**The Experimental and The Empirical: Arne Naess’s Statistical Approach to Philosophy[[1]](#footnote-1)**

Siobhan Chapman

University of Liverpool

**Abstract**

Experimental philosophy often draws its data from questionnaire-based surveys of ordinary intuitions. Its proponents are keen to identify antecedents in the work of philosophers who have referred to intuition and everyday understanding (e.g. Knobe and Nichols, ‘An experimental philosophy manifesto’). In this context, ‘Empirical Semantics’, pioneered by Arne Naess early in the twentieth century, offers striking parallels. Naess believed that much contemporary philosophy was unempirical because it made no reference to ordinary understanding or use of language. In response, he attempted large-scale studies of ordinary understanding of philosophically significant terms (e.g. Naess, *“Truth”*). This article will compare Empirical Semantics and experimental philosophy, and assess what such a comparison reveals that might not be apparent from a consideration of either in isolation. It will concentrate on two points of comparison: the wider philosophical ambitions of the two approaches, and the types of criticisms they have encountered. Naess was straightforwardly interested in folk intuitions as philosophical data, whereas many experimental philosophers are also interested in explaining the intuitions themselves. A comparison of the different types of criticism reveals attitudes to philosophical practice which have recurred across a time period of some eighty years, and those which are more specific to historical context.

**Keywords**

Arne Naess; Empirical Semantics; Experimental philosophy; Truth theories

**1. Introduction**

Beginning in the 1930s, the Norwegian philosopher Arne Naess investigated philosophically salient terms by collecting and analyzing ordinary, everyday intuitions. Using a methodology that he termed ‘Empirical Semantics’, he devised questionnaires and interview techniques and, together with a team of fellow researchers, used these to test subjects who were specifically not professional philosophers. The results of these tests were analysed both qualitatively and quantitatively. They formed the subject matter of his first major monograph, *“Truth” as Conceived by those who are not professional philosophers*, and continued to inform his philosophy of language for a number of decades.

Parallels between Naess’s work and much more recent developments in what has become known as ‘experimental philosophy’ are immediately apparent and have been remarked on by scholars. Barnard and Ulatowski (‘Tarski’s 1944 Polemical Remarks’, 460), for instance, note that Naess himself suggested the term ‘experimental philosophy’ to describe the type of future work he hoped to inspire, as does Appiah (‘Experimental philosophy’, 8). Murphy offers a detailed account of Empirical Semantics in its various manifestations, describing it as ‘deeply similar to today’s experimental philosophy’ (‘Experimental philosophy’, 2). There is no direct causal link between Empirical Semantics and experimental philosophy, but the similarities between the earlier and the more recent philosophical approaches are worth examining in some detail, not least because of what Naess’s undoubtedly pioneering work might now be able to contribute to very current philosophical debates.

In his comparison of the two schools of thought, Murphy considers Empirical Semantics as a whole, during the time period 1935-1965. He comments on its connections to present day experimental philosophy in terms of ‘subject matter’, ‘motivations and aims’ and ‘characteristic problems and objections’ (Experimental philosophy’, 2). In this article, I will consider just the work of Naess, rather than Empirical Semantics as a wider movement, and will in fact concentrate mainly on his 1938 study of the term ‘truth’. My main focuses will be on Naess’s wider philosophical ambitions, and the responses of his critics. That is, I will compare Empirical Semantics and experimental philosophy in relation to how their proponents have understood the relationship of their approach to the more mainstream work of their philosophical contemporaries, and their aspirations for the future of this relationship. I will also compare some of the often strongly negative contemporary reactions to what Naess was doing, with criticisms that have been encountered by experimental philosophy. Such a comparison shows that some objections arose separately in both the earlier and the current cases, while some are specific to the individual time periods. This in turn is suggestive of the differences between attitudes to philosophical practice which have changed, and those which have endured. Throughout the article, I will assess both the similarities and the differences between Empirical Semantics and experimental philosophy and suggest what might be gained from the comparison which would not be available from a consideration of either movement in isolation. Recent achievements in experimental philosophy may offer a new way of appreciating, evaluating and critiquing what Naess was trying to do. At the same time, Naess’s experiences, the decisions he made, and the ambitions he developed, have a contribution to make to present day debates concerning the nature, the purpose and the validity of experimental philosophy.

**2. Naess and Empirical Semantics**

Arne Naess’s philosophy of language, and the school of Empirical Semantics which he founded and led, have an important contribution to make to the story of mid- twentieth century analytic philosophy. In most appraisals of Naess’s legacy they are, however, overshadowed by his ecological writing and activism. In 1970, Naess resigned his chair at the University of Oslo and put aside his other philosophical interests, dedicating the rest of his long life to environmental philosophy. He coined the phrase ‘Deep Ecology’ to refer to a set of beliefs that see all life forms as equally important and worthwhile, regardless of their relationship to human beings, and that emphasise the essential inter-connectedness of all organisms. What was initially a small and rather marginal movement gained popularity and moved into the mainstream. By the time of his death in 2009, just short of his ninety seventh birthday, Naess was a celebrated figure of the Norwegian and international establishment, largely as a result of his work in ecology.

This pattern of pursuing unorthodox ideas and seeing them become established can be traced in the story of Naess’s philosophy of language, too. Empirical Semantics was in the ascendency in Norwegian philosophy for much of the middle part of the twentieth century, with all students of philosophy at the University of Oslo taking an exam in the subject as a compulsory part of their studies. But Naess had initially developed the ideas underlying Empirical Semantics in opposition to what he saw as some of the errors and flawed methodologies of contemporary academic philosophy, and in so doing had almost put an end to his own academic career before it had properly begun. His targets included many of the claims or the underlying assumptions of Western philosophy but he took particular issue with those he was in many ways philosophically closest to: the logical positivists of the Vienna Circle.

As a student, Naess studied both philosophy and the empirical sciences. He travelled to Austria in the early 1930s and attended meetings of the Vienna Circle. He was impressed by the commitment he found there to clarity of expression, to rigorous analysis, and to extending the methods of scientific study to all areas of inquiry. But there were other aspects of the movement that he found more troubling. He believed that logical positivism relied on a number of sweeping and unscientific assumptions, particularly assumptions about language. The principle of verificationism, for instance, depended on the idea that there were two types of meaningful statements. Analytic statements were meaningful because they could be judged true on the basis of their logical form, or of the meaning of the words they contained. Synthetic statements were meaningful if and only if a process of verification could be identified, by means of which they could be judged either ‘true’ or ‘false’ on empirical grounds. All other statements, regardless of whether they could be used and understood in ordinary communicative occurrences of language, were meaningless and therefore not amenable to serious philosophical scrutiny; any apparent issues introduced by an example such as ‘stealing is wrong’ were pseudo-problems because the example in question failed to make any meaningful statement.

For Naess, the doctrine of verificationism relied on theoretical judgments about meaning that were unsupportable, including judgments about the meaning of philosophically salient terms such as ‘true’ and ‘meaningful’. It valued *a priori* judgments of philosophers over any available evidence concerning the ordinary use and understanding of language. Recalling his reaction to logical positivists many decades later, Naess commented that: ‘They imagined they had perfect knowledge of ordinary language about their mother tongue. So, to me, they were antiempirical’ (Rothenberg, *Is it Painful to Think?* 28). In his search to establish a more empirical basis for philosophy, Naess drew inspiration from the then relatively new academic discipline of sociology. If philosophers are to study ordinary language, he reasoned, they should approach it in the same way as the study of any other type of social behaviour, using established methods of empirical investigation.

With a dedicated group of co-investigators, Naess began his empirical investigations of language in Oslo in the mid 1930s. He concentrated initially on the use of questionnaires to elicit information from participants about their use of linguistic expressions and their understanding of certain key terms. His initial focus, and the subject of his first major philosophical work, was the variety of everyday understandings and uses of the notion of ‘truth’. This is not a surprising choice, given the centrality of the concept of ‘truth’ to many concerns of philosophy, both ancient and modern. Moreover it was what Naess saw as the somewhat cavalier attitude to the meaning and applicability of the term ‘true’ that was largely responsible for his reaction against logical positivism. But Naess had more specific motivations, too. He had noticed that many philosophers were given to referring directly to ‘the opinion of the non-philosopher on the notion of truth’ or ‘the common-sense view of the notion of truth’, and to basing their philosophical claims on this construction, for instance insisting that their theory of truth had the advantage of according with everyday understanding, or rejecting a rival theory on the grounds that it went against common usage.

Naess was struck by the fact that, although the concept of ‘the common-sense view of the notion of truth’ had some currency in philosophy, there seemed to be little actual agreement as to what it comprised. It was generally taken to be in some way close to the correspondence account of truth, whereby a statement is true if it corresponds with the facts of the world, but there was little agreement as to its exact nature. Further, there was no evidence of philosophers using observation, or any empirical method, to establish what was the ordinary understanding of truth. It should be possible to establish how the term ‘truth’ is ordinarily used and understood, but instead successive philosophers were relying on unanalyzed and untested assumptions. ‘Perhaps some of them have asked their wives or assistants for their opinion of the truth-notion’, Naess suggested with faux naivety, but there was no concrete evidence of non-philosophers being consulted on the matter (*“Truth”,* 15).

The monograph *“Truth” as Conceived by those who are not professional philosophers* was published in Oslo in 1938. In it, Naess reported on, defended and analysed the results of a variety of different empirical studies. The question he set his participants in the first of these studies translates into English as ‘What is the common characteristic of that which is true?’ (Naess, *“Truth”*, 23). He followed this up with more specific questions about, for instance, examples of something that is true, and the possibility of something being absolutely true. He was particularly interested in using statistical analysis to make sense of his findings. So, for instance, he presented summaries of the proportion of respondents whose account of truth could be classified under general headings relating to ‘facts’, to ‘morals’ or to ‘tautologies’ (for more detail on the development of Naess’s research on truth, see Chapman, *Language and Empiricism*).

Naess’s primary and most controversial conclusions were as follows: far from there being a single entity, ‘the common-sense view of the notion of truth’, which was available for philosophers to use in their deliberations and arguments, ordinary people understood ‘truth’ in a wide variety of different ways. Further, and perhaps more strikingly, those understandings were often undistinguishable in substance from some of the most sophisticated theories of truth expounded by philosophers. Versions of the correspondence theory of truth could indeed be found, but so too could versions of theories of truth developed by pragmatists, by logical positivists and by relativists. A further conclusion suggested itself to Naess; ordinary people might be just as capable as professional philosophers of understanding sophisticated concepts, of discussing them and of defending them against objections. Philosophers had no basis for claiming any special access to philosophical understanding or aptitude for philosophical debate.

**3. Empirical Semantics and Experimental Philosophy: Ambitions**

Towards the end of his monograph on ‘truth’, Naess signaled what he saw as the potential longer-term and wider-ranging implications of his study. These amounted to no less than a new style of or approach to philosophy: ‘The diversity and consistency of amateur theories of truth, point to the possibility of an “experimental philosophy”’ (*“Truth”*, 161). Neither Naess nor his followers adopted this term as a title for their subsequent work, and it was not much heard of again until around the turn of the twenty-first century. There is no continuity in historical terms, then, between Naess’s work and experimental philosophy, nor is there any evidence of direct influence from Empirical Semantics on the development of the present-day movement. Nevertheless there are some striking similarities between ideas about methodology expressed in the two fields. These will be discussed in the next section, in relation to the types of criticism that the two approaches have encountered. There are both similarities and differences between the two in relation to the subject matters and aims of individual studies conducted, and the more general objectives and ambitions of the fields of study. These similarities and differences will be discussed in this section.

Analytic philosophy has long been concerned with the business of analyzing concepts. Philosophers have engaged in close consideration of human knowledge by means of attempting to produce formulas and descriptions as precise as possible of the meaning of certain terms or propositions. For some analytic philosophers, this type of conceptual analysis has involved close attention to the meaning and the use of certain words and expressions in ordinary language. The school of Ordinary Language Philosophy particularly associated with mid twentieth century Oxford, and with the work of J. L. Austin, for instance, attempted to introduce an empirical element to this type of study; the data under consideration could include not just individual introspective judgments, but also the outcome of collective discussion by a group of philosophers working together (see, for instance, ‘A plea for excuses’). In his own empirical work, Naess avoided the judgments of philosophers altogether, and sought data about ordinary meaning and usage from the philosophically lay residents of Oslo. Experimental philosophy, also, has the explicit aim of addressing specific questions about people’s everyday intuitions in relation to certain philosophically significant concepts. But its concern with intuitions goes further than Naess’s. Experimental philosophers may be interested in what folk intuitions suggest about philosophical contentious concepts. But they are interested also in finding out what psychological, cultural, social or other factors might cause people to make particular judgments, or to apply or withhold expressions in particular circumstances.

Joshua Knobe’s early, pioneering and much cited investigation of intuitions concerning intentional action provides a good example. Knobe was interested in the question of whether a predictable side effect of a form of behaviour could be described as being brought about intentionally. In relation to the question of ‘the proper analysis of the concept of intentional action’ he argued that, although not the only way to study the question, ‘the examination of ordinary language might provide us with some useful guidance’ (Knobe, ‘Intentional action, 190). Knobe presented his participants with vignettes. In each of these, a form of behaviour on the part of the main actor (in some cases the chairman of the board of a company, in some cases an army sergeant) produced side effects, some negative in nature and some positive. In each case, the participants were asked for two different types of responses: a scaled judgment as to how much blame (in the case of negative side effects) or praise (in the case of positive side effects) the actor deserved, and a binary judgment as to whether the actor brought about the side effect intentionally. Knobe was looking for data relevant to his hypothesis that ‘a person’s intuitions as to whether or not a given side effect was produced intentionally can be influenced by that person’s attitude toward the specific side effect in question’ (‘Intentional action’, 191). The data he collected appeared to support this hypothesis; his participants were more likely to judge a side effect as being produced intentionally if it was negative, and hence to judge the actor as blameworthy.

Since Knobe’s innovative study there have been many published reports of experimental work in philosophy. The methodological approaches taken in these studies vary widely, and there is no single or simple way of conducting experimental philosophy. Many studies, though, involve participants who are asked for their intuitive responses to certain vignettes, especially in relation to the terms they would or would not apply in different situations. Intentional action continues to be a major focus for such work, as too do causation, knowledge, free will and, in the philosophy of language, names and reference. Exclusively philosophical interests and issues, however, are not always paramount. Some experimental philosophers have engaged with more practical aims (see, for instance, Nadelhoffer, ‘Bad acts, blameworthy agents’). Others have emphasised the significance of the psychological processes underlying participants’ responses over that of philosophical problems themselves. Knobe has contrasted experimental philosophers with traditional philosophers who, when they have referred to intuitions, have not been interested in them for their own sake, but for their potential to be used to resolve philosophical questions. Experimental philosophers, however, are interested in ‘the internal processes that lead people to have the intuitions they do’, or in how people ordinarily understand the world (‘Experimental philosophy’, 90). More recently, he has emphasized further the difference between traditional philosophical analysis and experimental philosophy, arguing that the latter has its own, independent relevance and is best seen as a branch of cognitive science. Experimental philosophers are concerned with explaining intuitive responses in relation not just to individuals’ concepts, but to wider underlying cognitive processes (Knobe, ‘Experimental philosophy is cognitive science’).

In experimental philosophy, as also in Empirical Semantics, explicit emphasis is laid on the need for philosophers to become experimenters themselves: to devise their own experiments and to collect and analyse their own data. Sosa points out that philosophers have used scientific results for a long time, but what is novel is ‘that experimental philosophers do not so much borrow from the scientists as they become scientists.’ (‘Experimental philosophy and philosophical intuition’, 231). In a strikingly similar passage, several decades earlier, Naess argued that:

In contemporary philosophical literature questions are raised and answered which admittedly are empirical. Why not try to test the answers by procedures used in contemporary science? That is one way philosophy can be a mother science. But one cannot expect professional psychologists, physicists or others to do the job – they have their own favourite questions. The philosophically inclined must carry it out himself.

(Naess, *An Empirical Study*, 5).

But despite this similarity in motivation, there are some striking differences between the two approaches, particularly in relation to the topics investigated and the objectives of the undertaking.

Empirical Semantics maintained a much narrower focus than experimental philosophy. Naess had big ideas about its possible future, stating that ‘there are scarcely any of the traditional philosophic problems which are not suitable for this procedure’ *(“Truth”:* 162). In practice, however, he restricted his own work to topics in the philosophy of language, starting with the notion of ‘truth’, and continuing in later work to interrogate his participants on issues relating to synonymity (e.g. *Communication and Argument*). As well as being more limited in terms of the range of subjects to which it was applied, Naess’s approach was also more exploratory than that of most present day experimental philosophers. He set out to discover what people thought about a philosophically important term such as ‘truth’, or how they reacted to statements which might be judged to be synonymous, and drew generalisations from the data he collected. He did not start out with specific hypotheses about these issues and gather data with which to test such hypotheses. Further, and in still more marked contrast to the expressed concerns and motivations of many present day experimental philosophers, he did not have much to say about the possible psychological processes or social influences underlying his participants’ responses.

Naess’s hopes for the future of Empirical Semantics were perhaps more challenging to mainstream philosophy than are those of many present day experimental philosophers. Some experimental philosophers have argued that it is not significant whether or not their work can contribute to traditional philosophical problems because ‘certain aspects of the analytic project were actually misguided all along’ (Knobe ‘Experimental philosophy’, 89). Many, however, are keen to stress plurality in relation to philosophical method; experimental data collection and analysis can be a valuable tool for philosophers, but is not the only way of doing philosophy, and is proposed as an addition to rather than a replacement of more traditional methods. As Knobe and Nichols put it, ‘We are proposing another method (on top of all the ones that already exist) for pursuing certain philosophical inquiries.’ (‘An experimental philosophy manifesto’, 10). Naess saw Empirical Semantics as a necessary corrective to the errors of his fellow philosophers. The fact that the philosophical establishment was non-experimental meant that it was outdated and needed to be challenged:

What we aim at in this connection is to support the view that many “problems” which by tradition are localized outside fields demanding earnest research (of the sort demanded, for instance, of a scientist) and are made a privileged playground of those occupying the positions called “philosophic”, are reminiscences of types of thinking and feeling fundamentally foreign to the most valuable tendencies of our epoch.

(Naess, *“Truth”*, 58).

One of Naess’s supporters has complained that some contemporary criticism of the work on truth ‘was based on the misunderstanding that Naess wanted to *participate* in and contribute to the philosopher’s debate on truth, and set out to “find out what truth is” by asking non-philosophers. It was not understood that his aim was much more destructive in regard to traditional philosophy in general’ (Gullvåg, ‘Naess’s early philosophy’, 38, original emphasis). Naess, then, was not principally aiming at conceptual enlightenment, nor was he concerned to understand the psychological and other processes driving everyday intuitions about philosophical issues. His intention was to challenge the assumptions and what he saw as the complacencies of academic philosophy.

**4. Empirical Semantics and Experimental Philosophy: Criticisms**

Both Empirical Semantics in the 1930s and experimental philosophy in the present day have attracted criticism from those who see them as flawed or straightforwardly wrong-headed as approaches to philosophical inquiry. Despite the historical gap between the two movements, some of the criticisms leveled at experimental philosophy are strikingly similar to those that met the early results of Empirical Semantics, and suggest that some attitudes to philosophical practice have endured. Naess was not cowed by his opponents, and his intransigence can be seen as offering a certain degree of indirect support to experimental philosophy against its critics. Similarly, the current robust state of experimental philosophy can go some way towards vindicating what was at the time Naess’s controversial and largely autonomous approach. Other criticisms have been more closely associated with just Empirical Semantics or just experimental philosophy. These have something to say about changing attitudes to philosophical practice between the 1930s and the present day.

Naess’s insistence on pursuing an unorthodox and untried new approach to philosophy, and at that one which seemed directly to challenge the authority of philosophers themselves, did him no favours in academic circles. He was appointed to a chair at the University of Oslo soon after the publication of *“Truth”*, but this was in the face of strong and outspoken opposition, which focused specifically on that monograph. One member of faculty opposed his appointment to a chair on the grounds that: ‘The investigation is through and through characterized by a rather ironical attitude towards professional philosophy’ (E. Kaila in an evaluation of the applicants to the professorship in philosophy at the University of Oslo, September 3, 1938, quoted in Grimm, ‘On the application’, 118). Nagel concluded that, as a result of his way of doing philosophy, Naess: ‘will no doubt remain an outcast from the philosophical community’ (‘Review of *“Truth”*’, 79). Other philosophers focussed not just on Naess’s criticism of the unempirical generalisations of professional philosophers, but on his proposal instead to consult the intuitions of non-philosophers. Tarski was equivocal in his response to Naess’s work. He was quick to claim that it offered support for his own semantic conception of truth as according with common-sense usage, but was more cautious about the wider application of Naess’s empirical approach in relation to judging formal concepts of truth:

I should like to emphasise that in my opinion such investigations must be conducted with the utmost care. Thus, if we ask a highschool boy, or even an adult intelligent man having no special philosophical training, whether he regards a sentence to be true if it agrees with reality, or if it designates an existing state of affairs, it may simply turn out that he does not understand the question; in consequence his response, whatever it may be, will be of no value to us.

(‘The semantic conception of truth’, 360)

Within the monograph on ‘truth’ itself, Naess made it clear that he had no time for such fundamental objections to his work. He pointed out that if a professional philosopher and a school girl were to come up with the same definition, the logical conclusion of this position would be that the former but not the latter should be taken seriously. Summarising what seemed to him to be the necessary but ludicrous consequence of dismissing the judgments of non-philosophers, Naess wrote of the school girl’s account: ‘In no case may it be taken in earnest – if it is to be decided what criterion of truth is to be used by the non-philosopher, a philosopher must be required: a man with accepted capability to solve, or at least to discuss, problems of the kind at issue’ (*“Truth”,* 17).

There are clear parallels between the negative reactions to Naess’s use of the intuitions of non-philosophers and one of the major objections to experimental philosophy, the so-called ‘expertise objection’. According to this objection, philosophical training by its nature sharpens the ability to make fine judgments and utilize subtle distinctions, meaning that philosophers will necessarily arrive at clearer and more useful intuitions about concepts than non-philosophers. A leading critic of experimental philosophy on these grounds is Williamson, who has argued for instance, that:

Philosophy students have to learn how to apply general concepts to specific examples with careful attention to the relevant subtleties, just as law students have to learn how to analyze hypothetical cases. Levels of disagreement over thought experiments seem to be significantly lower among fully trained philosophers than among novices. […] We should not regard philosophical training as an illegitimate contamination of the data, any more than training natural scientists how to perform experiments properly is a contamination of their data. Although the philosophically innocent may be free of various forms of theoretical bias, just as the scientifically innocent are, that is not enough to confer special authority on innocent judgement, given its characteristic sloppiness.

(*The Philosophy of Philosophy*, 191)

Some experimental philosophers are more conciliatory to mainstream professional philosophers than Naess was, suggesting that professional judgments should be complemented by ‘folk’ intuitions (see, e.g. Knobe and Nichols, ‘An experimental philosophy manifesto’, 9), while maintaining that claims about the differences between philosophical and folk intuitions are themselves made unempirically, and should be subjected to rigorous testing (see work reported in Sytsma and Livengood, ‘Experimental philosophy and philosophical disputes’, 153). Others dismiss the expertise objection on the grounds that there is simply no need to make a distinction between the intuitions of philosophers and those of non-philosophers; ‘there are significant similarities in philosophers’ and nonphilosophers’ judgements’ (Rose and Danks, ‘In defense of a broad conception’, 522). But for many experimental philosophers, as for Naess, the intuitions of professional philosophers are best avoided altogether. Alexander and Weinberg, for instance, have argued that ‘philosophical practice is not concerned with understanding the nature of *knowledge* in some technical sense, but of *knowledge* as the concept is ordinarily understood’, and therefore that philosophers’ intuitions cannot be relied on to be representative (‘Analytic epistemology and experimental philosophy’, 58, original emphases). Other studies have suggested that ‘expertise about language does not appear to increase the reliability of people’s semantic intuitions’ and that on the contrary factors such as philosophical reading and background knowledge seem to interfere with the reliability of such intuitions (Sytsma and Livengood, ‘Experimental philosophy and philosophical disputes’, 153).

A further objection to present day experimental philosophy, related to the expertise objection but distinct from it, is that in consulting ordinary intuitions about the use of everyday expressions, experimental philosophers are in danger of mistaking context-dependent responses to pragmatic aspects of meaning for intuitions about the semantics of philosophically salient expressions. Sosa touches on this objection when he argues that experimental philosophers have not done enough to take account of the fact that ‘verbal disagreement need not reveal any substantive, real disagreement, if ambiguity and context might account for the verbal divergence’ (‘Experimental philosophy and philosophical intuition’, 234). Kauppinen sees the intrusion of pragmatic factors into intuitions about meaning as an inherent problem for experimental philosophers. It is not possible to assume a direct link between use and meaning, because ‘the appropriateness of what we say also depends on various pragmatic factors that are not part of the meaning or semantic content of the expression’ (‘The rise and fall of experimental philosophy’, 104). Bengson draws attention to work which suggests that participants’ responses may vary relative to different interpretations in context, suggesting responses to pragmatic rather than semantic aspects of the questions put to them (‘Experimental attacks on intuitions and answers’, 505).

Criticisms of Naess’s work were in general less focused than those of the work of experimental philosophers; his critics simply did not like what he was doing, seeing it as not a profitable approach to philosophy. Moore dismissed Naess’s work on the grounds that he ‘seems to have little to show for it when all is done and recorded’ (‘Review of *“Truth”*’, 490). Crockett protested that ‘it is not at all clear that the description of a stock use of an expression is assisted by counting the noses of those who employ it in this way’ (‘An attack upon revelation’, 109). There is perhaps a hint here of the later accusation that experimental philosophers are open to the confusion of the semantic and the pragmatic in their emphasis on participants’ intuitions about language use. Quine was even more dismissive on these grounds. He described Naess’s approach as ‘unimaginative’, and lampooned it as being ‘ask the natives’ (‘Methodological reflections’, 392). Urmson, acting in this case as a spokesman for Ordinary Language Philosophy, with its emphasis on philosophical group work, brought together a general criticism of Naess’s methods with a nod towards the expertise objection:

The device of a statistical survey of ‘what people would say’ by means of a questionnaire is no substitute for the group, (1) because there cannot be the necessary detail in the questionnaire, (2) because the untrained answerers can so easily make mistakes, (3) because we are raising questions where unanimity is both desirable and obtainable.

(Urmson et al, ‘A symposium’, 80)

Empirical Semantics and experimental philosophy, then, have both faced the criticism that the intuitions of non-philosophers are inappropriate data in philosophical inquiry, suggesting a continuing belief in some areas of the philosophical establishment that philosophical training is necessary to make reliable philosophical judgments. Naess faced other, less focused, objections that are not generally encountered by experimental philosophers, to the effect that statistical methodology was simply misplaced in philosophy. This difference may reflect a much more widespread acceptance now than in the 1930s of the place of quantitative research in academic study. But experimental philosophy faces one type of criticism that Naess was not confronted with, namely specific and detailed objections to the exact methodology being used. This in turn reflects more detailed present day thinking about the issues involved in the use of quantitative data.

It is no exaggeration to say that Naess was doing something unprecedented in introducing the methods of the sociology to philosophical study. Empirical methods had of course been explored by many earlier philosophers, but the use of questionnaires to interrogate large numbers of non philosophers and of statistical techniques to analyse the collected data, were unheard of in philosophical circles in the 1930s. These things form the mainstream of present day experimental philosophy. Experimental philosophers are explicit in their ambition to introduce into philosophy the empirical rigour of scientific study, particularly the methods of experimental psychology.

Murphy notes that both Empirical Semantics and experimental philosophy draw on the available techniques of their day; ‘At the center of empirical semantics and experimental philosophy is the use of the latest tools from psychology and social science’ (‘Experimental philosophy’, 42). Some techniques now available to experimental philosophers depend on recent advances in technology, and would have been unimaginable in Naess’s day. Rose and Danks have pointed out that some current work in experimental philosophy draws not just on participants’ responses, but also on measurements of their reaction times. They urge experimental philosophers to ‘avail themselves of the best experimental methods of cognitive science, including not just survey responses but also behavioral measures, neuroimaging data, and other measures of cognitive functioning’ (‘In defense of a broad conception’, 525).

Despite the availability of such advanced techniques, experimental philosophy has recently been challenged on methodological grounds. Woolfolk, for instance, states baldly that ‘the experiments conducted by experimental philosophers frequently fail to meet the methodological standards that are articulated by the experts on research design in those fields they would emulate’ (‘Experimental philosophy’, 80). Woolfolk voices one particular set of views on how experiments in philosophy could be made better and more reliable. Nevertheless, it is instructive to compare some of his specific criticisms of present-day methodology with some of the things Naess said about his research practice, in relation even to his early work on ‘truth’. It is not that Naess’s work straightforwardly succeeds in those areas where Woolfolk claims experimental philosophy fails. But it is certainly interesting to see that Naess was thinking about these methodological issues back in the 1930s, and either attempting to address them, or at least acknowledging shortcomings in relation to them, in his own work.

Firstly, Woolfolk complains that experimental philosophers frequently fail to pay adequate attention to the design and wording of their questionnaires:

Do the participants comprehend the questions in the manner they are expected to by the experimenter? It is rare that experimental philosophers conduct adequate pilot testing to answer this question of participant comprehension, a question psychologists and pollsters, through painful and embarrassing experience, have learned to ask before a final data collection is initiated.

(‘Experimental philosophy’, 81)

Although he does not use the term, Naess reports on something like an initial pilot study in his work on ‘truth’. Using his own idiosyncratic system of abbreviations, he explains how the questionnaires used in the survey were designed and refined through

experiment, and how the initial data involved in this process did not form part of the eventual findings:

Of the first 75 ps [‘test-persons’], a great many were examined by the oral method, but the questions did follow preconceived qts [‘questions’]. The aim of these examinations was mainly to find out the most effective types of questionnaires. The thoughts of the ps were to a certain degree followed without the l’s [leader of the tests] intervention. We generally reject the material collected by this method for statistical purposes.

(Naess, *“Truth”* 22).

Secondly, Woolfolk draws attention to the fact that experimental philosophers generally ignore qualitative methods, which in experimental psychology often complement quantitative techniques. The statistical analysis of questionnaire results can be enlightening, but may not be the best way of reaching more in-depth understanding:

Qualitative research would include focus groups, ethnographic methods and intensive case studies … if we are interested in a deeper understanding of what a given individual ‘really meant’ by her responses, qualitative researchers argue, some more subtle and individualized questioning of participants may yield a richer and more nuanced account, and one which is closer to the truth of the matter.

(‘Experimental philosophy’, 82)

Naess reports on trial runs of both the ‘individualized questioning’ and the ‘focus groups’ techniques during his study of ‘truth’. Initially, test persons were asked for their general responses to the nature of ‘that which is true’, and those responses were collected ‘by letting the ps talk without trying to lead the conversation by further questions. They were allowed to associate freely. The subjects thus touched upon by the ps were adopted as subjects for further questions’ (*“Truth”*, 19). In addition, ‘a certain group-technique was employed. 2, 3 or 5 persons were asked to consider the formulations. They were allowed successively to make their remarks and discuss the divergences’ (*“Truth”*, 31).

Further, Woolfolk complains that experimental philosophers are too ready unthinkingly to conduct their own empirical work: ‘I have yet to read an experimental philosophy study explicitly stating that administrators or scorers of questionnaires were blind to the participants’ assignment to conditions or to the experimental hypotheses’, which, he argues, is now a basic requirement of biomedical research (‘Experimental philosophy’, 83). Naess did not in general work with experimental hypotheses, his studies being much more speculative in nature, but he was aware of the potential for bias or suggestion if what he called the ‘leader of the tests’ were to be someone with a background in the issues at stake: ‘To avoid suggestion due to preconceived opinions on the truth-discussion, we deferred the task of systematic interviews to a person without any knowledge about the questions involved’ (*“Truth”,* 20).

Woolfolk also expresses concern over choices made in experimental philosophy about the actual test subjects used. Any researcher intending to work with data collected from human participants must address questions about how many participants to use, and from what demographic(s) they should be selected. Woolfolk argues that most studies use ‘opportunity samples’ chosen because of availability, and often consisting of undergraduate students, and that as a result ‘data from the aforementioned opportunity samples raise questions about the generalizability of the results they generate’ (‘Experimental philosophy’, 82). Naess was clearly deeply interested in issues about the choice of participants, and commented extensively on the make-up of his group of 250 ‘test-persons’. At the start of his report into his study he explains that ‘As “non-philosophers” (“amateurs”), persons of different age, education and sex, functioned’ (*“Truth”*, 22). Some were undergraduate students at the University of Oslo, although he tried to minimize use of students as participants, and he was adamant that none were to be philosophically educated to professional level. During the course of the monograph he suggested, although he did not always carry out, various possible comparative studies in relation to gender, age and nationality. For Naess, a diverse cohort of test subjects was essential, and crucially a cohort different from the trained philosophers who had previously had the monopoly on commenting on the applicability of philosophical terms. As he explained some time later in his career, ‘we all too commonly and rashly presume that what applies in our own case, or in that of some limited group, is also true generally’ (*Communication and Argument*, 21).

The majority of Naess’s test subjects were Norwegian, with just a few Austrian, Swedish, German and English participants. Naess seemed somewhat rueful about this, commenting that ‘This material was of course found too scanty to allow of conclusions of influence of nationality’ (*“Truth”*, 45). Present day experimental philosophers are increasingly interested in the cultural-specific nature of intuitions, and therefore in reaching precisely the types of conclusions that Naess hinted at. Machery et al, for instance, suggest that intuitions relevant to philosophical theories of reference may vary between Western and East Asian subjects, and offer comparative experimental data in support of this hypothesis. Their conclusions include a statement remarkably close to what Naess was saying about professional philosophers in the 1930s: ‘Since the intuitions philosophers pronounce from their armchairs are likely to be a product of their own culture and their academic training, in order to determine the implicit theories that underlie the use of names across cultures, philosophers need to get out of their armchairs’ (‘Semantics, cross-cultural style’ 54).

One striking aspect of Naess’s commentary on his own methodology, which he repeats a number of times during his monograph on ‘truth’, is his dissatisfaction with the number of test subjects he has been able to include, and his ambition to include substantially higher numbers. The task he had set himself of interrogating conceptions of truth among non philosophers ‘was overwhelmingly large, if the result were to be at all satisfactory from a scientific point of view’ (*“Truth”,* 11). It was, he noted, ‘impossible to ask *all* people about “truth”. ‘*None of our conclusions we therefore think is in anyway secure*’ (*“Truth”*, 20, original emphases). He would prefer, he admits, to increase the number of subjects from 250 to 1000 or 5000.

In comparison with Naess’s volubility on the subject, experimental philosophers have sometimes been rather terse when giving details of the selection and make-up of their group of test subjects. Here are a few brief, representative accounts from studies conducted in the early days of experimental philosophy:

Subjects were 78 people spending time in a Manhattan public park.’

(Knobe, ‘Intentional action’, 191 and 192)

All respondents were undergraduates at Florida State University who had not taken any philosophy courses.

(Cushman and Mele, ‘Intentional action’, 180)

Participants were 126 undergraduates.

(Nadelhoffer, ‘Bad acts, blameworthy agents’, 156)

Seventy-two (34 female and 38 male) University of California, Santa Cruz, undergraduates enrolled in philosophy classes participated in the experiment as volunteers.

(Woolfolk et al, ‘Identification, constraint and cognition’, 65)

Forty-eight (28 female and 20 male) Princeton University students participated in the experiment in order to fulfill requirements for participation in Psychology Department research.

(Woolfolk et al, ‘Identification, constraint and cognition’, 74)

Forty undergraduates at Rutgers University and 42 undergraduates from the University of Hong Kong participated.

(Machery et al, ‘Semantics, cross-cultural style’, 51)

Two factors are immediately striking here. Firstly, the number of participants in these studies is generally lower than the figure of 250 that Naess used and with which he was uneasy because of the restrictions it placed on the validity of his data. Secondly, many studies in experimental philosophy from the first decade of the twentieth century used undergraduate students, including those taking courses in philosophy or psychology as their subjects.

However, new resources present new opportunities. Amazon’s Mechanical Turk (AMT), established in 2005, has provided an alternative source of research participants, and is proving increasingly popular with experimental philosophers. It offers researchers access to a sample of a huge online workforce ready to complete short Human Intelligence Tasks (HITs) for small financial rewards. Studies suggest that AMT is a viable alternative to more traditional experimental methods in linguistics, in psychology and by extension in philosophy (Sprouse, ‘A validation of Amazon Mechanical Turk’; Paolacci and Chandler, ‘Inside the Turk’). AMT offers a means of addressing the problems associated with reliance on student participants because its workforce is ‘more demographically diverse than standard Internet samples and significantly more diverse than typical American college samples’ (Buhrmester et al ‘Amazon’s Mechanical Turk’, 4). In effect, recent technological developments now make available to experimental philosophers the massive, heterogeneous set of subjects that was Naess’s stated ideal. Paolacci et al note that researchers can restrict participation in a given HIT in relation to country, which ‘makes it easy to conduct cross-cultural studies’ (‘Running experiments on Amazon Mechanical Turk’, 413), which brings Naess’s aspiration of studying ‘the influence of nationality’ within reach. Recent studies include those that have used AMT participants to gather data on pragmatic inferencing (Cummins et al, ‘Granularity and scalar implicature’), to add comparative data to a study revisiting judgments about intentional actions (Mele, ‘Intentional, unintentional, or neither?’), to test for self-serving biases in trust games (Bicchieri and Mercier, ‘Self-serving biases’) and to investigate folk conceptions of weakness of will (Beebe, ‘Weakness of will’).

A moral that Naess drew from his own work on ‘truth’ was that empirical work in philosophy was of potentially great value, but would never be easy, and that the value would in fact be determined by the amount of effort put into collecting the data, the extent of that data, and the care and thought with which they were collected and anaylsed.

One of our main conclusions may be formulated thus: If one is seriously interested in the discovery of amateur-theories of truth, the amount of knowledge secured will be roughly proportional to the labour spent in collecting and analyzing material obtained by questioning the ps and by observing their behavior. Meditations and deductions from general principles will not do – just as little in this field of research as in botany.

(*“Truth”*, 153).

It is a moral that serves well also for present-day experimental philosophers, who remain keen to evaluate their own methodologies, scrutinize their own practices, and work to improve the reliability and quality of their data.

Reflecting in later life on the nature and the extent of the criticisms he had faced as a result of his work in Empirical Semantics, Naess was unapologetic. His questionnaire methodology, and his interest in the intuitions of non experts, were what stood out for him as the main problems:

I used questionnaires. At that time, 1937-8, they were looked upon as the absolute bottom of doing research. They couldn't be taken seriously at all. And then it implied that I had an undignified, really atrocious view of one of the great problems of humanity – namely, the problem of truth. Taking seriously what those schoolboys and housewives were saying was a kind of caricature of philosophy.

(Rothenberg, *Is it Painful to Think?,* 49).

By the time he made this statement, in the 1990s, Naess had moved on to other interests. But it is clear that he did not recant his commitment either to the questionnaire methodology or to paying serious regard to the intuitions of non-philosophers. As a result of these commitments he had remained something of a philosophical outcast, as Nagel predicted, at least in relation to mainstream analytic philosophy outside of Norway. The criticism that Naess encountered that non-philosophers are just not suitable subjects for philosophical study are still being encountered today by experimental philosophers, and indicate an on-going controversy about the relationship between philosophical training and the ability to make rigorous conceptual judgments. The view of questionnaires as ‘the absolute bottom of doing research’, however, is not a problem that is in general encountered by present day experimental philosophers; questionnaires are established practice in social psychology and a wide range of other academic disciplines. The detailed criticisms by scholars such as Woolfolk of the precise methods of experimental philosophy in fact highlights a widespread acceptance of questionnaires as a tool in quantitative methodology, and an aspiration to improve their use and interpretation. This aspiration is widely shared by experimental philosophers themselves, and is exemplified in their adoption of newly available types and samples of research participants.

**6. Conclusions**

A comparison between Empirical Semantics of the 1930s and present day experimental philosophy is instructive, and offers insights that would not be available from an assessment of either approach in isolation. The ambitions and methodologies of experimental philosophy lend some retrospective support to what was at the time Naess’s unusual and unpopular programme. Perhaps of more pressing current relevance, the fact that Naess said the things he said about philosophy, and did the things he did in his research, offers an interesting and underexplored antecedent for experimental philosophy. Differences between the two approaches are significant too. Naess’s interest in his participants’ intuitions as primary philosophical data, and particularly in their potential to challenge accepted philosophical thinking, is in contrast to and highlights the emphasis in experimental philosophy on the cultural and psychological underpinnings of folk intuitions. Criticisms which have been shared by the two disciplines, such as the argument from expertise, highlight attitudes to philosophical training and philosophical practice which are enduring. Other criticisms reveal changing attitudes. Empirical Semantics was challenged on the grounds that questionnaires were simply not an appropriate methodology for serious academic research. Experimental philosophy, in contrast, has been subjected to detailed methodological scrutiny. This indicates changes over the decades in the expectations on academic inquiry in general, and philosophical method in particular.

**Bibliography**

Alexander, Joshua and Jonathan Weinberg. ‘Analytic epistemology and experimental philosophy’, *Philosophy Compass* 2 (2007) 56-80.

Appiah, Kwame Anthony. ‘Experimental philosophy’, *Proceedings and Addresses of the American Philosophical Association* 82 (2008): 7-22.

Austin, J. L. ‘A plea for excuses’, *Proceedings of the Aristotelian Society*, 57 (1956): 1-30.

Barnard, Robert and Joseph Ulatowski. ‘Tarski’s 1944 polemical remarks and Naess’s “Experimental Philosophy”’, *Erkenntnis* 81 (2016). 457-77

Beebe, James. ‘Weakness of will, reasonability, and compulsion’, *Syntheses* 190 (2013): 4077-4093.

Bengson, John. ‘Experimental attacks on intuitions and answers’, *Philosophy and Phenomenology Research* 88 (2013): 495-532.

Bicchieri, Christina and Hugo Mercier. ‘Self-serving biases and public justifications in trust games’, *Synthese* 190 (2013): 909-922.

Buhrmester, Michael, Tracy Kwang and Samuel D. Gosling. ‘Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality data?’, *Perspectives on Psychological Science* 6 (2011): 3-5.

Chapman, Siobhan. *Language and Empiricism, After the Vienna Circle*, Basingstoke: Palgrave MacMillan, 2008.

Crockett, Campbell. 'An attack upon revelation in semantics', *The Journal of Philosophy* 56 (1959): 103-11.

Cummins, Chris, Uli Sauerland and Stephanie Solt. ‘Ganularity and scalar implicatures in numerical expressions’, *Linguistics and Philosophy* 35 (2012): 135-169.

Cushman, Fiery and Alfred Mele. ‘Intentional action: two-and-a-half folk concepts?’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 171-188. Oxford: Oxford University Press, 2007.

Grimm, Laura. 'On the application of the concept of precisation', *Synthese* 9 (1955): 104-20.

Gullvåg, Ingemund.'Naess's early philosophy of science and philosophy'. In *In Sceptical Wonder*, edited by Ingemund Gullvåg and Jon Wetlesen, 22-55. Oslo: Universitetsforlaget. 1982.

Kauppinen, Antti. ‘The rise and fall of experimental philosophy’, *Philosophical Explorations* 10 (2007): 95-118.

Knobe, Joshua. ‘Intentional action and side effects in ordinary language’, *Analysis* 63 (2003): 190-93.

Knobe, Joshua. ‘Experimental philosophy’, *Philosophy Compass* 2 (2007): 81-92.

Knobe, Joshua. ‘Experimental philosophy is cognitive science’. In A Companion to Experimental Philosophy, edited by Justin Sytsma and Wesley Buckwalter, Oxford: Blackwell, 2016.

[Knobe](http://philpapers.org/s/Joshua%20Knobe), Joshua and [Shaun Nichols](http://philpapers.org/s/Shaun%20Nichols). ‘An experimental philosophy manifesto’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 3-14. Oxford: Oxford University Press, 2007.

Machery, Edouard, Ron Mallon, Shaun Nichols and Stephen P. Stich. ‘Semantics, cross-cultural style’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 47-58. Oxford: Oxford University Press, 2007.

Mele, Alfred. ‘Intentional, unintentional, or neither? Middle ground in theory and practice’, *American Philosophical Quarterly* 49: 369-379.

Moore, Jared. ‘Review of *“Truth” as Conceived by Those Who Are Not Professional Philosophers* by Arne Ness’, *The American Journal of Psychology* 52 (1939): 489-90.

Murphy, Taylor. ‘Experimental Philosophy 1935-1965’. In *Oxford Studies in Experimental Philosophy: Volume 1*, edited by Joshua Knobe, Tania Lombrozo and Shaun Nichols, 1-58. Oxford: Oxford University Press Scholarship Online, 2014.

Nadelhoffer, Thomas. ‘Bad acts, blameworthy agents, and intentional actions: some problems for juror impartiality’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 149-167. Oxford: Oxford University Press, 2007.

Naess, Arne.*“Truth” as Conceived by those who are not professional philosophers*. Oslo: Jacob Dybwad, 1938.

Naess, Arne. *An Empirical Study of the Expressions “True”, “Perfectly Certain” and “Extremely Probable”.* Oslo: Jacob Dybwad, 1953.

Naess, Arne. *Communication and Argument*. Oslo: Universitetsforlaget, 1966.

Nagel, Ernest. ‘Review of *“Truth” as Conceived by Those Who Are Not Professional Philosophers* by Arne Ness’, *The Journal of Philosophy* 36 (1939): 78-80.

Paolacci, Gabriele and Jesse Chandler. ‘Inside the Turk: Understanding Mechanical Turk as a participant pool’, *Current Directions in Psychological Science* 23 (2014): 184-188.

Paolacci, Gabriele, Jesse Chandler and Panagiotis G. Ipeirotis. ‘Running experiments on Amazon Mechanical Turk’, *Judgment and Decision Making* 5 (2010): 411-419.

Quine, W. V. O. 'Methodological reflections on current linguistic theory', *Synthese* 21 (1970): 386-98.

Rose, David and David Danks. ‘In defense of a broad conception of experimental philosophy’, *Metaphilosophy* 44 (2013): 512-532.

Rothenberg, David. *Is it Painful to Think? Conversations with Arne Naess*. Minneapolis: University of Minnesota Press, 1993.

Sosa, Ernest. ‘Experimental philosophy and philosophical intuition’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 231-240. Oxford: Oxford University Press, 2007.

Sprouse, Jon. ‘A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory’, *Behavioral Research* 43 (2011): 155-167.

Sytsma, Justin and Jonathan Livengood. ‘Experimental philosophy and philosophical disputes’, *Essays in Philosophy* 13 (2012): 145-160.

Tarski, Alfred. ‘The semantic conception of truth and the foundations of semantics’, *Philosophy and Phenomenological Research* 4 (1944): 341-74.

Urmson, J. O., W. V. O. Quine and Stuart Hampshire. ‘A symposium on Austin’s method’. In *Symposium on J. L. Austin*, edited by K. T. Fann, 76-97. London: Routledge, 1969.

Williamson, Timothy. *The Philosophy of Philosophy*, Oxford: Blackwell, 2007.

Woolfolk, Robert L. ‘Experimental philosophy: a methodological critique’, *Metaphilosophy* 44 (2013): 79-87.

Woolfolk, Robert L., John M. Doris and John M. Darley. ‘Identification, situational constraint, and social cognition: studies in the attribution of moral responsibility’. In [*Experimental Philosophy*](http://philpapers.org/rec/KNOEP), edited by Joshua Knobe and Shaun Nichols, 61-80. Oxford: Oxford University Press, 2007.

1. I am grateful to Michael Beaney, to Alberto Vanzo, and to two anonymous reviewers for their very helpful comments on earlier drafts of this article. [↑](#footnote-ref-1)