ORIGINAL ARTICLE

The experimental turn in philosophy of language

Xiang Zhou*, Ya Gao

College of Foreign Languages, Shanghai Maritime University, Shanghai 201306, China

Abstract: With the rise of experimental philosophy in the twenty-first century, the past two decades have witnessed the experimental turn in the field of philosophy of language. We delineate in this paper the experimental turn in philosophy of language before distinguishing armchair theorizing from empirical testing and highlighting the complementarity between the two approaches, and then carry out an analysis of the experimental tools and methods available for philosophical experiments with examples by classifying them into three major types, viz., the method of survey, the method of big data, and the method of cognitive neuroscience.

Keywords: experimental philosophy; philosophy of language; experimental method

***Corresponding author:** Xiang Zhou, College of Foreign Languages, Shanghai Maritime University, Shanghai 201306, China; xzhou@shmtu.edu.cn

Received: June 14, 2020; **Accepted:** July 13, 2020; **Available online:** August 12, 2020

Citation: Zhou X and Gao Y (2020) The experimental turn in philosophy of language. *Forum for Linguistic Studies* 2(1): 19–26. DOI: 10.18063/fls.v2i1.1201

1. Introduction

The fundamental research method of philosophers of language is linguistic or conceptual analysis, that is, to carry out an in-depth analysis and elucidation of concepts through linguistic and logical analysis. From the perspective of philosophers of language, quite a number of philosophical puzzles can be ascribed to the obscurity of concepts, and hence they advocate conceptual analysis by means of thought experiments and by bringing up counter-examples. Dennett (1980) coined the term "intuition pump" to vividly describe what a thought experiment is and how it works—to pump (or elicit) one's intuition regarding a philosophical proposition or problem with hypothetical cases. However, there is no doubt that the reliability of intuition itself is dubious. As philosophers make a claim by using "people would say (in this situation)", "intuitively", or "ordinarily", they argue by appealing to their (own) intuitions. But how can we make sure that a theory or a thesis is plausible and well-founded if only a dozen of philosophers claim such and such out of intuition? With its initial aim to empirically test the reliability and generality of intuition, experimental philosophy (often abbreviated as X-Phi) has developed since the turn of the twenty-first century and is gaining momentum. Experimental philosophers adopt the experimental methods typically used in social psychology and cognitive science to investigate questions of philosophical interest and make an attempt to justify or falsify the theoretical claims and hypotheses made by traditional philosophers

with empirical evidence.

Since its emergence at the beginning of this century with the groundbreaking work of Weinberg et al. (2001), Nichols et al. (2003) and Machery et al. (2004), the approach of experimental philosophy has been adopted in many fields of philosophy, including not only philosophy of language, but also epistemology, ethics, philosophy of mind, philosophy of economics, and others. It also needs to be noted that this new philosophical movement has greatly contributed to the rethinking and reexamination of metaphilosophy and philosophical methodology. Many traditional philosophers might not readily accept the experimental approach, but the influential work and achievements of experimental philosophers should never be ignored. With the rapid development of experimental philosophy, experimental semantics and pragmatics, experimental philosophy of language is also gaining in popularity in the past two decades, with its main topics including (empirical tests of) the reference of proper names and natural kind terms, Kripke's Gödel case, the context-sensitivity of "knows" (knowledge), transparent ascriptions, etc. The unprecedented important role that experimentation plays in the study of philosophy of language and the emergence of an abundance of experimental research in philosophy of language both indicate that philosophy of language is taking an experimental turn. This, however, never means that the experimental approach will take place of conceptual and linguistic analysis, but simply that experimental philosophy has provided philosophers with an indispensable and useful toolkit, thus ensuring a higher validity and reliability of philosophical studies if they make good use of the experimental methods. In this paper we will first explain the distinction between armchair theorizing and empirical testing and their complementarity before illustrating three kinds of experimental tools available to philosophers, viz., the method of survey, the method of big data, and the method of cognitive neuroscience.

2. Armchair theorizing vs. empirical testing

Ever since the age of Plato, the appeal to intuition has been an essential, if not the only, source of evidence in doing philosophy. Philosophers generally see consistency/inconsistency with intuition as the default standard in making judgments about the reliability and acceptability of a theory or an argument. For philosophers of language, the major work they are engaged in over the years is conceptual analysis, and a common way to do conceptual analysis is to devise real or imagined cases (thought experiments) and to make intuitive judgments about the truth value or acceptability of some proposition of philosophical interest in the cases, on the basis of which arguments are developed, justified, or falsified. People may wonder where these intuitive judgments come from? Generally speaking, they are from themselves. As the traditional or typical way analytic philosophers theorize is through linguistic and logical analysis (which we can carry out all in an armchair with no need to go out to conduct experiments), traditional analytic philosophy was given a "nickname"—"Armchair Philosophy" ("AP" hereafter), contrasted with experimental philosophy. Experimental philosophers cast doubt on the reliability of intuition on which armchair philosophers rely in theorizing by the "Argument from Variation", which goes like this:

1) Experimental studies have shown that intuitions vary as a function of inappropriate factors like cultural background, emotional state, and so forth.

2) The above variation shows intuition to be an unreliable guide to philosophical truth.

3) Philosophers must abandon their reliance on intuition-based methods of theorizing. (Nado, 2016: 1–2)

It should be noted, however, that the Argument from Variation stands only if the traditional philosophers' method of philosophizing is really intuition-based. However, some analytic philosophers like Cappelen and Ichikawa have argued against the centrality of intuition by claiming that it is not intuition, but facts about philosophical phenomena which constitute the basis of evidence in philosophy (see Cappelen, 2012; Deutsch, 2009, 2010; Ichikawa and Jarvis, 2013; Ichikawa, 2014). But the real fact could be that facts and arguments are both intuition-related. Hence both intuitions and facts are supposed to come under empirical scrutiny.

As a matter of fact, conceptual analysis and empirical testing complement each other and should be combined in use in philosophical studies. Conceptual analysis and argumentation are never the privilege of philosophers. Likewise, experimentation is not the privilege of scientists, either. As Kornblith claims, "When philosophical theories in epistemology are constructed from the armchair, they run a serious risk of being divorced from the very phenomena they seek to illuminate. The only way to assure that we do not build elaborate castles in the air, unconnected to the real world phenomena which motivate our work in the first place, is to base our work on the best available experimental understanding of those phenomena." (Kornblith, 2014: 207) Arguably the rise of experimental philosophy symbolizes the return of the tradition of the continuity of philosophy and science. We agree with Papineau that philosophy is like science in several ways. First, the claims made by philosophy are synthetic, not analytic. Second, philosophical knowledge is a posteriori, not a priori: the claims established by philosophers depend on the same kind of empirical support as scientific theories. And finally, the central questions of philosophy concern actuality rather than necessity: philosophy is primarily aimed at understanding the actual world studied by science, not some further realm of metaphysical modality. (Papineau, 2014) The continuity of philosophy and science is clearly reflected by the fact that a variety of sub-disciplines with the form of "philosophy of X" have well developed since the twentieth century, such as the philosophy of science, the philosophy of psychology, the philosophy of physics, the philosophy of cognitive science, and the philosophy of information. If philosophy is consistent with science and can benefit from experimentation in a real sense, we would not have any reason to reject the available experimental methods of cognitive science, psychology, and neuroscience.

3. The methodology for experimental philosophy

Experimental philosophers do not have their own methodology, but they have borrowed in their research a wide variety of experimental tools and methods from social psychology and cognitive science. The most commonly adopted method in experimental philosophy has been questionnaire survey, which is used to elicit folk intuitive judgments about questions or propositions in devised cases. Nevertheless, it has been realized that survey is only one of the experimental tools available to philosophers, and there are still many different types of experimental methods available for philosophical investigations, such as corpus analysis, eye-tracking, VR technology, and fMRI. Ever since the beginning of this century, a number of philosophy laboratories have been established at world-renowned universities, such as the Experimental Philosophy Laboratory at the University of California at San Diego, the Experimental Philosophy Laboratory at Yale University and

the University of Arizona, the Experimental Epistemology Laboratory at Indiana University, Experimental Epistemology Research Group at University at Buffalo, the State University of New York, Pittsburgh Empirical Philosophy Lab at the University of Pittsburgh, to name but a few.

We classify the experimental methods applicable to philosophical experiments into three major types, viz., the method of survey, the method of big data, and the method of cognitive neuroscience. The method of survey is typically used to conduct behavioral and psychological experiments. By using the method of big data, we can collect data of real linguistic and conceptual use as empirical evidence for theoretical arguments. By adopting the method of cognitive neuroscience, data of brain activities can be precisely tracked with the aid of sophisticated equipment. Despite the wide range of experimental tools available, experimental philosophers have not broaden their use of the toolkit in their philosophical studies until the past decade or so. We will elaborate on these three types of experimental methods with specific cases of experimental studies of philosophy in the following sections.

3.1. The method of survey

The method of survey in social psychology has been the most commonly used method in experimental philosophy studies. With the aim to test the reliability of philosophers' intuition-based claims, researchers devise imagined cases and let the subjects make intuitive judgments on questions concerning some philosophical concept, usually by using a Likert scale. The survey method has the following characteristics: Firstly, it has a close relation to thought experiments in that many experimental cases adopted in survey studies are classic thought experiments, like the Gettier cases in Starmans and Friedman (2012) and the fake-barn cases in Colaço et al. (2014). Secondly, compared with other kinds of experiments that rely on sophisticated experimental instruments, it is simpler and more convenient to carry out a survey. The main procedures for conducting a survey are to propose a research question, make an assumption, devise cases, design the questionnaire, implement it, collect data, and analyze the results. The most critical and difficult part of an experimental study must be experimental design. Let's take a look at the pioneering work in experimental philosophy by Weinberg et al. (2001).

At the turn of the twenty-first century, Jonathan Weinberg and his colleagues conducted a groundbreaking experimental investigation into the cultural variation in epistemic intuitions. Considering that Nisbett (2001) demonstrates that there are differences in belief forming strategies in different cultural groups, Weinberg then tries to explore the possible differences in epistemic intuitions in different cultural groups. He put forward four hypotheses: 1) Epistemic intuitions vary from culture to culture; 2) Epistemic intuitions vary from one socioeconomic group to another; 3) Epistemic intuitions vary as a function of how many philosophy courses a person has had; 4) Epistemic intuitions depend, in part, on the order in which cases are presented. Three versions of Truetemp cases (the Individualistic Version, the Elders Version, and the Faluki Version) were designed to explore externalist/internalist dimensions of the subjects' intuitions in the surveys. Take a look at the individualistic version.

Individualistic Version: One day Charles is suddenly knocked out by a falling rock, and his brain becomes re-wired so that he is always absolutely right whenever he estimates the temperature where he is. Charles is completely unaware that his brain has been altered in this way. A few weeks later, this brain re-writing leads him to believe that it is 71 degrees in his room. Apart from his estimation,

he has no other reasons to think that it is 71 degrees. In fact, it is at that time 71 degrees in his room. Does Charles really know that it was 71 degrees in the room, or does he only believe it? (Weinberg, Nichols and Stich, 2001)

The subjects in this experiment (undergraduates at Rutgers University, East Asians and Westerners) are required to choose from whether "really knows" or "only believes" (that it is 71 degrees). The results suggest that there exists a highly significant difference between East Asians and Westerners in the individualistic version, and there are no significant differences in the Elders version and the Faluki version. The research findings indicate that in the individualistic version Charles' reliable mechanism is epistemically external. Therefore, the group who choose "only believes"'s folk epistemology may be internalist; in the elders version, the reliability of epistemological mechanism is social sanction; in the Faluki version, it is shared by everyone in the community. The results of this experiment confirmed their hypothesis that intuition is culture-specific.

Before and while conducting a survey study of philosophy, we should pay particular attention to the following few aspects. First, strict and effective control should be taken of all the possible disturbing variables in the process of experimental design. Second, presentation effects such as the ordering effect (the influence of the order in which the vignettes are presented) and the framing effect (influence of wording in particular) should be avoided. Third, the relationship of theory and experimental cases must be seriously considered and well handled. Though the cases (often called "vignettes" in X-Phi research) are used in a survey to collect data to justify or falsify a thesis or a theory, we should be careful to avoid theory-laden formulation of them in the experimental design. Finally, experimental cases and control cases should be well selected and matched.

It should never be ignored that the method of survey also has some obvious deficiencies. First, the data we obtain from a questionnaire survey is not real language use or authentic use of the concept under investigation, but only participants' intuitive judgments about a proposition or their intuitive response to a question in devised cases. Second, a relevant important problem is the validity of survey. We would not know (1) whether the participants have really understood the vignettes (though a common practice is to ask a comprehension question to check their reading after the presentation of each vignette) and (2) whether the judgment they give is a true report of what they think. Still another issue is that there are too many disturbing variables in a survey that might prove difficult to be avoided altogether. If the disturbing variables cannot be taken the best possible control of in experimental design, then the validity of a survey will be devaluated and challenged. Fortunately, the above-mentioned deficiencies of questionnaire survey can be offset by the adoption of other experimental methods available.

3.2. The method of big data

As Devitt (2014) argues, philosophers of language should rely more heavily on evidence from actual usage—both sentences from the corpus and those elicited in controlled experiments. Evidence from survey data are evidently participants' intuitive judgments about questions and propositions rather than actual language use. To collect data of authentic language use, we need to appeal to the method of big data. Even though the big data technology is still in its early stage of development and application, it has provided a useful research approach for social scientists. The concept "Big Data" was first put forward by Fayyad in 1995. The primary function of the big data

method is data mining and its basic objectives are descriptive and predictive. The main techniques include classification, association analysis, cluster analysis and anomaly detection. There are four characteristics of Big Data, namely volume, variety, velocity and value (often abbreviated as "4V"). Compared with the survey method, the adoption of the method of big data enables us to provide huge amounts of data. The method of big data is closely related to computer science. For example, computational simulation experiment and corpus can be applied to empirical studies of philosophy. Devitt suggests the use of corpus both informally and scientifically gathered to study the theory of reference. One way to gather direct evidence is to look at the corpus of usage. (Devitt, 2012) Leonelli (2012) and Pietsch (2013) have discussed the function of big data in philosophy of science, but it should also be realized that big data can be equally well applied to X-Phi studies. Devitt (2012) demonstrated the application of corpus to philosophical study with an example.

The 'Godel' vignette contains eight uses of the name 'Godel' (and one mention). Now consider the question: Who do these uses refer to? MMNS (short for Machery, Mallon, Nichols, and Stich) are the authors of this vignette and there can be no doubt that these philosophers are fully competent with the name 'Godel'. And the referent of this name out of the mouths of the fully competent is to the eminent logician who did in fact prove the incompleteness of arithmetic and spent many years at Princeton. So that is whom MMNS's eight uses of the name in the vignette refer to. But then their use of the name in the following passage disconfirms the description theory: "Now suppose that Godel was not the author of this theorem. A man called 'Schmidt', whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question." For, if MMNS's use of 'Godel' refers to that eminent logician in virtue of their associating with it the description, 'the prover of the incompleteness of arithmetic', this passage is not something that MMNS would be disposed to say. They would not, in one and the same breath, both refer to Godel and suppose away the basis of that reference. Similarly, according to the theory that the reference of 'bachelor' is determined partly by its association with 'unmarried', competent speakers would not be disposed to say: "Suppose that the bachelors in Iceland are married." But here the description theory seems to survive because we would not be disposed to say this. (Devitt, 2012: 28)

This example shows how we may use corpora to help investigate a theory of reference. Hacquard and Wellwood (2012) carried out a corpus-based study on epistemic modals in English, with the aim of empirically investigating whether epistemics contribute to the truth conditions of the sentences they appear in. In a sense, the method of big data and the method of survey well complements each other, and the two experimental methods can provide intuition-based claims with important empirical evidence.

3.3. The method of cognitive neuroscience

Cognitive neuroscience aims to uncover the basis of the brain function of the mind and emotional activities. The study of cognitive neuroscience covers such research areas as the study of language, thinking, and intelligence, the relationship of language and the brain, the neural basis of language processing, etc. The experimental methods of cognitive neuroscience are also more and more extensively used in philosophical research, including brain wave, brain imaging, microelectrode, eye movement, ERP, and fMRI. A great merit of the neuroscientific method is that it is able to visualize brain activities. By recoding different material change, we are able to observe the physical images of specific parts of the brain carrying out advanced functional activities from outside the brain,

and study cognition from the way the brain works. Crutch and Warrington (2004) has conducted a neuroscientific experiment on the semantic organization of proper names.

The subject of this experiment is an old lady named AZ. A CT scan revealed a large cerebral infarction in the territory of the left middle cerebral artery involving the parietal, temporal and posterior frontal regions but sparing the basal ganglia. Influenced by the damage in the brain district, AZ's naming abilities were severely impaired. The researchers designed seven experiments to investigate AZ's semantic ability: 1) semantic distance effects in identification of person names; 2) refractoriness in identification of person names; 3) spoken word-written word matching in the broad category of brand names; 4) verbal knowledge of person names; 5) semantic relatedness of historical figures; 6) is person knowledge organized by principles other than occupation? 7) does phonemic similarity influence the semantic organization of proper nouns? The results show that all of the procedures reported in this study have probed for the presence or absence of a semantic distance effect upon AZ's response accuracy.

Greene et al. (2001) and Lindquist et al. (2012) adopted fMRI to study the engagement of emotion and the nature of emotion respectively. In a recent study, Diaz (2019) analyzed the possibilities of neuroimaging methods with the above two experimental studies as examples within the practice of experimental philosophy.

4. Conclusion

We examined in this paper the experimental turn in philosophy of language by (1) analyzing the distinction between and complementarity of armchair theorizing (conceptual analysis in particular) and empirical testing (experimentation); (2) proposing that three types of experimental methods in social sciences and cognitive science are available to philosophers and have been adopted in philosophical investigations, that is, the method of survey, the method of big data, and the method of cognitive neuroscience. The growing popularity of experimental philosophy (and the use of "the experimental turn") does not mean that experimental philosophy is to replace armchair philosophy, but that the former complements the latter by providing necessary empirical evidence. The rise of experimental philosophy symbolizes the return of the tradition of the continuity of philosophy and science and there will continue to be waves of more X-Phi studies with a wider variety of experimental tools and methods in the near future.

Conflict of interest

No conflict of interest was reported by all authors.

Acknowledgements

This work is supported by Philosophy and Social Science Fund of Shanghai (Grant number: 2019EZX008).

References

Cappelen H (2012) *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Colaco D, Buckwalter W, Stich S, et al. (2014) Epistemic intuitions in fake-barn thought experiments. *Episteme* 11(2): 199–212.

Crutch SJ and Warrington EK (2004) The semantic organisation of proper nouns: The case of people and brand names. *Neuropsychologia* 42(5): 584–596.

Dennett DC (1980) The milk of human intentionality (commentary on Searle). *Behavioral and Brain Sciences* 3: 428–430.

Deutsch M (2009) Experimental philosophy and the theory of reference. *Mind and Language* 24(4): 445–466.

Deutsch M (2010) Intuitions, counter-examples, and experimental philosophy. *Review of Philosophy and Psychology* 1(3): 447–460.

Devitt M (2012) Whither experimental semantics? *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia* 27(1): 5–36.

Devitt M (2014) Linguistic intuitions are not "the voice of competence". In: Haug MC (ed) *Philosophical Methodology: The Armchair or the Laboratory?* New York: Routledge, pp.268.

Diaz R (2019) Using fMRI in experimental philosophy: Exploring the prospects. In: Fischer E and Curtis M (eds) *Methodological Advances in Experimental Philosophy.* London: Bloomsbury Academic, pp.131.

Greene JD, Sommerville RB, Nystrom LE, et al. (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293(5537): 2105–2108.

Hacquard V and Wellwood A (2012) Embedding epistemic modals in English: A corpus-based study. *Semantics and Pragmatics* 5(4): 1–29.

Ichikawa JJ (2004) Who needs intuitions? Two experimentalist critiques. In: Booth A and Rowbottom D (eds) *Intuitions*. Oxford: Oxford University Press, pp.232–256.

Ichikawa JJ and Jarvis B (2013) *The Rules of Thought*. Oxford: Oxford University Press.

Kornblith H (2014) Is there room for armchair theorizing in epistemology? In: Haug MC (ed) P*hilosophical Methodology: The Armchair or the Laboratory?* New York: Routledge, pp.195.

Leonelli S (2012) Introduction: Making sense of data-driven research in the biological and biomedical sciences. *Studies in History & Philosophy of Science Part C: Studies in History & Philosophy of Biological & Biomedical Sciences* 43(1): 1–3.

Lindquist KA, Wager TD, Kober H, et al. (2012) The brain basis of emotion: A meta-analytic review. *The Behavioral and Brain Sciences* 35(3): 121.

Machery E, Mallon R, Nichols S, et al. (2004) Semantics, cross-cultural style. *Cognition* 92(3): B1–B12.

Nado J (2016) *Advances in Experimental Philosophy and Philosophical Methodology.* London: Bloomsbury.

Nichols S, Stich S and Weinberg J (2003). Metaskepticism: Meditations in ethno-epistemology. *The Skeptics* 227–247.

Nisbett RE, Peng K, Choi I, et al (2001) Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review* 108(2): 291.

Papineau D (2014) The poverty of conceptual analysis. In: Haug MC (ed) *Philosophical Methodology: The Armchair or the Laboratory?* New York: Routledge, pp.166.

Pietsch W (2013) Big data–the new science of complexity. Unpublished.

Starmans C and Friedman O (2012) The folk conception of knowledge. *Cognition* 124(3): 272–283.

Weinberg JM, Nichols S and Stich S (2001) Normativity and epistemic intuitions. *Philosophical Topics* 29(1/2): 429–460.