Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

School of Social Sciences and Philosophy
Assignment Submission Form

Student Name:	Chenxi Li
Student ID Number:	23330541
Programme Title:	Applied Social Data Science
Module Title:	Research Design for the Social Sciences
Assessment Title:	Assignment 1: Reflective Essay
Lecturer (s):	Dr Martyn Egan
Date Submitted	October 3, 2023

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write

Signed: Chenxi Li

Date: October 3, 2023

Assignment 1: Reflective Essay

Measuring human society with nature science's approaches is just like using a finely woven bamboo basket for water and will finally miss a lot of key points. This is because natural society is bright, orderly, and shows a tendency to be knowable, human society, on the other hand, tends to be vague, and obscure, with unknowability and uncertainty constituting the norm of its operation. The operation of the social world is filled with subjective initiative between people, and many facts are constructed by human activities, but I think it is possible for us to recognize social facts in uncertainty. So, this means that I don't agree with mechanical positivism or constructivism. In fact, the first set of views I want to criticize is exactly classical positivism, ontological realism, and functionalism. The great common denominator of these doctrines is that they abstract a notion called society from daily life, and then carefully remove humans from this conception, so in fact the object of study is the philosophical concept of human society, and under a suspension and bracketed. A classic series of oppositions have been explored around this theme, namely, social objects are knowable or unknowable, society is an entity or an imaginary body, research methods are empirical or interpretive, etc. Just like Della and Keating mentioned (Della & Keating, 2008: 21-23):

For nominalists, categories only exist because we arbitrarily create them. For realists, the categories are there to be discovered… in the natural sciences, where theories have often been formulated and applied before the underlying causal mechanisms have been explicated.

However, as time progresses, this set of opposing views also merges with each other. As Dalla and Keating point out in the same passage, post-positivism is closer to modern scientific approaches, which accept a degree of uncertainty (Della & Keating, 2008: 24). And going one step further, even if the mechanical ontology and methodology of traditional positivism do not work, they start to implement a very important trend of falsificationism, which suggests that scientific theories should be testable and rebuttable. Poper (Poper, 1963: 34-35) made the conclusion beautifully that challenge and verification are very beneficial to the scientific nature of research, and only continuous challenge and verification can prove the correctness of the conclusion. I think falsificationism goes a long way toward alleviating tensions about methodological uncertainty, and people realize that instead of pursuing a grand unified narrative, it is better to turn to more operational and mid-level concepts that can be continuously verified.

And when we start to focus on uncertainty in society, machine learning in statistics is a topic that we must discuss. Unlike traditional research methods, machine learning is a postmodern empirical study in which there is an invisible black box. When the huge arithmetic power of the program intervenes in the research methodology, the research process may become more flexible, but it must also bear the uncertainties involved (Desai, Watson, etc., 2022: 469). Indeed, although machine learning may suffer from technical problems such as overfitting and sample bias, etc., on the one hand, in terms of feasibility, we have the potential to change these technical and overfitting problems as the program iterates and statistical techniques improve (Dietterich & Kong, 1995: 1-2), and on the other hand, in terms of necessity, as the numbers of input variables and possible associations among them increase, the model that captures these relationships becomes more complex which leads statistical inferences become less precise (lj, 2018: 233-234). I hold the view that machine learning in statistics is a huge progress in post-empirical methods, and within this condition, it makes a research method combine, or pluralism, become more and more possible.

So, pluralism is introduced into my thinking at this logic. The feature of pluralism that appeals to me most is that it is a high-risk and at the same time high-reward approach. Researchers are able to choose the most efficient approaches in their projects and can combine different styles if necessary. But the premise of this utopia vision is to understand each side, both quantitative and qualitative, and know the exact situation to choose the most fixed methods. There were scholars already dedicated to making a bridge between these two traditions by listing different assumptions and research goals underlying the two traditions, and just like they mentioned (Mahoney & Goertz, 2006: 245-246):

Misunderstandings across the two traditions are not inevitable. In so far as scholars are conversant in the language of the other tradition and interested in exploring a peaceful and respectful dialogue, they can productively communicate with one another.

Their efforts take the pluralism step further, and there is much research has started to mix these two traditions and get a significant result with it (Caruth, 2013: 112-113; Dawadi, Shrestha & Giri, 2021: 25). However, we must notice the risk in the approaches mixed. If we simply combine methods without considering the applicability of the various methods themselves, and

the maneuverability of working together, then we will only end up with data of very low reliability and validity, which will be a disaster for the final conclusions.

Furthermore, based on my dissatisfaction with positivism and my approval of postmodern positivism, especially in machine learning in statistics, I also need to respond to those who are fixed on social construction theory, on nominalism, or on quantitative methods. I must admit that their methods are full of humanistic care and quite literary, even artistic. But the most important thing is that they fell into an overly subjective situation, which means they could only understand the event from a corner, and lacking a macro-control of the overall thing. Like the blind men and the elephant, without a rigorous demonstration, sample inference will often lead to serious distortion of the population.

After discussing my attitude towards various methods, I also hope to talk about my research interests. Basically, I want to study some feminist issues, and specifically, I am planning to use a large amount of quantitative data, supplemented by interviews with some specific cases, to determine whether gender inequality exists in some social events. For example, one of the questions I am currently curious about is whether women, as a vulnerable group in most societies, are also more susceptible to ideological influences, such as consumerism, which further leads to their difficult situation. More specifically, the body shape and appearance anxiety sold to customers by cosmetics, clothing, and other products seem to have a serious gender orientation, that is, this anxiety is sold more to women and less to men.

For this research interest of mine, classical positivism and functionalism are unreasonable, because gender issues and ideological issues are difficult to simplify, or conclusions can be derived using a few simple formulas; at the same time, construct Theory, nominalism, and the purely qualitative method it points to are also difficult to achieve my goal, because the typicality obtained using this method cannot cover a wider range, and for my research, universality is obviously more important than typicality. Finally, the combination of statistics and machine learning, as well as scientific falsificationism and pluralism clearly meets my needs. On the one hand, the introduction of machine learning makes it easier for me to use various models to describe variables, correlation levels, and changing trends; on the other hand, my research will continue to advance while verifying previous relevant theories. All in all, as far as my research direction is concerned, I hope to avoid useless narratives that are too grand or too precise.

Reference

Caruth, G. D. (2013). Demystifying mixed methods research design: A review of the literature. *Online Submission*, *3*(2), 112-122.

Dawadi, S., Shrestha, S., & Giri, R. A. (2021). Mixed-methods research: A discussion on its types, challenges, and criticisms. *Journal of Practical Studies in Education*, *2*(2), 25-36.

Della Porta, D., & Keating, M. (2008). How many approaches in the social sciences? An epistemological introduction, 21-23

Della Porta, D., & Keating, M. (2008). How many approaches in the social sciences? An epistemological introduction, 24

Desai, J., Watson, D., Wang, V., Taddeo, M., & Floridi, L. (2022). The epistemological foundations of data science: a critical review. *Synthese*, *200*(6), 469.

Dietterich, T. G., & Kong, E. B. (1995). Machine learning bias, statistical bias, and statistical variance of decision tree algorithms, 233-234

Ij, H. (2018). Statistics versus machine learning. *Nat Methods*, *15*(4), 233.

Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political analysis*, *14*(3), 227-249.

Popper, K. R. (1963). Science as falsification. *Conjectures and refutations*, *1*(1963), 33-39