	(A) = (B) (%)	(A) = (C) (%)	(A) = (B) = (C) (%)
Cluster 1	87	84	83
Cluster 3	81	78	77
Cluster 6	85	80	77

Table 2: percentage of (A)/(B)/(C) synonymy judgments in clusters 1, 3, and 6

These show that for a strong majority of respondents, appending either 'It is true that' or 'It is perfectly certain that' to a sentence doesn't change the assertion in question. One hypothesis that might come to mind here is that 'true' and 'perfectly certain' are taken to mean the same thing by the respondents. Yet the negated clusters 2 and 5 refute that interpretation, given that the (B)/(C) synonymy judgments there are among the lowest in the whole study.

¹⁸ Synonymy is also emphasized in Toulmin's (1956) brief review.

¹⁹ See also Mates 1950: 215, and Næss 1956/1958 for his rebuttal.

Mean percentage of (B) = (C) judgments (all clusters)	65.7 (SD ²⁰ = 39.8)
Percentage of (B) = (C) judgments in cluster 2	7 (1.48 SD below mean)
Percentage of (B) = (C) judgments in cluster 5	8 (1.45 SD below mean)

Table 3: comparing (B) = (C) judgments in negated clusters to the mean across all clusters

‘True’ and ‘perfectly certain’ don’t mean the same if ‘not true’ and ‘not perfectly certain’ don’t.²¹ But the (A)/(B) synonymy judgments remain high in the negated clusters:

Mean percentage of (A) = (B) judgments (all clusters)	72.1 (SD = 14.9)
Percentage of (A) = (B) judgments in Cluster 2	75 (0.19 SD above mean)
Percentage of (A) = (B) judgments in Cluster 5	69 (0.21 SD below mean)

Table 4: comparing (A) = (B) judgments in negated clusters to the mean across all clusters

What this suggests is that respondents are strongly inclined to uphold (A)/(B) synonymy across negated and non-negated cases, and uphold (B)/(C) synonymy only in non-negated cases.

These results perfectly cohere with my claim that Næss’s subjects were looking for co-assertibility. Being in a position to properly assert that something is not perfectly certain is weaker than being in a position to properly assert that something is not true. It might be presumptuous to declare that (it’s true that) my lottery ticket is not a winner, but perfectly fine to assert that it’s not perfectly certain that it’s not a winner. The standards for assertibility between ‘not true’ and ‘not perfectly certain’ are quite different. But the respondents seem to be telling us that the criterion for asserting that something is true is perfect certainty. That’s why there is a strong pull to judge that (A) = (B) = (C) in non-negated cases. If you are in a position to assert one of them, you are in a position to assert them all. On this view, the operant norm of assertion is perfect certainty. For example, the

²⁰ ‘SD’ abbreviates ‘standard deviation’.

²¹ Næss might dispute this claim (which I must admit I find incontrovertible). He writes: “The proportion of persons who do not equate untruth and uncertainty but equate truth and certainty, is large” (1953a: 27). If ‘equate’ means ‘judge to be expressive of the same assertion as indicated in the directions’, then this is a straightforward report of the data. But if it means ‘equate’, then I can’t imagine coherently interpreting the subjects this way. If ‘true’ and ‘perfectly certain’ are two names for the same thing, then ‘not true’ and ‘not perfectly certain’ are two names for the lack of that thing.

respondents seem to be saying that you may assert ‘My ticket is a loser’ (as well as ‘It’s true that my ticket is a loser’) if and only if you are also in a position to assert ‘It’s perfectly certain that my ticket is a loser’. (This matches typical presentations of the lottery paradox—we lack certainty that we will lose, and this seems to ground our disinclination to claim definitively that we will lose, despite the high probability.) This interpretation shouldn’t be that surprising, given that Næss frames the instructions in terms of *warranted* acceptance, thereby inviting epistemic considerations into the semantics-focused project.

When we look at the first six clusters, we see a strong tendency to judge that $(A) = (B)$, a strong tendency to judge that $(A) = (B) = (C)$ in non-negated cases, and a strong tendency to judge that $(B) \neq (C)$ in negated cases:

	(A) = (B) (%)	(A) = (B) = (C) (%)	(B) ≠ (C) (%)
<i>Non-negated clusters</i>			
Cluster 1	87	83	9
Cluster 3	81	77	9
Cluster 6	85	77	13
<i>Negated clusters</i>			
Cluster 2	75	6	93
Cluster 5	69	5	92

Table 5: comparing (A), (B), and (C) judgments between negated and non-negated clusters

This last result reveals that ‘It is true that’ and ‘It is perfectly certain that’ are not taken to mean the same thing; so the cases where it’s judged that $(B) = (C)$ indicate the presence of something other than synonymy, namely, co-assertibility, paired with a ‘perfectly certain’ norm of assertion. If that norm is in place, then (A), (B), and (C) are all co-assertible for non-negations: if you’re able to assert that you’re perfectly certain that p , then you’re able to assert that p and that it’s true that p . But if you’re *not* able to assert that you’re perfectly certain that p , then you’re not able to assert that p or that it’s true that p . Hence, my explanation here predicts that respondents would find ‘My lottery ticket is a loser, though

I'm not perfectly certain that it is' as awkward and/or unassertible (or perhaps "Moorean-paradoxical").²²

Finally, though (B) and (C) come apart via negation, there remains a strong tendency to connect (A) and (B) across a variety of contexts. On my interpretation, this is strong evidence that respondents judge 'p' and 'It is true that p' to be co-assertible. They *might* take them to be synonymous as well, but that would require further investigation.

But now look to cluster 7. Here the (A) = (B) judgments fall sharply, to below a majority.

Mean percentage of (A) = (B) judgments (all clusters)	72.1 (SD = 14.9)
Percentage of (A) = (B) judgments in cluster 7	44 (1.88 SD below the mean)

Table 6: comparing (A) = (B) judgments in cluster 7 to the mean across all clusters

As noted above, the format for this cluster differs from the earlier cases in at least two significant ways. The target sentences, again, are:

- (7A) (Here you may imagine that the selection theory of Darwin be formulated.)
- (7B) Darwin's selection theory is true.

Notably, (7A) doesn't give the subjects a specific prompt to work with, and we don't know how in fact they interpreted it. To be consistent with the intention behind the earlier (A)/(B) pairs, (7A) and (7B) ought to differ only in having or lacking an instance of 'true'. One way to do this would be for

²² It might be objected here that this reading imposes an infallibilist understanding of knowledge on the subjects. If there is a knowledge norm of assertion present, and it's operating in conjunction with a perfect certainty norm, that indicates a commitment to knowledge requiring perfect certainty. I for one wouldn't find that surprising, and it's a worthwhile empirical enterprise to detect which norms of assertion are operant in the wild. (See Turri 2016 and 2017.) My own speculation here fits well with the sentiment David Lewis expresses when he writes: "it seems as if knowledge must be by definition infallible. If you claim that *S* knows that *P*, and yet you grant that *S* cannot eliminate a certain possibility in which not-*P*, it certainly seems as if you have granted that *S* does not after all know that *P*. To speak of fallible knowledge, of knowledge despite uneliminated possibilities of error, just *sounds* contradictory" (1996: 549). I would add, furthermore, that the study drives the participants into a high-standards epistemic context, where we should expect more stringent epistemic norms to be in force, if contextualists are correct.

the subject to stipulate that Darwin’s selection theory is identical to, say, ‘Species evolve by natural selection’. Then the subject could rewrite (7A) as:

(7A*) Species evolve by natural selection.

The choices for (7B) are now multiple, depending on how closely subjects chose to adhere to conformity with the earlier prompts, or with the current prompt. They could have left it alone, or could have rewritten it as:

(7B*) It is true that species evolve by natural selection.

(7B**) ‘Species evolve by natural selection’ is true.

In my view (which I explain below), (7A*) and (7B) are certainly not synonymous; whether they are co-assertible or not depends on if one takes semantic facts to be fixed in the realm of possibilities imagined. Furthermore, I think that (7A*) is co-assertible with (7B*), but not necessarily with (7B**), again depending on what we hold fixed. So I can find “disagreement” within myself regarding cluster 7, depending on how I choose to resolve the uncertainties introduced by Næss’s format. Finally, it’s not difficult to imagine some subjects interpreting (7A) along the lines of:

(7A**) Darwin’s selection theory is that species evolve by means of natural selection.

If that is what they had in mind, then they’d be right to judge that (7A**) and (7B) are not expressive of the same assertion, since one is saying what the theory is, and the other is asserting it.²³ Given that there are multiple ways that subjects may have responded to the uncertainties in Næss’s prompt, it’s reasonable to think that it was multiply interpreted.

Let me now explain why the different wording between (7B*) and (7B**) is crucial to judgments of synonymy. Earlier I pointed out that clusters 7 and 8 shift from operators to predicates,

²³ Cf. Næss: “It seems that some subjects have come to the conclusion that in question 7A, the theory of Darwin is supposed to be formulated but not affirmed, whereas in question 6A a theory or hypothesis is not only formulated but also affirmed” (1953a: 13). Whether this seeming is due to Næss’s following up with participants or just speculation to explain the divergence is unknown.

a shift that effects a crucial difference for synonymy and co-assertibility. Consider the following three sentences:

- (G) Tigers are mammals.
- (O) It is true that tigers are mammals.
- (P) 'Tigers are mammals' is true.

Philosophers divide over the relationship between them. For a pure disquotationalist (e.g., Field 1994), all three are synonymous, and so co-assertible as well. For the primitivist (e.g., Asay 2013), (G) and (O) are synonymous with each other, but not (P). This is because while (G) and (O) are about the biological classification of a certain species of large cat, (P) is about the alethic status of an English sentence. They may nevertheless be co-assertible, but they do not mean the same thing. Still others (e.g., Kühne 2003: 51) dispute the synonymy between (G) and (O). I do not expect, of course, that non-philosophers (or even most philosophers) have nuanced views about the effects of truth operators versus predicates on sentence meaning. But the shift from operators to predicates in clusters 7 and 8 introduces a potentially confounding factor, especially when paired with the lack of clarity regarding whether participants should be searching for synonymy or co-assertibility.

It's worth noting that Næss himself was more skeptical of the results found in clusters 7 and 8 (and 6 as well): "less weight should be attached to the statistics referring to them" (1953a: 7). Between-subject agreement was the smallest in clusters 7 and 8, suggesting to Næss that "There is evidence that the wording of those questions is bad, and that answers reflect differences of interpretation" (1953a: 12). Furthermore, Næss observes that "The wording of question 7A is unfortunate, and caused uncertainty and misinterpretation among the test persons" (1953a: 18).

Overall, I see Næss's data as most directly revealing what norms of assertion were operant in the epistemic context occupied by his student subjects. The (B)/(C) differences in negated and non-negated cases show that subjects weren't tracking synonymy, and so the best way of explaining the (A) = (C) data in non-negated cases is that '=' is actually co-assertibility. Despite the name of the questionnaire, there is no obvious connection to Tarski's work to be drawn, as I will explain below. Nor does there seem to be much to learn as regards the nature of truth, save for some support for the idea that appending 'It is true that' to sentences doesn't change their assertibility. (I'm not sure any theory of truth would suggest otherwise. The "transparency" of truth isn't disputed by non-deflationists.)

4.3. Barnard and Ulatowski's analysis

Let's now consider the more expansive conclusions reached by Barnard and Ulatowski, and what implications they think the study holds for the philosophical investigation of truth. First, they believe that it has mixed results for Tarski. Second, they think the study (like the 1938 one) bolsters the case for some form of alethic pluralism. They report that the study (in conjunction with Næss's earlier work) "reveals a lack of consensus on the nature of truth," and "calls into question the assumption that truth is a singular and simple monolithic notion" (Barnard and Ulatowski 2021: S720, S730).

Let's first consider the implications (or lack thereof) for Tarski. Barnard and Ulatowski describe Næss's study as "his work on intuitive responses to questions asking about Tarski's semantic conception of truth" (2019: 149), and claim that it shows that people have varying levels of acceptance of instances of the equivalence schema ' φ is true if and only if φ ', depending on the particular φ in question (2019: 152). As has been shown, however, this is not what the study does. I take it that Barnard and Ulatowski have in mind by 'Tarski's semantic conception of truth' what they describe as its orthodox interpretation, which combines "a formal and generalizable schema with the material adequacy condition for truth" (2019: 145). But it's clear that Næss's subjects were not asked about *that*. They were asked questions about whether various sentences exhibited something in the neighborhood of synonymy, and which (I've argued) is best thought of as co-assertibility. Most of the cases involving truth—and *all* of the non-problematic ones, even in Næss's own view—do not even take a form relevant to Tarski's schema. The biconditionals relevant to Tarski's project take the form ' X is true if and only if p ', where ' X ' is to be replaced by the name of a sentence of a given object language for which 'true' is being defined, and ' p ' replaced by a translation of that sentence into the metalanguage that is providing the definition. In the case of a natural language like English that can serve as its own metalanguage, ' X ' can be replaced by " p " (resulting in the familiar schema " p is true if and only if p ") since we can take advantage of our quote-name convention for sentences. These biconditionals are relevant to Tarski's project because he thinks an adequate theory of truth for a language needs to imply them for each sentence of that language. So what matters to Tarski is that these biconditionals follow from one's theory of truth.

Næss's study mainly asks subjects whether pairs of sentences of the form ' p ' and 'It is true that p ' are expressive of the same assertion. It's unclear, therefore, how to connect the data to Tarski's project. The pairs Næss asks about can't even be formed into the biconditionals relevant to Tarski. Again, Næss's questions (mostly) use the truth operator; Tarski's biconditionals involve a truth

predicate. As noted, many philosophers maintain that sentences of the form ‘*p*’ and ‘‘*p*’ is true’ are not synonymous—one is about whatever ‘*p*’ is about, and the other is about the sentence ‘*p*’. (This is especially clear in cases where the object language and metalanguage are distinct.) But they are *materially equivalent*, and so the material biconditionals formed from them are true. So I fail to find any relevance of Næss’s study to the evaluation of Tarski’s project, which makes no claims whatsoever regarding synonymy or assertibility.²⁴

What Barnard and Ulatowski might be thinking is that for Tarski’s project to succeed in capturing the ordinary concept of truth, the biconditionals need to be judged true by users of ‘true’. If the two sides of the biconditional are judged synonymous, then the biconditionals themselves would be judged true. If we then conflate ‘It is true that *p*’ and ‘‘*p*’ is true’, we can take Næss’s study to have shown that one possible route of falsifying Tarski’s project has not panned out. And perhaps that’s worth something. But notice that *even if* Næss had found deep reluctance in his subjects to detect synonymy or co-assertibility between ‘It is true that *p*’ and ‘*p*’, there wouldn’t have been a problem for Tarski. Tarski’s biconditionals are material, which implies that the relationship between them need not be so strong as synonymy or co-assertibility. Someone who judged ‘*p*’ and ‘*q*’ to not be co-assertible is not forced to conclude that ‘*p* if and only if *q*’ is false. All that follows is that the biconditional is not necessary. And Tarski’s biconditionals are not necessary. Consider:

‘Schnee ist weiß’ is true-in-German if and only if snow is white.

This biconditional is not necessary; had ‘Schnee ist weiß’ meant that seven is even, but snow remained white, it would be false. Nothing Næss studied puts the *truth* of the biconditionals at stake, and the truth of the biconditionals is the only factor in the neighborhood of being relevant to Tarski’s project.

Recall, however, that cluster 7 uses a truth predicate, and (A) = (B) synonymy judgments plummet there, as compared to clusters 1-6. (See again *Table 6*). Barnard and Ulatowski take the cluster 7 data to reveal a pretty striking fact, if it is one: “Næss’s study shows that people seem to think about a partial and informal rendering of Tarski’s semantic conception of truth differently, depending upon features of the expressions contained in [the equivalence schema]” (2019: 152; cf. Ulatowski 2018: 495). Their idea is that Næss’s research shows that people might accept the Tarski biconditionals for

²⁴ In the schema, the sentence ‘*p*’ needs to be a translation of *X*, so Tarski may be committed to ‘*p*’ being synonymous with *X*. But that’s not to say that ‘*p*’ is synonymous with ‘*X* is true’.

some sentences but not others, depending on their content. Næss had hypothesized that in “theory” contexts (i.e., clusters 6 and 7) people would be less likely to endorse (A)/(B) synonymy than in “matter of fact” contexts (i.e., clusters 1-5), but cluster 6 disconfirmed that hypothesis (1953a: 16-18). Plus, recall Næss’s repeated dissatisfaction with cluster 7, claiming that it was poorly worded and ambiguously interpreted by the subjects. Barnard and Ulatowski, by contrast, double down on the importance of cluster 7. They write: “There seems to be something peculiar having to do with Darwin’s evolutionary theory that doesn’t seem to apply to other scientific theories, such as the wave theory of light, in testing people’s intuitive responses to practical variants of Tarski’s semantic conception of truth” (2019: 153). In other words, their explanation for why (A)/(B) synonymy drops in cluster 7 is that people have hyper-specific understandings of the truth schema such that many take its instances to be false when it comes to some scientific theories (like Darwin’s) but not others. My explanation, by contrast, is that cluster 7 was confusing to subjects and ambiguously interpreted (Næss’s own view). Some may have rightly judged that ‘*p*’ and ‘*p* is true’ are not synonymous, while others may have interpreted the cluster to match the format of the earlier ones (a charitable exercise on their behalf, given Næss’s unfortunate conflation of operators and predicates). But again, all of the cluster 7 data are beside the point so far as Tarski is concerned. Suppose the data are perfect, and people are truly at odds regarding whether ‘*p*’ and ‘*p* is true’ express the same assertion for certain instances of ‘*p*’. That has no bearing on the question of whether people uphold the Tarski biconditionals. Whether the biconditionals are true or not doesn’t turn on whether or not their components are synonymous or co-assertible.²⁵

Moving beyond Tarski, Barnard and Ulatowski contend that “the collected empirical data seem to show that people’s intuitive responses regarding synonymy between “*p*” and “It is true that *p*” tends to fluctuate according to *p*’s specific content. So sensitive are people’s responses that they seem to recognize a distinction between different scientific theories” (2019: 154; cf. Ulatowski 2017: 76 and 99). This finding is then interpreted as giving support in some form to alethic pluralism, the idea that truth is not in some sense the same in all its applications. But as is clear from Næss’s results

²⁵ To illustrate somewhat evocatively, Barnard and Ulatowski seem to be inferring from ‘S judges that ‘*p*’ and ‘*q*’ are not synonymous’ to ‘S would judge that ‘*p* if and only if *q*’ is false’. Since I think the two sides of the Tarski biconditionals are not synonymous (and I highly suspect you, as well as Tarski, agree with me), I am inferred to be rejecting as false *all* of the Tarski biconditionals. That is certainly news to me. This line of thinking is made explicit at Barnard and Ulatowski 2019: 164.

in Table III, nothing of the sort is indicated. The operator cases—clusters 1-6—all show high levels of (A)/(B) synonymy.

	(A) = (B) (%)	SDs from Mean
Mean across all clusters: 72.1 (SD = 14.9)		
Cluster 1	87	0.99 above
Cluster 2	75	0.19 above
Cluster 3	81	0.59 above
Cluster 5	69	0.21 below
Cluster 6	85	0.86 above
Cluster 7	44	1.88 below
Cluster 8	64	0.54 below

Table 7: comparing (A) = (B) judgments in each cluster against the mean across all clusters

They're not *identical*, of course, but are all within a standard deviation of the mean. Barnard and Ulatowski take the differences between cluster 7 and the others to be the basis for their claim, but cluster 7 and its switch to the truth predicate is of an entirely different logical form than the others, and thus irrelevant to the exact conclusion they are drawing, which is formulated in terms of the truth operator. In cluster 7, the (A)/(B) sentences *aren't* synonymous (*pace* pure disquotationalists). This, in addition to its confusing nature, would seem to provide a straightforward and simple explanation for why synonymy judgments drop when it comes to cluster 7. Given this crucial difference between 7 and the others, and the potential sources of corruption in the cluster 7 data, there is no reason to see alethic pluralism as providing the best explanation for the numbers.²⁶

Summing up, Barnard and Ulatowski look to Næss's study to find an empirical footing regarding how non-philosophers would accept or reject Tarski biconditionals of the form “*p*’ is true if and only if *p*’”. They conclude that Næss finds reason to suppose that people are more or less inclined to accept the biconditionals, depending on the specific ‘*p*’ in question, and that this result indicates

²⁶ Barnard and Ulatowski replicated Næss's study, and avoided Næss's operator/predicate switch. They highlight that they still found a statistically significant difference between their versions of clusters 1 and 7 (2019: 159). But note that whereas Næss found (A)/(B) synonymy to drop from 87% to 44% between clusters 1 and 7 (a 49% decrease), Barnard and Ulatowski found a drop from 94.2% to 80.9% (a 14% decrease).

that ordinary people deploy a concept of truth that conforms to some form of alethic pluralism. I have lodged three main objections to this interpretation. First, testing for judgments of synonymy (or even co-assertibility) between the two sides of the biconditionals is an ineffective means for testing acceptance of the biconditionals, since their truth does not turn on the synonymy or co-assertibility of their components. (Moreover, the majority of synonymy test cases did not match the form of the two sides of the relevant biconditionals.) Second, the variance on which Barnard and Ulatowski base their claim (namely, the difference in results between clusters 1 and 7) corresponds not just to a variance in the particular '*p*' in question, but also a variance in both the structure and logical form of the synonymy test. The experiment, therefore, was not properly controlled. In cluster 1 subjects are asked to evaluate whether '*p*' and 'It is true that *p*' are synonymous (and they are, at least in many philosophers' views); in cluster 7 subjects are *at best* asked to evaluate whether '*p*' and "'*p*' is true' are synonymous (they aren't, and it's unclear what question the subjects even thought they were answering since no explicit content was provided). Finally, the data from cluster 7 are compromised, as Næss himself asserts several times. They cannot serve as the basis for Barnard and Ulatowski's radical claim that people's individual *concepts* of truth are so finely-grained and variant between any two people that they accept or reject instances of the truth schema on the basis of whether they involve the particular contents of various scientific theories.

5. Conclusion

Næss was correct, in my view, to wonder whether philosophers' claims concerning the "ordinary conception of truth" were empirically adequate. They are empirical claims, and thus subject to empirical appraisal. If Næss's work is reasonably representative, then non-philosophers (which is to say, the overwhelming majority of humanity) don't speak with a single voice when talking about truth—I doubt such accord is found regarding any topic of substantial or philosophical interest. Philosophers don't speak with a single voice, either, of course. And when we *say* different things about truth, that's evidence that we disagree about truth, that we have different philosophical beliefs and theories about it. It's not at all obvious, however, that those disagreements are evidence of the more provocative thesis that we lack a shared concept of truth. After all, to conclude that we had different concepts all along is to conclude that our disagreements have been illusory this entire time.

On my view, the familiar sort of data that experimental philosophers can and have assembled is relevant primarily to uncovering what people's conceptions of truth are.²⁷ Such data are welcome, and help regulate the empirical claims we may wish to make about which conceptions of truth are more or less intuitive or commonsensical. Whether those data are relevant to the traditional projects of understanding TRUTH and *truth* is another matter; the empirical traces of concepts and logical properties are far less obvious than are the empirical traces of conceptions and naturalistic properties. It's true that philosophers have paid lip service to a commitment to developing theories that closely hew to the folk view. Perhaps what we can learn from Næss is that there is no need for such a commitment: there is no folk view of truth, and therefore no constraint on philosophical theories of truth to conform to it.

I have urged extreme caution when it comes to using experimental data to draw conclusions that connect to longstanding philosophical theories about the nature of TRUTH and *truth*. Be that as it may, it doesn't follow that those data tell us nothing of philosophical interest.²⁸ Perhaps the theory of truth's focus on giving theories of the nature and constitution of the concept of truth and its associated property has been far too parochial. If indeed there is no dominant folk theory of truth, the significance of that fact need not be limited to its role in, say, undermining one plank of the correspondence theorist's platform. That there is considerable diversity in how people wield 'truth' in ordinary thought and speech is itself an important discovery, and one that raises myriad questions that can be further explored. If so, then it may be that the real value in the empirical study of truth is not that it helps us answer the age-old questions about the nature of truth, but rather that it enables us to formulate new questions about truth that we didn't even realize we should ask.

Consider, for example, Joshua Knobe's (2016) contention that most experimental philosophy is a kind of cognitive science. Rather than serving as a handmaiden and/or corrective to the projects of traditional philosophy, experimental philosophy's real value, for Knobe, "consists of identifying surprising effects in people's intuitions and explaining those effects in terms of underlying cognitive processes" (2016: 50). For example, Barnard and Ulatowski (2013) report findings of statistically significant variations in their respondents' answers due to gender. These data *might* be evidence for the claim that gender plays a role in what alethic concepts a person has; but even if they're not, they still

²⁷ Barnard and Ulatowski have also developed their own empirical studies covering a number of alethic issues, including attempts to replicate Næss's 1953 study (Barnard and Ulatowski 2013, 2019, and 2021). See also Reuter and Brun (forthcoming).

²⁸ Thanks go to a referee for the journal for pushing me on this point.

suggest the modest claim that “gender is among the factors that will influence how a person thinks about truth” (2013: 632). Even if such factors don’t help to settle the longstanding disputes between, say, deflationists and substantivists, or monists and pluralists, they can help us better understand what it is we are doing when we think about truth.

Acknowledgments

Many thanks go to Kristy Lam for her assistance in putting together the statistical analyses and tables for the paper. I’m grateful also to the referees of the journal for their helpful, constructive feedback. Thanks also go to Bob Barnard, Max Deutsch, and Joe Ulatowski for helpful discussion of the paper. The research presented here was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (project number HKU 17618420), and funding from the University of Hong Kong’s Outstanding Young Researcher Award.

References

- Alexander, Joshua, Ronald Mallon, and Jonathan M. Weinberg. 2010. Accentuate the negative. *Review of Philosophy and Psychology* 1: 297-314.
- Asay, Jamin. 2013. *The Primitivist Theory of Truth*. Cambridge: Cambridge University Press.
- Barnard, Robert, and Joseph Ulatowski. 2013. Truth, correspondence, and gender. *Review of Philosophy and Psychology* 4: 621-638.
- . 2016. Tarski’s 1944 polemical remarks and Naess’ “experimental philosophy”. *Erkenntnis* 81: 457-477.
- . 2019. Does anyone really think that ‘ $\ulcorner\varphi\urcorner$ is true if and only if φ ?’”. In *Advances in Experimental Philosophy of Logic and Mathematics*, eds. Andrew Aberdein and Matthew Inglis, 145-171. London: Bloomsbury Academic.
- . 2021. The objectivity of truth, a core truism? *Synthese* 198: S717-S733.
- Bar-On, Dorit, and Keith Simmons. 2007. The use of force against deflationism: assertion and truth. In *Truth and Speech Acts: Studies in the Philosophy of Language*, eds. Dirk Greimann and Geo Siegwart, 61-89. London: Routledge.
- Barsalou, Lawrence W., and Jesse J. Prinz. 1997. Mundane creativity in perceptual symbol systems. In *Creative Thought: An Investigation of Conceptual Structures and Processes*, eds. Thomas B. Ward, Steven M. Smith, and Jyotsna Vaid, 267-307. American Psychological Association.

- Chapman, Siobhan. 2008. *Language and Empiricism: After the Vienna Circle*. Houndmills: Palgrave Macmillan.
- . 2011. Arne Naess and empirical semantics. *Inquiry* 54: 18-30.
- . 2018. The experimental and the empirical: Arne Naess' statistical approach to philosophy. *British Journal for the History of Philosophy* 26: 961-981.
- Crockett, Campbell. 1959. An attack upon revelation in semantics. *Journal of Philosophy* 56: 103–111.
- Deutsch, Max. 2015. *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge: MIT Press.
- Field, Hartry. 1994. Deflationist views of meaning and content. *Mind* 103: 249-285.
- Greene, Joshua D. 2007. The secret joke of Kant's soul. In *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, 35-79. Cambridge: MIT Press.
- Horwich, Paul. 1998. *Truth*. Second edition. Oxford: Clarendon Press.
- Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press.
- Kauppinen, Antti. 2007. The rise and fall of experimental philosophy. *Philosophical Explorations* 10: 95-118.
- Knobe, Joshua. 2016. Experimental philosophy is cognitive science. In *A Companion to Experimental Philosophy*, eds. Justin Sytsma and Wesley Buckwalter, 37-52. Oxford: Wiley Blackwell.
- Künne, Wolfgang. 2003. *Conceptions of Truth*. Oxford: Clarendon Press.
- Lewis, David. 1996. Elusive knowledge. *Australasian Journal of Philosophy* 74: 549-567.
- Lynch, Michael P. 2009. *Truth as One and Many*. Oxford: Clarendon Press.
- Machery, Edouard. 2017. *Philosophy Within Its Proper Bounds*. Oxford: Oxford University Press.
- Mates, Benson. 1950. Synonymity. In *University of California Publications in Philosophy, 25: Meaning and Interpretation*, 201-226. Berkeley: University of California Press.
- Moore, Jared S. 1939. [Review of] "Truth" as Conceived by Those Who Are Not Professional Philosophers by Arne Ness. *American Journal of Psychology* 52: 489–490.
- Murphy, Taylor Shaw. 2014. Experimental philosophy 1935–1965. In *Oxford Studies in Experimental Philosophy, Volume 1*, eds. Tania Lombrozo, Joshua Knobe, and Shaun Nichols, 325–368. Oxford: Oxford University Press.
- Næss, Arne. 1938a. Common-sense and truth. *Theoria* 4: 39-58.
- . 1938b. "Truth" as Conceived by Those Who Are Not Professional Philosophers. Oslo: Jacob Dybwad.

- . 1953a. *An Empirical Study of the Expressions "True", "Perfectly Certain" and "Extremely Probable"*. Oslo: Jacob Dybwad.
- . 1953b. *Interpretation and Preciseness: A Contribution to the Theory of Communication*. Oslo: Jacob Dybwad.
- . 1956/1958. Logical equivalence, intentional isomorphism and synonymity as studied by questionnaires. *Synthese* 10a: 471-479.
- . 1981. The empirical semantics of key terms, phrases and sentences. In *Philosophy and Grammar: Papers on the Occasion of the Quincentennial of Uppsala University*, eds. Stig Kanger and Sven Öhman, 135-154. Dordrecht: D. Reidel.
- Nagel, Ernest. 1939. [Review of] *"Truth" as Conceived by Those Who Are Not Professional Philosophers* by Arne Ness. *Journal of Philosophy* 36: 78–80.
- Rasmussen, Joshua. 2014. *Defending the Correspondence Theory of Truth*. Cambridge: Cambridge University Press.
- Reuter, Kevin, and Georg Brun. Forthcoming. Empirical studies on *truth* and the project of re-engineering *truth*. *Pacific Philosophical Quarterly*.
- Tarski, Alfred. 1944. The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research* 4: 341-376.
- . 1956. The concept of truth in formalized language. In his *Logic, Semantics, Metamathematics: Papers From 1923 to 1938*, trans. J. H. Woodger, 152-278. Oxford: Clarendon Press.
- Toulmin, Stephen. 1956. [Review of] *An Empirical Study of the Expressions "True," "Perfectly Certain" and "Extremely Probable"* by Arne Naess. *Philosophical Review* 65: 116-118.
- Turri, John. 2016. Knowledge, certainty, and assertion. *Philosophical Psychology* 29: 293-299.
- . 2017. Experimental work on the norms of assertion. *Philosophy Compass* 12: 1-9.
- Ulatowski, Joseph. 2016. Ordinary truth in Tarski and Naess. In *Uncovering Facts and Values: Studies in Contemporary Epistemology and Political Philosophy*, eds. Adrian Kuźniar and Joanna Odrowąż-Sypniewska, 67-90. Leiden: Brill.
- . 2017. *Commonsense Pluralism About Truth: An Empirical Defence*. Cham: Palgrave Macmillan.
- . 2018. Is there a commonsense semantic conception of truth? *Philosophia* 46: 487-500.
- Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich. 2001. Normativity and epistemic intuitions. *Philosophical Topics* 29: 429-460.
- Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge: Harvard University Press.