Polling, Power and Legitimacy: The Politics of Representation and the Making of Publics

Lukas Griessl

A thesis submitted for the degree of Doctor of Philosophy in Sociology

Department of Sociology

University of Essex

Date of submission for examination: September 2023

Table of Contents

List of Abbreviations

AAPOR	American Association for Public Opinion Research
ADM	Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (Working Group of German Market and Social Research Institutes)
CAWI	Computer-assisted web interviewing
CPAC	Conservative Political Action Conference
DFB	Deutscher Fußball-Bund e.V. (German Football Association)
ESOMAR	European Society for Opinion and Marketing Research
FPÖ	Freiheitliche Partei Österreichs (Freedom Party of Austria)
GESIS	Leibniz Institute for the Social Sciences, formerly Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (German Social Science Infrastructure Services)
ISC	International Statistical Congress
ISI / IIS	International Statistics Institute / Institut international de statistique
MAR	Missing at Random
MCAR	Missing Completely at Random
NMAR	Not Missing at Random
NIM	Nürnberg Institut für Marktentscheidungen (Nuremberg Institute for Market Decisions), formerly GfK-Verein
ÖVP	Österreichische Volkspartei (Austrian People's Party)
RDD	Random Digit Dialing
SOEP	Socio-Economic Panel
SPÖ	Sozialdemokratische Partei Österreichs (Social Democratic Party of Austria)
TI	Transparency Initiative, AAPOR

Abstract

This dissertation is situated at the cross-section of Science and Technology Studies (STS), philosophy and sociology of science, and social theories of knowledge and ignorance. Building on historical literature on statistical thought and qualitative interviews, this research explores the history and sociology of survey research and polling. The central aim is to trace the emergence and potential dissolution of what is seen as the established scientific criteria governing survey research and polling, with particular attention to a longstanding, contentious controversy surrounding probability and non-probability sampling. Methodologically, the thesis draws on semi-structured interviews with polling experts based in industry and the academy, juxtaposed with a genealogical analysis of the history of different ontological and epistemological assumptions surrounding the validity of different polling techniques. The thesis explores competing paradigms throughout modern history for understanding the representativeness of polls. Drawing on the concept of boundary work, I analyse how proponents of competing polling programmes seek to legitimize their own approach while simultaneously delegitimizing the opposing viewpoint. In looking at different polling controversies and developing the new notion of 'forging stable statistical chains', the thesis explores questions around ontological politics, examining the formation and sustenance of publics through surveys and polls. The research presented in this thesis seeks to provide a sociological analysis of the making of social scientific knowledge. In so doing, I make novel conceptual contributions to sociological understandings of survey research and polling.

Acknowledgements

Since I started working on this dissertation, many people have helped me in conducting this research, be it in the form of supervision, moral support, feedback or funding. In the following paragraphs, I wish to mention all those who have supported me in this journey.

First and foremost, I would like to thank my supervisors Linsey McGoey and Nick Allum. It's difficult to acknowledge their efforts and my appreciation in just a few lines, it simply has been a great journey of learning and development. Right from the start, they believed in me and my project's potential, helped me acquire funding, kept me motivated and helped me grow as a sociologist. Their ideas, guidance, feedback and patience have been invaluable, and this project would not have happened without them. Thank you very much!

I would also like to thank my Board Chair Michael Halewood, whose thoughts and regular commentary on my progress have played a crucial role in shaping this project. I would also like to thank Katy Wheeler and Bethany Morgan Brett for introducing me to the craft of qualitative interviewing and whose comments and advice on the preparation and execution of my empirical research have been essential. I also thank Pablo Cabrera Alvarez for stimulating discussions and feedback on the history of polling. I would also like to thank the leaders of the departmental colloquium, which was a very encouraging environment to present and discuss research and everything concerning the life and struggles of a PhD student.

Outside Essex, I would like to thank Matthias Groß and Sergio Sismondo, who hosted me at their institutions in Leipzig, Germany and Kingston, Canada as a visiting researcher and provided me with thoughts, feedback and an inspiring environment. I would also like to thank Christoph Schiebel, Danny Otto, Manuel Jung, Max Gropper, Niclas Rautenberg, Sebastian Klimasch and Yintan Fan for their conversations, reading, feedback and comradeship throughout the last years and beyond. A great thanks also goes to Julian Molina, for the many conversations, discussions and feedback on my work. I also thank Dominic Lusinchi for his reading and

discussion of an earlier draft of this dissertation. I would like also to thank Joost van Loon, especially for his enduring efforts as a mentor when drafting my first ideas for this project. Last but not least, I would like to thank all my interviewees and informants, whose narratives, knowledge and insights form the core of this dissertation.

I would also like to extend my gratitude to my examiners, William Davies (Goldsmiths, University of London) and Tara Mahfoud (University of Essex), for their efforts, time and thorough review and insightful feedback on my dissertation. Their questions and suggestions have been invaluable.

I also want to thank my parents, Maria and Johann, and my siblings Philipp and Johannes for their love, support and for always being there for me. Lastly, I wish to express my gratitude to my wife and best friend, Merve, for her support, encouragement, and for standing by me throughout this long journey and beyond.

CHAPTER 1

1 Introduction

On the 14th of May 2018, one day before the German national football team was set to announce its preliminary nominations for the football World Cup in Russia, pictures of German national players Mesut Özil and İlkay Gündoğan and then-Everton striker Cenk Tosun, posing with the Turkish president Recep Tayyip Erdoğan appeared. The pictures were taken at an event in London's Four Seasons hotel and show the football players with the Turkish president, whom they presented football shirts from their respective English clubs Arsenal, Manchester City and Everton F.C. The shirt that Gündoğan gave President Erdoğan, who was at this time campaigning for Turkish elections to be held around one month later, bore the message: "Sayın Cumhurbaşkanım'a saygılarımla", which translates to "To my president, with my respects". To Erdoğan, who is banned from holding campaign rallies in Germany, this was a welcome opportunity to reach the around 1.2 million Turkish citizens in Germany who were eligible to vote during these elections.

Those pictures led to a controversy over the footballers' questioned loyalty towards Germany and their place in the German national football team. Some even called for Özil and Gündoğan to be eliminated from the German National Team. Not knowing or ignoring the fact that football players are prohibited from switching national teams after having represented a country at the senior level, Sebastian Münzenmaier, an MP for the right-wing *Alternative für Deutschland* (AfD), for instance, commented in a video statement that if these players would not stop "flattering" the Turkish president, they should start playing for the Turkish national team. In a tweet in which he shared this statement, he directly addressed the two football players: "To the

two [passport] German[1] footballers #Özil & #Gündogan: Dear Mesut, dear Ilkay, you have to decide - the president of Germany is called Steinmeier and is based in Berlin! If #Erdogan is your president, why do you play football for our country?"

The controversy surrounding these pictures was pervasive in Germany at this time. The issue of whether the players should remain part of the German national football team during the World Cup became a matter of public concern. This controversy also led to accusations that the debates over these pictures were responsible for Germany's poor performance during this World Cup, a debate that culminated in Özil resigning from the national team on the 22nd of July in 2018.[2]

In what seems like an important occurrence at the intersection of politics, identity and football, it also constitutes an important episode in contemporary debates around polling and surveys, touching on the question of how to best 'know' a population and their opinions by means of a sample: One day after the photo shoot, the German News Magazine Focus Online published a survey executed by a public opinion and market research company called Civey. Civey, a newcomer in the field was founded in 2015 and conducts all its surveys via the internet. According to this survey, which was declared to be representative, 80% of the over-18s in Germany no longer wanted to see Mesut Özil and İlkay Gündoğan to play for the German national team. According to another survey conducted by a company called Forsa, a much-

---

[1] Münzenmaier uses the pejorative term "passdeutsche", which states that people are only German due to their identity documents, not their heritage. The term is difficult to translate into English. The term 'German citizen of non-German origin', for instance, does not grasp the pejorative and negative connotation the German term entails. It is rather used to express that there are "real Germans" and Germans of inferior quality, referring to German citizens who themselves or their ancestors were not born in Germany, who hold foreign-sounding names and who therefore are not seen as real citizens (for a discussion, see Martens, 2018).

[2] In his resignation letter, which he wrote in English and shared via Twitter, he stated that "having a picture with President Erdoğan wasn't about politics or elections, it was about me respecting the highest office of my family's country. […] The treatment I have received from the DFB and many others makes me no longer want to wear the German national team shirt. I feel unwanted and think what I have achieved since my international debut in 2009 has been forgotten. […] I am German when we win but I am an immigrant when we lose" (Özil, 2018).

longer-standing polling company in Germany, founded in 1984 in Cologne but now headquartered in Berlin, the situation was somewhat different. According to this survey, 61% stated that the footballers should not have taken part in the photo shoot, but despite that, 71% saw no reason not to nominate the two players for the World Cup. Despite their incompatible findings, both outcomes were declared to be representative of the German population. If taking both surveys at face value, there thus seem to be people who want these two football players to play and at the same time not play for the German national team, a polling version of Schrödinger's cat.[3] The two polls, to put differently, enacted two different versions of German public opinion.

Following the publication of the first survey, three established German opinion research companies, the already mentioned Forsa, Infas and Forschungsgruppe Wahlen filed a complaint against Focus Online with the German Press Council for publishing the mentioned survey conducted by Civey. The complaint brought to the German Press Council stated that the survey uses a design, "which, according to the generally accepted scientific criteria of empirical social research, is fundamentally not suitable for delivering 'representative' results" (marktforschung.de, 2018, Translation). According to this complaint, the magazine thus crossed a line in publishing a survey whose methodology was violating what, according to the complainants, corresponds to the generally accepted scientific criteria of empirical social research. On the 4th December 2018, the German Press Council decided that from a press ethics perspective, the usage of Civey surveys is unproblematic, since, as they argue, examining the scientific basis of surveys is not part of journalistic due diligence.

---

[3]   Mark Pack (2022, pp. 93–94) discusses the notion of Schrödinger's cat in a discussion of polls regarding public sentiments towards Brexit. This issue often arises in cases of agree–disagree-type questions and is termed acquiescence bias, meaning that people tend to agree to certain kinds of statements.

Manfred Güllner, founder and director of Forsa, interpreted this decision in an interview with me when he stated that Civey "has not been whitewashed [by the Press Council] in saying that these were representative results, they have only said that the media do not have to check whether they are representative". What this means, however, so Güllner, is that "if an institute says that something is representative, then the press is entitled to accept that and does not have to verify it further" (Güllner, Interview, Translation). To rephrase Güllner's statement using a different terminology, he seems to suggest that in the contemporary climate, representations of publics by surveys and polls do not need to be anchored in something tangible, as long as they are labelled to be representative. Instead, Güllner and his co-complainants demand those representations to be anchored in a particular methodology. In pressing that Civey's survey design violates those accepted scientific criteria, one can interpret this not only as a complaint directed towards Civey or Focus Online per se but towards the used survey design.

The matter of contention concerning survey design revolves around the distinction between probability and non-probability approaches to sampling. The disparity between these approaches stems from the knowledge of the likelihood of selecting a participant. When representing a population statistically through a probability sample, every individual in the population must have an equal or known chance of entering the sample. The most effective method for achieving this is by randomly selecting a sample from a predetermined sampling frame. On the other hand, non-probability samples do not grant each potential participant an equal or known chance of selection; instead, participants typically recruit themselves, usually leading to a highly skewed sample, which is adjusted after the data has been gathered. Probability sampling is widely regarded as the benchmark in the field since it is based on mathematical principles. Non-probability sampling deviates from those core principles that have guided the landscape of surveys and polling since the 1950s. The reason for the rapid increase in non-probability sampling in recent years lies in the practical problems – such as increasing non-response and

thus increased costs – that traditional surveys are more and more prone to encounter. Furthermore, the internet offers promising new ways to gather data, an opportunity that companies like Civey tap into. While the justification for reaching accurate results in probability sampling lies in its design, informed by statistical theory; the justification for non-probability sampling lies in model assumption, as to how the sample differs from the population.

The transition from probability to non-probability sampling signifies a profound change in the manner in which survey researchers and pollsters attain knowledge about a population. This transformation not only engenders a fundamental change in methodology but also ignites polarization and fragmentation within the community, giving rise to practices of boundary work and delegitimizing alternative approaches. Within this intricate context, questions of who possesses the authority and legitimation to represent publics and their opinion arise. This issue becomes even more pressing in light of the evolving accessibility to survey instruments and new ways of tapping into public opinion beyond classical surveys and traditional survey providers. Those developments challenge the long-standing monopoly of those traditionally endowed with the authority to represent publics and their opinion. The political relevance of those developments becomes clear when considering the rising number of controversial polls and surveys being published and disseminated around all sorts of issues. Therefore, this dissertation aims to disentangle the multifaceted components inherent in the current transformation experienced by polling and survey research.

My research questions are as follows:

1) *What is the historical development of survey sampling as a technique for constructing representations of publics?*
2) *How has the legitimacy of polling as a measure of public opinion been established and consolidated over time?*

3) *Which actors are endowed with the power and authority to create legitimate representations of public opinion?*

4) *How has the shift towards non-probability sampling occurred and which discursive positions developed in the field?*

To answer these questions, I adopted a combination of qualitative interviews, a re-evaluation of historical and philosophical literature on the history and epistemology of polling and other types of representative design, and an analysis of archival material and media reporting on polling controversies across late modernity. I conducted 20 semi-structured interviews with survey researchers, pollsters and statisticians as a way to access the field and to get an understanding of the different viewpoints and perspectives. Furthermore, literature on the history of statistical thought and the development of survey research and polling allowed me to put those findings into context and to triangulate statements from the interviews with the literature.

The structure of the thesis is as follows: Following a chapter dedicated to the Literature Review, Methodology, Study Design and Epistemology, the dissertation is structured into six substantive chapters (chapters 3-8), each of which contributes to the overarching objective of this thesis, which is to investigate the challenges faced by pollsters and survey researchers in their attempts to gain a comprehensive understanding of a population's opinions through sampling. The third chapter explores the historical trajectory of survey sampling as a method to know populations. Therefore, it undertakes a historical genealogy that traces the origins of sampling as an accepted scientific method, observing its division into probability and non-probability sampling, a distinction that was laid down early on. While there were earlier precursors in the 17th and 18th centuries, sampling gained substantial traction towards the end of the 19th century within the realm of the International Statistics Institute (ISI). Initially meeting resistance from the majority within the field, both sampling approaches eventually gained acceptance by the mid-1920s. Already at this juncture, debates and controversies regarding the

superiority of these two approaches pervaded the field, eventually concluding in 1934, favoring probability sampling.

The fourth chapter explores the solidification of probability sampling as the gold standard in survey research and polling and investigates how such developments have contributed to the formation of a specific understanding of public opinion. This chapter aims to examine the process by which the credibility of polling as a means to gauge public opinion has been established and reinforced. In this chapter, I contend that the early pioneers in election forecasting and polling redefined the notion of public opinion as the statistical aggregation of individual opinions, pointing at a conceptual framework that aligns with the epistemological assumptions inherent in probability sampling. In so doing, pollsters were able to assert a monopoly over both the access to and representation of public opinion.

The fifth chapter analyzes scholarly literature on the topic of representation, setting the stage for my later exploration, in chapter 8, of three recent polling controversies: a manipulated poll in Austria, involving the former chancellor Sebastian Kurz, a - most likely - invented poll in the US, involving the rap-rock musician Kid Rock and Donald Trump's relationship to polls. Chapter 5 also introduces two conceptualizations of representation: representation as mirroring and representation as performance. While the former assumes that survey and polling outcomes mirror the characteristics of the population being studied, providing a reflection on the individuals represented, the latter posits that processes of representation actively construct and shape the realities that they purport to describe. Furthermore, this chapter delves into the factors that have contributed to the growing trend towards non-probability sampling, a development that started in the 1990s. In this context, this chapter examines the perception and the sentiments of practitioners towards those changes.

While those first chapters are mostly focused on the historical emergence of what the plaintiffs in the case concerning the photo shoot and the associated Civey survey deemed as the

"accepted scientific criteria of empirical social research," chapters five through eight shift the focus to the present era, drawing empirically on my interviews conducted with pollsters and survey researchers. Chapter six is guided by the question of *how the shift towards non-probability sampling occurred and which discursive positions developed in the field*. At the heart of this chapter lie two narratives of conversion, featuring individuals who transitioned from advocating for probability sampling to endorsing non-probability sampling, and vice versa. These conversions serve as instances where the practices of boundary work, aimed at establishing one approach as the exclusive and legitimate means of generating knowledge about a population, begin to manifest. This phenomenon will be the central focus of the seventh chapter.

The seventh chapter builds upon the previous chapter, delving into the intricate dynamics of the controversy at hand. It examines how individuals involved in the discourse defend their positions within the realms of probability or non-probability sampling, shedding light on the way boundaries and legitimacy are negotiated in the field. In so doing, this chapter revolves around different themes that were apparent in the empirical material conducted for this research. Its overarching goal is to address the question of *how power struggles over the epistemology of polling shape competing statistical method*s. This chapter engages with recent work of second-generation scholars from within the economies of convention approach, highlighting how social science methodology evolves from the collective experiences gained in scientific research, rather than following a universal logic. The controversy is characterized by participants engaging in forms of boundary work, establishing criteria for what counts as scientific and what does not.

The eighth chapter aims to synthesize those various strands of thought to propose a conceptual framework for understanding the making and stabilization of public opinion in relation to polls and surveys. In this chapter, I argue that previous treatments of statistical chains from the history and sociology of quantification are insufficient for explaining how publics are created through processes of quantification. To address this, I add a focus on 'forging' and

stability to previous discussions of statistical chains. This chapter therefore revolves around questions of *what the discursive and material conditions in the formation of publics through polling practices are.* Through several steps, I introduce and defend the concept of 'stable statistical chains', drawing on recent work in the tradition of the French Economies of Convention, Actor-Network Theory, and work by political philosopher Jacques Rancière. Those different, yet related, approaches are central to my discussion of ontological politics, where I draw on the contemporary examples mentioned above – recent controversies involving Kurz, Kid Rock and Trump – to illustrate the theoretical value of my notion of stable statistical chains.

Together the research presented in this thesis seeks to provide a sociological analysis of the making of social scientific knowledge. In so doing, I make novel conceptual contributions to sociological understandings of survey research and polling.

CHAPTER 2

2 Literature Review, Methodology, Study Design and Epistemological Approach

2.1 Introduction

This chapter provides an overview of the literature, methodology, study design and the epistemological approach applied to this research project. This chapter is divided into four parts in which I first introduce various core literatures from the philosophy and sociology of science and the construction of scientific facts that the subsequent dissertation draws and builds upon. After this, I introduce my methodology in detail, followed by an elaboration on the study design, including a presentation of interviewees and the epistemological approach used. In so doing, I aim to provide a throughout background on which this dissertation developed.

2.2 Literature Review: The Making of Publics Through Quantification

To offer a comprehensive historical and sociological examination of contemporary controversies within the realm of polling and survey research, this dissertation develops in dialogue with four interconnected strands of scholarly inquiry encompassing sociology, science and technology studies, political theory, and philosophy. In particular, those different streams of intellectual thought deal with scholarship concerned with the social studies of science and knowledge collectives; social studies of quantification and the making of truth and facts, political theories of representation, social studies of survey research and polling, as well as the field of ignorance studies. By weaving together these strands, the dissertation not only enhances understanding of the emergence, practice and significance of polling and survey research but also provides a

framework through which it is possible to grasp the relationship between the outcomes of surveys and polls and the population being surveyed.

2.2.1 The Social Construction of Facts

A central tenet for this dissertation is the field of Science and Technology Studies (STS), which explores the "institutions, practices, meanings, and outcomes of science and technology and their multiple entanglements with the worlds people inhabit, their lives, and their values" (Felt *et al.*, 2017, p. 1). Social studies of science usually view scientific developments to be influenced by the social conditions and context in which they take place. This view is found not only in contemporary work in the sociology of science and STS, but can be found in many classical and canonical texts in sociology (see e.g. Baber, 2000). In *Grundrisse*, first published posthumously in 1939, Marx, for instance, had already seen the rise of industrial capitalism as a crucial factor in "the development of the natural sciences to their highest point" (Marx, 1993, p. 409) and contends that the "theoretical discovery of autonomous laws appears merely as a ruse so as to subjugate it under human needs" (Marx, 1993, p. 410). Instead of viewing scientific discoveries and laws as being purely objective and independent of human influence, he views them as being shaped to serve human needs within a capitalist society.

Another noteworthy development was the Sociology of Scientific Knowledge (SSK). While the idea that much of people's ordinary knowledge is merely a result of the society people are born into and raised, constitutes a crucial element to the sociology of knowledge, the sociology of scientific knowledge (SSK) takes this further and claims that this also holds true for scientific knowledge.

To exemplify this development, it is worth looking into how Mannheim (1998), first published in 1928, for instance, distinguished between the 'existential determination of thought' by 'extra-theoretical factors' and the development of 'immanent laws' derived from the 'nature of things' of 'pure logical possibilities'. Some forms of knowledge are perceived to be grounded

in society, whereas knowledge in the natural sciences is believed to be guided by nature or logic, which is why Mannheim kept the physical sciences and mathematics out of the realm of the sociology of knowledge. To the early sociology of knowledge, scientific and mathematical knowledge was thus understood as purely governed by the scientific method. Starting from the 1970s, sociologists, philosophers and historians initiated programs of analysis and research, resisting this epistemological exceptionalism and engaged in an anti-positivist turn with the demand of treating scientific knowledge on a par with other forms of knowledge.

The emergence of SSK is also associated with what has become known as the science wars in the 1990s (Mermin, 2008; see e.g., Lynch, 2020). Those wars revolved around the clash between natural scientists, and scholars from fields such as literary studies and the social sciences. The former group positioned themselves as representatives of scientific knowledge and launched assaults on work conducted by the latter. Prominent figures included Paul Gross, Norman Levitt, Alan Sokal, and Lewis Wolpert. One significant event during this period was physicist Alan Sokal's publication of a deceitful article in the journal *Social Text* (Sokal, 1996). This article aimed to expose the lack of critical analysis performed by the journal's editors. The development of SSK in the 1970s played a crucial role in triggering the controversy, as it sought to investigate the social context and content of scientific endeavours, which garnered criticism from scientists who perceived it as a direct challenge to scientific truth.

What we take from those preliminary remarks is that a central element within the social studies of science is to view scientific truth as the consensus within a scientific community and not as an objective fact. Throughout this section, I explore five influential streams of intellectual discourse, which I will take up throughout this dissertation. The following section begins with an exploration of the work of Ludwik Fleck (1896-1961), followed by the work conducted by Thomas Kuhn (1922-1996), Imre Lakatos (1922-1974), Bruno Latour (1947-2022) and Karen Barad (1956-). While this enumeration of scholars is not intended to be exhaustive, their seminal

contributions hold significant value in elucidating the origins, evolution, and decline of developments within the realm of survey research and polling, which this thesis sets out to do.

2.2.1.1 Fleck: Thought Styles and Thought Collectives

Ludwik Fleck was a Polish biologist and physicist but also an epistemologist, whose work, in particular his *Genesis and Development of a Scientific Fact,* first published in 1935, has been foundational in the philosophy and sociology of science. Based on his own laboratory experiences and a historical reconstruction of how the concept of "syphilis" emerged, he argued that there is no pure and direct observation, but rather that scientists are always influenced by historical and environmental factors. On this basis, scientists develop a *thought style*, understood as a "directed perception, with corresponding mental and objective assimilation of what has been so perceived" (Fleck, 1981, p. 99). A *thought style* corresponds to a *thought collective,* a "community of persons mutually exchanging ideas or maintaining intellectual interaction" (Fleck, 1981, p. 39). A thought collective, in Fleck's understanding, is the carrier of a particular *thought style*. In this sense, there are different ways of looking at the same facts, which can appear to be *incommensurable.*[4] Furthermore, a scientific fact is defined as a "thought stylized conceptual relation which can be investigated from the point of view of history and from that of psychology, both individual and collective" (Fleck, 1981, p. 83). This means that while scientists may think their work is solely grounded in scientific evidence, it is actually the cognitive mechanisms that allow for the construction of scientific facts. A particular *thought style* frames observations and knowledge systems. Fleck's *thought styles* entail a sociological examination of science, which is one of the differences from Hacking's rather epistemological notion of *styles of reasoning* (for a discussion on the differences and similarities, see: Sciortino, 2017).

---

[4] It is interesting to note that the notion of *incommensurability* first appeared in the writings of Fleck and not in the work of Kuhn, where it is oftentimes thought to have originated.

2.2.1.2   Kuhn: Paradigms, Crises and Revolutions

Fleck's *Genesis and Development of a Scientific Fact* was rather unknown until it found prominence, particularly through its rediscovery by Thomas Samuel Kuhn, who stated that Fleck's essay "anticipates many of my own ideas" (Kuhn, 2012, p. xii) and which has been highly influential for his work. Kuhn was an American physicist, who later became one the most influential historian and philosopher of science. Similar to Fleck, his seminal work *The Structure of Scientific Revolutions* puts forth the argument that science is always a community-based activity and when studying science, one must "first unravel the changing community structure of the sciences over time" (Kuhn, 2012, pp. 178–9). Science is hereby understood "as a product as much of culture as of nature" (Collins, 2015, p. 308). An important notion in this work is that of *incommensurability*, which is already prevalent in Fleck, meaning that scientists who belong to different paradigms "practice their trades in different worlds" (Kuhn, 2012, p. 149) and "must fail to make complete contact with each other's viewpoints" (Kuhn, 2012, p. 147).

To Kuhn, scientific truth constantly evolves through paradigm shifts, which happens when a dominant paradigm or theoretical framework in which scientists work is replaced by a new emerging paradigm. A paradigm offers a stable and structured framework in which scientists can successfully conduct research. Such periods are called *normal science* and can be characterised by the practice of *puzzle-solving,* meaning that problems can be solved by drawing on the framework provided by the paradigm. A crisis in a discipline can arise when the confidence in the ability of a certain paradigm to solve particularly worrying puzzles declines. Those situations are called 'anomalies'. At some point, anomalies accumulate and become real problems that cannot be solved by drawing on the framework available. Those situations of crisis, usually lead to younger scientists developing alternative frameworks, which might eventually constitute a new paradigm and thus a new period of normal science. The most "radical implication" found in Kuhn's work is, as Sismondo puts it, that "[s]cience does not track the truth but creates different

partial views that can be considered to contain truth only by people who hold those views" (Sismondo, 2011, p. 16). Understood in this way, science is the practice of doing science, rather than the discovery of objective facts.

As I will show later in this dissertation, the notion of paradigms, as developed by Kuhn, has also been mobilised by my interviewees when talking about developments and changes in the field of survey research and polling. Ulrich Rendtel, for instance, stated that he was impressed by the idea that "mostly paradigms do not change because people have been convinced, but because the followers of a certain paradigm simply retire, die and others grow up [laughs]." In so doing, he draws on a conceptualisation of scientific progress that implicitly refers to that of Kuhn, and uses it to legitimate certain approaches with the argument that it is just the way things are. The argument that paradigms simply change without committing to an underlying concept of science, has also become subject to criticism, for example by Imre Lakatos, whom I will introduce in the next section. Another interviewee, Rainer Schnell, who has a completely different perspective on sampling than Rendtel directly referred to Kuhn to argue that the sampling approach that he opposes does not constitute a paradigm in a Kuhnian sense: "Well, Kuhn's paradigm concept, as vague as it is in the original, has as its essential criterion that there must be a successful model solution. That is, someone must have demonstrated that it actually works. And the term paradigm is always used so vaguely, in the sense of a worldview. But that's not what Kuhn means. […] [I]t means that there is actually a successful application […] and with non-probably samples I don't see one. So, I don't see a successful demonstration of solving a problem, so beyond magic tricks." In those two examples, we can identify different ways in which Kuhn's work is understood and even politicized: his criteria to delineate scientific practices from non-scientific ones, as well as the idea of progress and revolutions have here been mobilised in a way to legitimate one's own approaches and delegitimate others' approaches in the discipline.

Even though Fleck and Kuhn's work share many similarities and parallels, it is particularly Fleck's radical understanding of science as a social process through which both conceptions differ. Whilst to Fleck, there is no fundamental difference between scientific knowledge and other forms of knowledge, the emergence of a paradigm in Kuhn's sense can be seen as a social closure and thus a characteristic of science. To Fleck, thought styles are always in flux; however, thought constraints might characterise forms of scientific thinking.

2.2.1.3 Lakatos: Not one, but Multiple Paradigms

The publication of Kuhn's *Structure* led to heated discussion among those studying science and scientific practice. One view that emerged in interaction with and as a response to Kuhn was the work of Hungarian mathematician, physicist and philosopher of science Imre Lakatos. In *The Methodology of Scientific Research Programmes*, Lakatos (1978) suggested the idea of a research program, which, different to a paradigm in Kuhn's sense, is usually multiple.[5] Thus, while to Kuhn, there is one paradigm in a given scientific field, to Lakatos, there is usually more than one at the same time, there are "larger paradigm-like units [that] operate in parallel and compete in an ongoing way" (Godfrey-Smith, 2003, p. 102).

Whilst acknowledging Kuhn's historical arguments, Lakatos was critical of Kuhn's work in the sense that science appeared as an endeavour in which the loudest and most numerous voices would dictate the course. Whereas to Kuhn, one paradigm replaces another after anomalies accumulate, to Lakatos, what we find when observing science is a competition between different research programmes. Scientists who operate within a research programme are committed to this

---

[5] As we shall see in due course, Lakatos' view seems intuitive from the perspective of the history of survey sampling. Whilst one may argue that the idea of sampling constitutes a paradigm in a Kuhnian sense, one that replaced the previous paradigm of full enumeration, we can see that once the idea of sampling was generally accepted, different versions of this idea coexisted and replaced each other as the dominant one, but none of these ever became the exclusive way of doing sampling. By exploring the historical development of sampling and polling, it becomes clear that the notion of a single paradigm at any given time fails to account for the actual dynamics and processes observed within this field.

programme, trying to modify their theories in the case of anomalies and problems. Despite this, research programmes are sometimes abandoned, just like paradigms in Kuhn's sense.

When it comes to changes within a research programme, change must only occur to the protective belt of the programme, not its hard core. He therefore distinguishes between a negative and a positive heuristic. The *negative heuristic* of a research programme forbids "to direct the modus tollens at this 'hard core'" (Lakatos, 1978, p. 48), meaning that the core of the programme must remain stable and unchanged. Lakatos also describes a *positive heuristic*, which "consists of a partially articulated set of suggestions or hints on how to change, develop the 'refutable variants' of the research programme, how to modify, sophisticate, the 'refutable' protective belt" (Lakatos, 1978, p. 50). This heuristics "saves the scientist from becoming confused by the ocean of anomalies" (Lakatos, 1978, p. 50). Thus, whereas the protective belt is 'refutable', the hard core is 'irrefutable' by the methodological choices of its proponents. Changes in a research programme ought to be progressive, which means that its application must extend to a larger set of cases. While changes must only be made to the protective belt, those should also be progressive. While a progressive research programme "is one that is succeeding in increasing its predictive power", a degenerating research programme is one in which changes "are being made […] only [to] serve to cover existing problems" (Godfrey-Smith, 2003, p. 105). In this sense, while a degenerative research programme struggles to catch up with anomalies, a progressive research programme is able to extend its own range to new phenomena. Importantly, to Lakatos, it is permissible to protect a research programme, even when it is degenerative, since there is always the chance of it recovering. This is insightful when exploring the history of sampling and polling. As we shall see, up until the end of the 1940s, random and purposive sampling were both research programs that scientists and practitioners could choose from. After the 1948 US presidential elections, the purposive sample lost its predictive power and became a degenerative

research programme. It was, however, never abandoned and is now, in the context of the rise of the internet, being recovered again.

Another relevant perspective, which also assumes multiple research practices was developed by Knorr-Cetina (1999), who suggests looking at the *epistemic cultures*, defined as "amalgams of arrangements and mechanisms-bonded through affinity, necessity, and historical coincidence which, in a given field, make up how we know what we know" (Knorr-Cetina, 1999, p. 1). This approach highlights that there is a "diversity" of epistemic cultures and thus a "disunity" in science. Epistemic cultures are the "machineries of knowledge construction" and thus the practices and cultures of generating scientific knowledge.

2.2.1.4 Latour: A Radical Constructivism

While to Kuhn, science processes through revolutions, and to Lakatos through research programmes, Latour describes exercises of enrolment, interessement and other elements involved in fostering scientific progress. What those have in common, however, is, as argued by Lynch (2012), that scientific success is contingent: "[M]any would-be revolutions fail to catch on, and many efforts to enrol allies are ignored or rebuffed" (Lynch, 2012, p. 450). Crucial for Latour's perspective is to account for a greater role of nature or non-humans in the making of scientific facts. This is best visible in the title of the new edition of Latour and Woolgar's ethnographic study. Its 1979 edition *Laboratory life. The social construction of scientific facts* (Latour and Woolgar, 1979) turned to *Laboratory life. The construction of scientific facts* (Latour and Woolgar, 1986) in the 1986 edition. This slight amendment of the title, the removal of the word "social", thus signifies the idea of a radical constructivism, in which facts are mutually constructed by humans and non-humans. The way in which material entities are included in Latour's Actor-Network Theory is in recasting humans and non-humans as actants, which treats non-humans as quasi-agents. This will also be crucial for this research, as it emphasises e.g., technology or models and other non-human actors in the development of statistical knowledge.

Furthermore, Latour's (2005) definition and understanding of the notion of *constructivism* bears important relevance: "When we say that a fact is constructed, we simply mean that we account for the solid objective reality by mobilizing various entities whose assemblage could fail" (Latour, 2005, p. 91). To Latour, "constructivism [is a] synonym for an increase in realism" (Latour, 2005, p. 92), a view that goes hand in hand with Desrosières' (1998) argument of statistical numbers having a dual nature of *constructivism* and *realism,* which plays a crucial role in the social studies of quantification.

2.2.1.5 Barad: Ontological Entanglements and Intra-acting Performativities

An additional important source for this dissertation is the work of Karen Barad. Influenced by Nils Bohr's philosophy-physics, Barad develops an approach to understanding measurement: For Bohr, quantum physics needs a new logical framework, which considers processes of observation, a notion from which Barad develops the concept of agential-realism. Observation is possible with the help of an apparatus in between the object and the observer, which cannot be subtracted from the result: "Since observation involves an indeterminable discontinuous interaction, as a matter of principle, there is no unambiguous way to differentiate between the 'object' and the 'agencies of observation' - no inherent/naturally occurring/fixed/universal/Cartesian cut exists. Hence, observations do not refer to objects of an independent reality" (Barad, 1996, p. 170). The apparatus and so the measurement is the instance where "matter and meaning meet" (Barad, 1996, p. 166). The observed phenomenon exists only together with the whole apparatus which does not only include, in our case the sample survey and opinion polls, but also the historical and political background which led to its very development. A polling chart, to remain at this example, does not provide a representation of public opinion, but must be seen as a wholeness. The apparent cut between public opinion on one hand and the sample survey on the other is a constructed one that creates the very phenomenon of public opinion. The questionnaire, the sampling methodology and it's

dissemination are in fact part of the object one wants to further know. This is what Barad describes with the notion of onto-epistemology, no longer separating between epistemology and ontology, but highlighting the influence of how we know on what we know. The constructed cut introduces the notion of an observer and an observed object. As Barad (1996) points out, "[r]eality is not composed of things-in-themselves or things-behind-phenomena, but things-in-phenomena" (Barad, 1996: 176). Following this framework, polling and survey research does not reveal the reality of public opinion, it performs different versions of it. It is not a fixed ontology which stands alone – it is always reconstituted through 'material-discursive intra-actions'. What exists, which would fall under the realm of ontology is not separable from how things are known about the world, which falls under the realm of epistemology. Due to those entanglements, scientists are not only responsible for the creation of knowledge but also for what exists, for which Barad introduces the notion of ethico-onto-epistemology (Barad, 2007, p. 90).

2.2.2　Social Studies of Quantification

2.2.2.1　Quantifying Everything

A seminal reference to the history of probability and quantification is Ian Hacking's observation that in the late-nineteenth century "almost no domain of human enquiry is left untouched by the events that I call the avalanche of numbers, the erosion of determinism and the taming of chance" (Hacking, 1991, p. 189). In his work, Hacking shows how during a certain period of time, it became possible to think of statistical patterns as explanatory and to understand the world as no longer deterministic. During this period, populations were not only counted, the 'avalanche of numbers' also led to a new ordering and classification of populations. In this sense, statistics became crucial for governance and "helped determine the form of laws about society and the character of social facts [so that it became] itself part of the technology of power in a modern state" (Hacking, 1991, p. 181). Hacking refers to the emergence of statistics as the emergence of a new "style of reasoning", which conditions of possibility he describes. Understanding how

apparent chance effects could be described as statistical regularities, one was able to formulate and test hypotheses about the behaviour of groups and individuals. In a similar vein, Daston (1988) is also interested in the emergence of probability, however, arguing that there was a "classical" style of mathematical thinking, which was abandoned by 1840. Hacking, instead, describes a continuous development of probability theory from 1660 to the present.

The history of statistics can, according to Prince (2019), be divided into three periods. Even though the work of Prince originated in the field of human geography, it plays a key role in this thesis, as it draws together a variety of themes, authors and current developments in the social studies of quantification. The first spans from the late 18th century to the 19th century and consists of the rise of statistics as a crucial component of public life in Europe and led to the creation of governmental departments and ministries to collect statistics. This set the stage for the second period, which Prince (2019) locates in the 20th century. During this period, a shift in scale took place, leading to a significant increase in the production of statistics. Institutions produced statistics on a larger scale, surpassing the confines of the nation-state. The third period spans from the late 20th century to the present. Since this period, the production of statistics is no longer mainly in the hands of state institutions and universities, and an expansion of agencies dedicated to data collection and analysis took place, resulting in the generation of statistical knowledge across numerous interconnected networks. This also highlights how the rise of opinion polling, which is and was largely done by private companies, raises issues regarding the implications of for-profit forms of data collection.

2.2.2.2   Numbers as Authoritative Facts

In addition to the observation of how numbers became a central point of reference, another important issue to bear in mind is the question of what there is to numbers that gave them such authoritative power. In her 1998 book *A History of the Modern Fact*, Mary Poovey sets out to ask "how knowledge was understood so that it seemed to consist of both apparently

noninterpretive (numerical) descriptions of particulars *and* systematic claims that were somehow derived from those particularized descriptions" (Poovey, 1998, p. XII italics in original). Poovey's work raises the question as to how, when and why the sense got established that "numbers are somehow *epistemologically* different from figurative language, that the former are somehow value-free whereas the excesses of the latter disqualify it from all but the most recreational or idealist knowledge-producing projects" (6). In doing so, she shows that the factual is a category that emerged historically by the early eighteenth century, when "numbers had acquired a set of connotations that would soon make them central to what counted as knowledge in numerous domains" (143). Her argument suggests that relying on the long-standing belief that simple and straightforward language is most effective in describing the world, was, supported by the esteemed status of mathematics, replaced by the view that numbers and mathematical language are presented as neutral and unbiased. The historical argument is that the development of double-entry bookkeeping "helped confer cultural authority on numbers" (Poovey, 1998, p. 54) and that those apparently noninterpretive numbers also "embody theoretical assumptions about what should be counted, how one should understand material reality, and how quantification contributes to systemic knowledge about the world" (Poovey, 1998, p. XII). Numbers have thus come to "epitomize the modern fact" (Poovey, 1998, p. XII) because they have both been seen as pre-interpretive and as the foundation of systematic knowledge.

Those views can be supplemented by looking at the work of Theodore M. Porter. In his 1986 book *The Rise of Statistical Thinking, 1820-1900,* Porter (1986) describes how in the nineteenth century, the possibilities for modern statistical innovations emerged and how statistics evolved through an exchange of ideas between natural and social scientists. Similar to Poovey, Theodore Porter reflects on how numbers became an authoritative and apparently objective means of representation. Porter describes quantification as a "technology of distance" which minimizes the need for personal trust and thereby entails more authority. The idea that a "highly disciplined

discourse helps to produce knowledge independent of the particular people who make it" (Porter, 1995, p. ix) became particularly "compelling to bureaucratic officials who lack the mandate of a popular election, or divine right" since decisions informed by numbers include "at least the appearance of being fair and impersonal" (Porter, 1995, p. 8). Porter describes quantification as a way to make decisions "without seeming to decide", and thus lending "authority to officials who have very little of their own" (Porter, 1995, p. 8).

2.2.2.3 Are Numbers Losing Authority?

While numbers have gained authority as apparently impartial and objective facts, their robustness becomes contested in situations of crisis. By this, I refer to the kind of situations described in the introduction to this dissertation, involving contradictory measurements of public sentiments. In those situations of crisis, both the authority of and trust towards numbers decline and the social mechanisms involved in their making become apparent. In order to understand those developments, I turn to recent scholarship, such as the work of William Davies (2018, 2020), who took up Poovey's work in a manner that bears particular importance for this thesis.

Statistics as an apparently neutral and apolitical instrument helps to settle arguments and offer points of reference, on which people from different backgrounds can agree. In recent years, however, some claim that statistics partly lost this ability, and "[r]ather than diffusing controversy and polarisation, it seems as if statistics are actually stoking them" (Davies, 2020, p. 84). Importantly, "[s]tatistics can only develop a positive enlightenment effect on the condition that their quality is trusted" (Radermacher, 2020, p. 2). This trust is dependent on giving answers to "[h]ow can we know that we know what we know (or do not know)", answers that "will protect statistics against inappropriate expectations and to address false criticism" (Radermacher, 2020, p. 2). If statistics and numbers no longer serve as a way of anchoring the representation of people and their opinions in something tangible, something stable, those very publics become fragile.

In an article in *The Guardian*, William Davies put it as such: "With the authority of statistics waning, and nothing stepping into the public sphere to replace it, people can live in whatever imagined community they feel most aligned to and willing to believe in" (Davies, 2017). Referring to Poovey, Davies (2020) argues that for numbers to maintain their authority, they need continuous policing "ideally through centralisation in the hands of statistics agencies or elite universities", adding that "the rise of commercial polling in the 1930s was already a challenge to the authority of 'facts' in this respect" (Davies, 2020, pp. 54–55). What we experience now, however, is a "politicisation of social sciences, metrics and policy administration" which means that "the 'facts' produced by official statistical agencies must now compete with other conflicting 'facts'" (Davies, 2020, p. 55). In the example presented at the beginning of this thesis, we have to do with conflicting facts within the same realm of production, which is public opinion polling. Those controversies mainly revolve around diverging conventions in the way numbers are produced, with, nevertheless, potentially strong effects on the question of which realities they bring about.

2.2.2.4 The Making of Numbers

The examination of the specific characteristics underlying the process of the making of numbers can be observed through the lens of scholars in the tradition of the economies of convention, a framework rooted in pragmatism. This approach seeks to explore the social practices surrounding quantification and measurement, drawing connections to other scholarly endeavours focused on the construction of facts. In 2008, Espeland and Stevens stated that "[q]uantification is a constitutive feature of modern science and social organisation, yet sociologists have generally been reluctant to investigate it as a sociological phenomenon in its own right" (Espeland and Stevens, 2008, p. 402). They suggest that the reason for this is that "like many scientists, we [the sociologists] have been more concerned with the accuracy of measures than with their social implications" (Espeland and Stevens, 2008, p. 402). In recent years, this

reluctance has changed and the field of the social sciences of quantification has become a prolific and more and more interrelated area of research.

The field of social sciences of quantification embodies a transdisciplinary approach that investigates the intricate processes of quantification, exhibiting the interconnectedness of various social and socio-economic phenomena. Consequently, one can refer to it as a comprehensive "political economy of statistics, quantification, and categorization" (Diaz-Bone and Didier, 2016, p. 7), recognizing the profound associations between these domains. This perspective towards numbers emphasises how they "are based on conventions" and "call into question their assumed obviousness by examining the exercises in quantification that produced them" (Bruno, Jany-Catrice and Touchelay, 2016, p. 1). Conventions are "logics of evaluation, valuation and interpretation" (Diaz-Bone, 2019, p. 116) and also underlie social science research methods, such as those employed by pollsters and survey researcher. Importantly, its aim is not to undermine but to study the possibility of quantification. The central conviction is thus that quantification is a social activity involving societal negotiations, aiming at illuminating the black boxes constituted by "indicators, categories, scoreboards and other accounting or statistical tools" (Bruno, Jany-Catrice and Touchelay, 2016, p. 1). To put it differently, it is about exploring the negotiations, agencies and mobilisations that play a role in the making of numbers.

The French Social Sciences of Quantification have one of its intellectual origins in the work of Alain Desrosières, whose book *The Politics of Large Numbers*, first published in French in 1993, drew inspiration from a variety of previous work, as well as STS and governmentality studies. Following this background, the discipline is, as put by Bruno, Jany-Catrice and Touchelay (2016, p. 3), is based on two postulates: First, data are not 'something given' (as the Latin origin suggests), but constructed in accordance with a variety of conventions that guide the process of quantification. In this sense, quantification is both *technical* and *social*, as it involves

measurement and agreement. Second, science and politics are not mutually exclusive spheres, but ought to be conceptualised and considered together.

One of the key aspects of Desrosières' work is his perspective of *realism* and *constructivism* towards statistical numbers. His work is "intended to avoid the recurrent dilemmas encountered by the people preparing the figures, if they wish to answer it fully. On the one hand, they will specify that the measurement *depends on conventions* concerning the definition of the object and the encoding procedures. But on the other hand, they will add that their measurement *reflects a reality*. The paradox is that although these two statements are incompatible, it is nonetheless impossible to give a different answer" (Desrosières, 1998, p. 12, italics in original). Instead of focusing on *objectivity*, Desrosières suggests focusing on *objectification,* through which this paradox can be viewed in a different light. He goes on stating that "[r]eality appears as the product of a series of material recordings: the more general the recordings – in other words, the more firmly established the conventions of equivalence on which they are founded, as a result of broader investment – the greater the reality of the product" (Desrosières, 1998, p. 12). This perspective goes along with, as previously mentioned, Latour's notion of constructivism as a "synonym for an increase in realism", and an "account for the solid objective reality by mobilizing various entities whose assemblage could fail" (Latour, 2005, pp. 91–92).

It is interesting, however, how much of what the Social Sciences of Quantification claim as their discovery, has already been formulated by, for instance, proponents of the Radical Statistics movement, such as in the introduction to the book *Demystifying Social Statistics* from 1979. In a contextualisation of the book, they write that "[l]ike other scientific practices statistics needs to be situated squarely in its social and historical context. Statistical practices are social in nature and their conceptual and technical instruments, orientation and uses all need to be seen in social terms" (Irvine, Miles and Evans, 1979, p. 3). Members of the radical statistics movement, have made the case for a social exploration of statistics early on. Irvine, Miles and Evans (1979) stated

that "[d]ata are therefore conceived of as social products: statistics are not *collected*, but *produced*; research results are not *findings*, but creations" and that "[b]ehind every statistician there is a constraining institutional and cultural framework; behind every technique and textbook on its use there are implicit ideological structures" (Irvine, Miles and Evans, 1979, pp. 3–4, italics in original). The Radical Statistics Group supported the publication of three books in the course of the last 40 years: *Demystifying Social Statistics* (Irvine, Miles and Evans, 1979), showing how social statistics are constructed and controlled in ways that serve the powerful, *Statistics in Society* (Dorling and Simpson, 1999), showing how social statistics are generally collected with a particular underlying purpose and *Data in Society* (Evans, Ruane and Southall, 2019)*,* considering the rise of Big Data and how data can be misleading and used for certain purposes*.*

There are several traditions, identified as "influential precursors which can be regarded as classical studies" (Diaz-Bone and Didier, 2016, p. 10) in the Social Sciences of Quantification. According to the same authors, Desrosières "mixed all together these four traditions of social studies of quantification" (Diaz-Bone and Didier, 2016, p. 11) and gave the field the shape it still has today. The first is the French tradition (Durkheim and Mauss, 2009; Bourdieu, 2010), considering social categorisation and social enumeration as "a social product of special importance, and […] (as) crucial objects of sociological inquiry" (Diaz-Bone and Didier, 2016, p. 10). The second is the American tradition, consisting of symbolic interactionism (Blumer, 1969; Becker and Horowitz, 1972) and ethnomethodology (Cicourel, 1964; Garfinkel, 1967). This tradition began "to make quantification an object of sociology, in a clear critical tone, aiming at questioning the monopole of the quantitative criteria of proof" (Diaz-Bone and Didier, 2016, p. 10). The third tradition is based in the history of science and discusses "whether the apparition of probability was itself a scientific revolution" (Diaz-Bone and Didier, 2016, p. 10) in a Kuhnian sense and culminated in the publication of two volumes entitled *The Probabilistic Revolution* (Kruger, Daston and Heidelberger, 1987) and the creation of the 'Bielefeld Group'.

The fourth tradition has been named the internalist tradition (Lazarsfeld, 1961; Duncan, 1984), which developed when statisticians, economists and quantitative sociologists "got themselves interested in the question of their own history" (Diaz-Bone and Didier, 2016, p. 10).

2.2.2.5   Social Studies of Survey Research and Polling

Another crucial field of research is the social studies of survey research and polling. Whilst much in this area developed in dialogue with work in the social studies of quantification, there is also a bulk of literature particularly dedicated to social histories of polling and survey research (Blondiaux, 1998; Keller, 2001; Igo, 2007), current controversies in the field (Lusinchi, 2016, 2017a) and historical narratives from proponents in the field of survey research and polling (Saris, 1998a; Kalton, 2019). Lusinchi, a retired independent social research consultant and instructor, for instance, explores current controversies, such as online surveys within the field of sampling and polling from a perspective informed by the sociology of science (Lusinchi, 2016, 2017a). Lusinchi (2012, 2017b, 2018, 2021) also investigated the history of survey research and polling, providing a sociological inquiry into the development of the field. In doing so, this line of research adds to already existing historical accounts of the field (Kruskal and Mosteller, 1980; Stigler, 1986; Bellhouse, 1988; Smith, 1990; Desrosières, 1998), including earlier classical sociological critiques of the practice of polling (Blumer, 1948; Bourdieu, 1972; Adorno, 2005). Of these mid to late-century 20th-century sociological treatments of polling, I draw in particular on Bourdieu's analysis of the social functions of opinion polls.

Another area of research that I would like to mention in this section are recent studies of polling hailing from political theory. In this sense, Mondon's (2022) work on the mainstreaming of reactionary politics and the insight that "rather than following 'what the people want', elite actors play an active part in shaping and constructing public opinion and legitimizing reactionary politics" (Mondon, 2022, p. 1), played an important role in my thinking on the role of polling and its symbolic value. Also relevant to my analysis is Yudin's research on the role of polls in

contemporary Russia (Yudin, 2020), and Ellwanger's work on the force that polling results exert on policy making (Ellwanger, 2017).

2.2.3   Concepts of Representation

When discussing processes of quantification, the concept of representation is inherently implicated in the discourse. Quantification is always the quantification of something in numerical form, thereby serving as a representation of reality through the medium of numbers. Throughout this section, I aim to briefly review central literature on the concept of representation, which will become central to the subsequent analysis. In doing so, I will pay a particular focus on two conceptualizations: representation as *response* and representation as *mobilisation*.

2.2.3.1   Representation as Response: The Classical View

The view of representation as response is mainly associated with the work of American political theorist Hannah Pitkin and her seminar work on *The Concept of Representation* (Pitkin, 1967). Pitkin characterises modern liberal democracies by a form of responsiveness, in the sense that the political leadership acts "in the interest of the represented, in a manner responsive to them" (Pitkin, 1967, p. 209). Thus, seminal work on the notion of representation understands it as an interaction in which the representant is responsive to the represented. Lisa Disch, an American political scientist, terms this norm, that political representatives must be driven by constituent demand, the "bedrock norm" (Disch, 2011). This can be understood as the "common-sense notion that representatives in a democratic regime should take citizen preferences as the 'bedrock for social choice'" (Disch, 2011, p. 100). To Pitkin, this responsiveness involves that the represented must be logically prior, that the representant responds to the represented and not the other way around.

Pitkin (1967) presents four views of representation, which are formalistic, symbolic, descriptive and substantive representation. Whereas the first is about the institutional settings in which representation takes place, the second is concerned with the meaning that a representative has for the represented, the third is about the way a representative resembles the represented, and the fourth is about what the representatives do for the represented. Pitkin defines the concept of representation following its etymological origin as *re-presentation* in the sense of "making present again". Pitkin further differentiates between views of representation as *acting for* and as *standing for.* In the case of the former, representation is either concerned with *authorisation*, when the representative is authorised to act or with *accountability*, when the representative is held accountable for their actions. Distinct from this view, representation as *standing for* does not involve action at all but is concerned with the representatives' characteristics; it is about being, rather than about doing.

2.2.3.2 Representation as Mobilization: Inverting the Classical View

The view of representation as mobilization is best exemplified in Disch's work, for instance in her recent book *Making Constituencies: Representation as Mobilization in Mass Democracy* (Disch, 2021). Disch develops her understanding of representation in dialogue with and in opposition to Pitkin's view that the representation of the interest of the constituencies is the starting point and the representation of those interests by representatives its endpoint. Disch's view is motivated by the observation and everyday experience that representatives are not as responsive as they ought to be to particular groups of society. The notion of *representation as mobilisation* thus interrogates this view in stating that "in mass democracies, acts of political representation often do not take constituencies and their interests as a starting point: they begin by making constituencies and specifying their demands" (Disch, 2021, p. 1).

A good starting point to explore the consequences of this view in relation to polls and surveys is the notion of 'public opinion as a dependent variable' (Margolis and Mauser, 1989). Instead

of viewing polls as a reflection of an independent public opinion, the focus is to "start with elites and presume that they try to manipulate public opinion through the mass media and by other means" (Margolis and Mauser, 1989, p. 87). Similarly, Maza and Cook (2002) mention that the "one-way flow of influence from public opinion to policy", which corresponds to the classic view, as proposed by Pitkin, implicitly ignores "the possibility that politicians themselves may influence the policy preferences of citizens" (Manza and Cook, 2002, p. 639; see also: Jacobs and Shapiro, 2000).

In her introduction to the edited volume on *The Constructivist Turn in Political Representation* (Disch, Sande and Urbinati, 2019), Disch counts several intellectual traditions and scholarship among the sources of this new paradigm, which I will briefly summarise in the following. The constructivist turn goes back to de Saussure's (2011) semiotics and the suggestion that language creates meaning differentially and not referentially. Sense does not stem from a correspondence between words and things, but from the opposition and linkages among terms. Signs and reality are in an arbitrary relation to reality, which is symbolically mediated. A subject is thus not merely a receiver of messages but plays an active role in constituting the meaning of the message. In light of those considerations, Disch (2019) points attention to empirical findings on public opinion creation, stating that individuals form their preferences in response to the communication of political leaderships and campaigns. This "turns the classic model of democratic legitimacy on its head" (Disch, 2019, p. 6), as it shows that citizens respond to politicians and not the other way around. A further source in the intellectual trajectory is to be found in the work of Berger and Luckmann (1991), emphasising the "mutually constitutive relationship between the perceptions and material practices of human subjects and the 'objective facticity' of the social world", offering "another way for constructivist approaches to arrive at understanding political representation as a mutually constitutive process that defines representative and represented alike" (Disch, 2019, pp. 6–7). This goes in line with the work of

Saward (2006) and his analysis of the 'representative claim'. Representative claims do not speak for and thus represent a constituency, but "solicits a constituency to recognise itself in a claim and to support the person who made it" (Disch, 2019, p. 9).

2.2.4   Ignorance Studies

Lastly, an area of research that played an important role in the development of this thesis, especially in the beginning, was the field of ignorance studies. My original plan for this dissertation was to study statistics as a way to rationalize the unknown, a notion that has also been repeatedly mentioned by many of my interviewees. Rather than studying the making of knowledge, the field of ignorance studies explores the social conditions underlying the construction and distribution of ignorance and developed to become a thriving cross-disciplinary field. An early articulation of this line of research was the introduction of the term "agnotology" (Proctor and Schiebinger, 2008) by which they describe a reversal of the notion of epistemology. Different to epistemological questions which aim to understand the conditions of knowledge, those questions aim to understand the conditions of non-knowledge and ignorance. The development and variety of the field can be viewed in the publication of two international handbooks on ignorance studies (Gross and McGoey, 2015, 2022). Important for the development of this dissertation was especially McGoey's book, *The Unknowers* (McGoey, 2019), which informed my thinking about the role of pollsters and the monopolization of oracular power. In 4.3.4, I contend that pollsters have contributed to a distinct interpretation of public opinion, effectively establishing themselves as the exclusive interpreters and spokespersons of public opinion.

Scholars in the field of ignorance studies have also drawn connections to work presented earlier in this literature review. Ravetz (2015), for instance, shows that "[a]lthough Thomas Kuhn did not discuss ignorance, his picture of 'normal' scientific research, and by extension scientific education, is one of the systematic inculcation of ignorance-of-ignorance" (Ravetz, 2015, p. 58).

This means that paradigms constitute the principles and structures in which science and thought can take place, rendering things outside the paradigm unknowable.

2.3   Methodology

2.3.1   Methods - Overview

Methodologically, this research follows a two-pronged approach involving secondary analysis of literature on the history of statistical thought, especially on the history of survey research and polling, as well as the empirical research based on 20 semi-structured qualitative interviews. The study also included some archival research, studying the protocols of the International Statistics Institute (ISI), which are digitally accessible through their website. This archival material, dating from 1895, when Anders Kiaer, who played a central role in the development of sampling, helped me to understand the process of sampling to become an accepted approach. I also draw on media reporting, mainly used as a way to illustrate recent developments from the 1990s onwards, which my informants have talked about in the interviews.

These interviews enabled me to access the field of survey research and polling and to get hold of the inner dynamics and positions within various controversies over polling methods and results. The interviews were based on semi-structured interview guides, which to some extent contained similar questions and topics, but some of which were particularly tailored to the interviewees. The interviews usually started with a general question on the biography of the interviewees, followed by a question as to how they would define statistics and survey research. By the end of the interview, I would always ask if they had anything to add that we did not discuss throughout the interview and whether they would recommend other potential interviewees to me. Most interviewees were recruited through this snowball sampling technique.

The analysis of interview data follows the general tenets of grounded theory, in remaining open to allow the field to express itself in the form of the interview material. In this sense, a first

cycle of analysis was done completely inductively, without the aim of finding predefined patterns in the data. Throughout the course of the data analysis, several general themes emerged, which became the main themes through which I structured the presentation of the empirical material. In a second round, which, in terms of periodisation, happened after an intense study of the literature, the analysis of interview material gained a stronger deductive nature. In this sense, including historical literature in the grounded theory analysis means including data "which were originally collected for other purposes" (Glaser, 1963: 11) and allows for induction, deduction, and verification and thus includes the essential "three aspects of inquiry" (Strauss, 1987: 12).

2.3.2   First Encounter with the Field

My interviews grew largely out of tracking the developments that opened this thesis: a growing controversy in Germany over the rise of Civey. I first learned of the controversy through an article by Robert Pausch and Fritz Zimmermann (Pausch and Zimmermann, 2020), investigating conflicts between Forsa and Civey, published in the German weekly newspaper DIE ZEIT. From a statistics standpoint, the controversy was about probability vs. non-probability sampling, a discussion that is strongly related to questions around randomisation and how to choose participants. This controversy fascinated me not primarily because of its scientific dimensions, but because it displayed a variety of aspects that a sociological investigation into statistical fact-making could be interested in: Proponents of both companies mobilised highly interesting issues such as technological changes affecting the way surveys and polls are to be conducted, changing societal patterns and preferences, as well as path-dependencies when it comes to the choice of statistical methodologies. In addition, this controversy displays strong emotional attachments and particularly harsh language, as well as disingenuous practices, which one usually does not expect to come across in scientific settings.

For instance, one actor in this controversy was a Twitter account called @civey_watch. This account was opened in April 2018 and its description states: "Civey is a junk institute. Civey

surveys are not representative. We keep an eye on it. Whistleblowers welcome." The first sentence of the bio later changed to "Civey is the imitation of a survey institute". The identity behind this Twitter account is unknown, but the goal to discredit Civey is clear. On the 26th of November 2018, @civey_watch published a screenshot of Civey's internal input mask, claiming that Civey determines its results prior to the survey. The screenshot is claimed to have been leaked by a former Civey intern. As Pausch and Zimmermann (2020) highlight, Janina Mütze, the co-founder and CEO of Civey, has dismissed the accusations made by @civey_watch as baseless and has emphasized that Civey regularly grants access to scientists for scrutiny of their methodologies. Civey has initiated an investigation to identify the source of the leaked screenshots, suspecting an intern who had previously worked in the Forsa call center as the sole possible culprit. Consequently, the intern was terminated from their position and a legal case was filed against them. Although the prosecution supported this assertion, they decided not to pursue the lawsuit due to insufficient public interest.

2.3.3   Recruitment: Surveying the Surveyors

After reviewing a variety of media coverage, scientific publications and other publicly available material on this controversy, a core set of actors emerged, which were central to the controversy in Germany. Even more, there appeared to be well-known antagonists, who are also perceived as such in the field and who regularly confronted each other and did not seem to have much understanding and patience for the other side. This is particularly interesting, since the field is quite small, which means that it can be assumed that most potential interviewees know each other and have a clear idea of the other's standpoints.

In this sense, I contacted potential interviewees from three different areas: 1. Researchers advocating for either of the two research programmes (probability or non-probability sampling), 2. researchers investigating the advantages and disadvantages of either of those research programmes and 3. representatives of the two companies in question. Since this took place at the

height of the COVID-19 pandemic and face-to-face interviews were not permitted and could have been a health risk, all interviews had to take place via Zoom, a situation that made it easier to reach people, but which also came with limitations. One interview took place via landline phone. I discussed the identification of informants with my co-supervisor, Professor Nick Allum, an expert in quantitative methods and survey design, who also advised me on potential interviewees and introduced me to some of them.

To find interviewees from the mentioned areas, I first contacted Prof. Dr. Ulrich Rendtel, a retired statistics professor at the Free University of Berlin and Prof. Dr. Rainer Schnell, a professor of empirical social research at the University of Duisburg-Essen. Both have publicly defended either of those approaches and are known as antagonists in the field. This has not only become visible throughout the interviews I conducted with them, in which they talked about each other, highlighting their conflictual relationship; other participants in the field also emphasised this in my interviews with them. After mentioning Rendtel in my interview with Schnell, he laughed and said he did so, "because there is practically no point where Rendtel and I agree" (Schnell, Interview, Translation). Since Rendtel serves as an adviser for Civey, Schnell also made clear to me that he does not take his evaluations of Civey seriously: "Mr Rendtel has personal and financial ties to Civey. And how can I say, you must not ask the frogs if you want to drain the swamp." (Schnell, Interview, Translation) Rendtel said about Schnell that he has a "decidedly different opinion on the matter" (Rendtel, Interview, Translation). In interviews with other participants in the field, this antagonism between the two has also been raised: "I am prepared to pay admission when Uli Rendtel and Rainer Schnell discuss" (Münnich, Interview, Translation).

The general culture of how this controversy unfolded in Germany has also been vividly described by another interviewee, Prof. Dr. Sabine Zinn: "And then there was a discussion between Civey, Ms Mütze [Founder and CEO of Civey], Rainer Schnell and one […] from Forsa

[...]. And the only thing missing was someone selling popcorn. I sat inside at the end and thought that the most professional one was the youngest, namely Ms. Mütze, who argued everything most professionally and also asked questions, while the others just mocked around and I thought that was really bad. I was a bit ashamed because that's not a culture we should have as scientists." (Zinn, Interview, Translation)

I then carried out an interview with Dr. Carina Cornesse, who was part of GESIS, University of Mannheim when doing the interview and who is now at the German Institute for Economic Research (DIW Berlin). She is the main author of a recent review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research (Cornesse *et al.*, 2020), which puts her in a great position as an interviewee. The mentioned study will be also of great importance for this dissertation. The interview with her greatly helped me to map the different positions and arguments and to be able to see certain positions from their positionality in the field. Carina Cornesse also pointed me towards some of the relevant figures in the US, which greatly helped me to gain a broader perspective and to recruit more interviewees from outside of Germany. I have also stayed in contact with her after the interview and participated in a summer school that she taught on the very topic of probability and nonprobability sampling. We also met at a workshop and a conference in the years after the interview.

I then contacted proponents of the two companies in Germany, Civey and Forsa. I was very quickly able to get interviews with the managing director Thorsten Thierhoff and the founder and director Manfred Güllner. Both were very approachable and told me about the history of the company, their perspectives on the current state of polling and their competitor Civey. Unfortunately, I was unable to obtain an interview with Civey. I sent out a vast number of E-Mails and LinkedIn messages, not only to the founders and directors, but also to other general addresses, but I never received a response. I received one response from a person from within the inner circle who was very interested at first and asked if I could send them my questions in

advance. I thus sent the general topics and issues I was interested in but did not hear from the person anymore after this. I assumed that Civey decided not to give me an interview, so I eventually stopped reaching out to them. Instead, I was able to obtain interviews with people close or with relations to Civey, such as the already mentioned Prof. Dr. Ulrich Rendtel and Prof. Dr. Sabine Zinn. Another interviewee in this context was Prof. Dr. Raimund Wildner, vice-president of the Nuremberg Institute for Market Decisions (previously GfK Verein) and honorary professor at the University of Erlangen-Nürnberg. He is chairman of the Council of German Market and Social Research (Rats der Deutschen Markt- und Sozialforschung) whose goal is preserving the reputation of market and social research. He is also a member of ESOMAR's Professional Standards Committee. ESOMAR stands for its original name, European Society for Opinion and Marketing Research and is an important organization for market, social, opinion, and data analytics professionals and researchers.

To access the US context, I followed a similar strategy. I also aimed to interview known antagonists in the field and proponents of companies with different positions towards the controversy in question. I thus contacted Dr. Drew Linzer, survey scientist and statistician, and the Director of Civiqs, an online polling company based in Oakland, California. I reached out to Linzer after he gave a keynote at a summer school at the University of Essex, where I participated. In the interview, we not only spoke about the survey and polling field in general and about how there is a strong need to develop new methods besides traditional probability sampling, but also about some of the more sociological critiques towards polling. Another person I contacted and who also directly agreed on giving me an interview was Gary Langer, a long-time director of polling at ABC News and president of Langer Research Associates. Langer is a known proponent of probability sampling approaches and thus to some extent distinct from Linzer.

At the same time, I reached out to Professor Douglas Rivers, a professor at Stanford University and CEO of YouGov and Professor Jon Krosnick, also a professor at Stanford University. The contact to Krosnick was established through Gary Langer. Throughout the interviews with Krosnick and Rivers, it quickly became clear that both stand in a similar relationship, both personally and professionally, like Schnell and Rendtel in Germany. The fact that many people advised me to talk to both people showed how important they are as opinion leaders in the field. Due to the importance of both for the developments in the field of survey research and polling, the interviews with Rivers and Krosnick also served as a way to establish the recent developments in the field itself. I also stayed in contact with Gary Langer, Jon Krosnick and Douglas Rivers, to discuss some follow-up questions via E-Mail and Jon Krosnick participated in a panel I co-organised, together with Nick Allum, at the conference of the European Survey Research Association (ESRA) in July 2023 on the philosophy and history of survey research.

Through an introduction by my co-supervisor, Prof. Nick Allum, I also conducted an interview with Prof. Dr. Willem Saris, a retired professor at the University of Amsterdam and the Pompeu Fabra University in Barcelona. Saris developed the Telepanel, one of the main precursors to online sampling and online methodologies, which is why his perspective was of particular importance for my research.

In addition to the interviews conducted with individuals involved in the particular controversy, I also aimed at getting interviews from general statisticians, who can respond to general questions about the field and statistical fact-making. Therefore, I was looking for renowned statisticians who are able to speak, to some extent, for their field of expertise, that is, statistics. Therefore, I interviewed Professor David Spiegelhalter, board member of the UK Statistics Authority and statistics professor at the University of Cambridge, Prof. Dr. Ralph Münnich, statistics professor at the University of Trier and chairman of the German Statistics

Society (Deutsche Statistische Gesellschaft), as well as Prof. Dr. Gerd Bosbach, statistics professor at the Koblenz University of Applied Sciences. I conducted one further interview with a statistician, who, however, asked me to remain anonymous.

In the course of my data gathering, I also started to explore the role of randomisation in the case of estimating COVID-19 case numbers. This took place when I spent some time as a visiting researcher under Prof. Dr. Matthias Gross at the Helmholtz Centre for Environmental Research (UFZ) in Leipzig, Germany. By this time, there was a strong discussion about the reliability of case numbers, as they were calculated from positive tests taken at test stations and hospitals. Since those numbers necessarily did not represent the whole population, as they did not estimate cases of non-tested individuals, wastewater-based estimations were discussed as a more reliable sampling method, since the wastewater was believed to constitute a random sample of whole populations within certain catchment areas. Wastewater was thus seen as a remedy to the undercoverage inherent in other, more standard means of measurement. This discussion was often being brought up in relation to the controversy over probability and non-probability sampling in survey research, which is why I started exploring it. Therefore, I conducted interviews with researchers at the UFZ and affiliated institutes. I thus conducted four interviews with natural scientists who were working on new methods and models to estimate COVID-19 case numbers, which informed my thinking about sampling and the difference between design-based and model-based approaches. In this context, I conducted interviews with Dr. René Kallies, Dr. Lennart Schüler and Prof. Dr. Hauke Harms from the UFZ and Björn Helm from the Dresden University of Technology, who, however, collaborated with the scientists at the UFZ. In the course of this research, I also visited a wastewater treatment plant in Langenreichenbach, where I could, besides others, see the sampling and probing processes and was given a tour by its director Ted Linke. In the end, I, however, decided to use those insights and interview material rather as background information, since there were many interesting statements about this matter

involved. I did not, however, explore the case of Covid-19 wastewater-based epidemiology in its own right and decided to potentially investigate it in a separate side project.

In total, my thesis drew on 20 interviews, summarized in the following table:

	Name	Affiliation	Date	Duration
1	Ulrich Rendtel	Free University of Berlin, Retired	02.10.2020	1h 45m
2	Carina Cornesse	GESIS, now German Institute for Economic Research (DIW Berlin)	23.10.2020	1h 18m
3	Rainer Schnell	University of Duisburg-Essen	27.10.2020	1h 01m
4	David Spiegelhalter	University of Cambridge	12.11.2020	0h 37m
5	Drew Linzer	Civiqs	07.12.2020	0h 47m
6	Willem Saris	University of Amsterdam; Pompeu Fabra University, Retired	14.12.2020	1h 04m
7	Gary Langer	Langer Research Associates	11.01.2021	1h 01m
8	Jon Krosnick	Stanford University	26./27.01.2021	2h 18m
9	Douglas Rivers	YouGov; Stanford University	03.02.2021	0h 32m
10	Thorsten Thierhoff	Forsa	12.02.2021	0h 56m
11	Gerd Bosbach	Koblenz University of Applied Sciences, Retired	02.03.2021	1h 11m
12	René Kallies	Helmholtz Centre for Environmental Research	04.03.2021	0h 39m
13	Lennart Schüler	Helmholtz Centre for Environmental Research	17.03.2021	0h 42m
14	Hauke Harms	Helmholtz Centre for Environmental Research	23.03.2021	0h 40m
15	Björn Helm	Dresden University of Technology	31.03.2021	0h 57m

16	Sabine Zinn	German Institute for Economic Research (DIW Berlin)	23.04.2021	1h 12m
17	Anonym	-	11.05.2021	0h 55m
18	Raimund Wildner	University of Erlangen-Nuremberg, Council of German Market and Social Research; ESOMAR; NIM	17.05.2021	0h 57m
19	Manfred Güllner	Forsa	17.05.2021	0h 54m
20	Ralf Münnich	University of Trier; Chairman of the German Statistical Society (DStatG)	26.05.2021	1h 31m

All interviewees were given a participant information sheet and were requested to sign a consent form, indicating that I would be allowed to use their statements for the purpose of this dissertation, for academic publications and conferences, as well as that I would be allowed to use their real names and affiliations. While almost all interviewees signed the consent form, there were three cases where I received only recorded oral and written consent via E-Mail. In one of those, the interviewee stated the following to me via e-Mail: "The consent form is silly. You're free to quote me or use anything I tell you, so long as you attribute anything original to me. I am not participating in a project; I don't care where you store your notes or what you use it for now or in the future. When did university bureaucrats stop understanding how scholarship works and introduce this kind of nonsense?" In three cases, the interviewees asked me to send them the used statements for approval. All my interviews took place between October 2020 and May 2021 and lasted between 30 minutes and 2.20 hours. After conducting the interviews, I transcribed them and implemented them to MAXQDA for subsequent analysis. The Interviews were conducted in English and German. In the case of the interviews conducted in German, I provided English translations for the excerpts presented in this dissertation.

Besides the 20 main recorded interviews themselves, I had several informative background interviews with scientists at the UFZ, researchers working on the history of sampling and statistics, as well as with survey researchers, which helped me to discuss my findings and ideas and to explore new possible paths to investigate my dissertation topic.

2.3.4 Positionality

I started my PhD in January 2020 and moved to Colchester at the end of February 2020. The onset of the COVID pandemic was both a methodological advantage and a disadvantage. It made it easier to reach potential interviewees and it seemed as if they were very willing and interested to use Zoom, especially since everybody had just adjusted and got used to online meetings. On the other hand, the online environment also potentially made it more difficult to build a relationship with some interviewees and to discuss certain matters in depth.

My own academic background is in philosophy and sociology, focused on qualitative analysis. To ensure sufficient statistical knowledge and training, I completed intermediate-level classes in quantitative social research during the first two years of my PhD, mainly in the form of summer schools at the Essex Summer School in Social Science Data Analysis and at GESIS – Leibniz Institute for the Social Sciences, as well as through self-study. Furthermore, being a native German language speaker proved important for approaching German-speaking participants.

2.4 Epistemological Approach

The methodological principles of Actor-Network Theory (ANT) have been guiding throughout this dissertation. Actor-Network Theory, whilst being denoted a theory, can rather be conceived of as a methodological framework for conducting social research in diverse and complex settings. ANT aims to transcend conventional dualistic distinctions, such as subject/object or

nature/culture and instead focuses on assemblages of heterogeneous actors into actor-networks, in which practices and associations achieve stability.

The primary interest of ANT lies in understanding how networks transform themselves through the enrolment or dissolution of actors. A crucial term within this framework is "actant," which Latour employs to acknowledge the role of non-humans in shaping scientific theories. ANT employs the notion of actants as a means to methodologically treat humans and non-humans on equal terms. An actant is defined as "something that acts or to which activity is granted by others. It implies no special motivation of human individual actors, nor of humans in general. An actant can literally be anything provided it is granted to be the source of an action" (Latour, 1996, p. 373). The identity of actants is shaped and influenced by these transformations and the relationships between various actors in the network. Their differences are not predetermined but arise as effects of actor-networks. Only when such networks solidify and attain stability, do they become actors themselves.

Throughout this dissertation, the notions of stability and durability, as conceived of in Actor-Network Theory, are crucial: The first two chapters, for instance, explore the early history of survey research, polling and election forecasting, in which the establishment and dissolution of predictive networks can be observed. While, for instance, the magazine Literary Digest was among the first and successful actors when it came to the prediction of election forecasts, they mispredicted the outcome of the 1936 US election, contributing to the waning circulation and demise of the magazine. While the magazine's election predictions were not based on a stringent methodology, they were highly successful in that they were able to establish a stable network. At around the same time as the Digest's failures, Georg Gallup was able to successfully predict the outcome by drawing on quota sampling, a sampling strategy that was theoretically deemed to be inferior to random sampling. Similar to the Digest, Gallup was also able to establish his predictions within stable networks, up to the point where he also mispredicted the outcome of

the 1948 US presidential elections. At this point, the network started to disintegrate and new mobilizations and enrolments took place with the result of establishing a stable culture of polling, involving random sampling, as a methodology that was able to hold the various stakeholders, publics and more together. In more recent developments, we can see similar processes. For instance, with the decline of landline phones, traditional sampling techniques lost some of their efficacy as participation rates declined, thereby destabilizing and necessitating numerous reconfigurations within the polling apparatus. Following ANT's perspective that establishing scientific facts involves "placing these actors in a stable network" (Detel, 2001, p. 14265), I show how the attainment of stability is pivotal for successfully establishing a methodological apparatus in the realm of (social) science research methodology.

Another central aspect that guides this dissertation is the notion of boundary work, a term coined by Thomas F. Gieryn, who asks how science could gain "intellectual" (Gieryn, 1983, p. 783), "cognitive" (Gieryn, 1983, p. 784) or "epistemic" (Gieryn, 1999, p. 1) authority, thus how science could position itself as the sole producer of truth and become successful in "credibility contests" (Gieryn, 1999, p. 1). Distinct to essentialist understanding of science, Gieryn's account is a constructivist account, in that it looks at the "attributions of selected characteristics to the institution of science for purposes of constructing a social boundary that distinguishes 'non-scientific' intellectual or professional activities" (Gieryn, 1983, p. 791). While Gieryn's concept of boundary work was aimed at the social construction of a boundary between science and non-science, this was further developed by other thinkers, such as Jasanoff, who defined it in a broader sense as "a communally approved drawing of lines between 'good' and 'bad' work (and, not trivially, between good and bad workers) within a single discipline, between different disciplines, and between 'science' and other forms of authoritative knowledge" (Jasanoff, 1995, p. 53). According to Jasanoff, boundary work thus describes not only the construction and maintenance of the distinction between science and non-science, but also disputes over authority

and recognition within a scientific community.[6] Adding yet another dimension to the social practice of boundary making, recent work in the field of ignorance studies adds notions of power, as to who are the people and groups endowed with the power to define where those boundaries lie. This relates to the notion of "oracular power", as coined by sociologist Linsey McGoey, or the shaping of a "social consensus about where the boundary between ignorance and knowledge lies" (McGoey, 2019, p. 61). Oracular power is about the institutions and individuals viewed as the most trusted authorities, who monopolise this very capacity of delineating what counts as legitimate knowledge. To McGoey, those are especially found among natural and social scientists, people in financial markets and religious figureheads. Throughout this dissertation, I will discuss pollsters as a subcategory of social scientists, endowed with the power to define the truth about public opinion. In adapting a certain technique, the random sample, and establishing a particular view of what public opinion is, pollsters created a monopoly over the authority to speak in the name of publics and their opinions. They successfully established a boundary that delineates who is best placed to speak in the name of public opinion.

---

[6]   Gieryn also mentioned the use of boundary work to create boundaries within a discipline. Gieryn notes that "the utility of boundary work is not limited to demarcations of science from non-science. The same rhetorical style is no doubt useful for ideological demarcation of disciplines, specialties, or theoretical orientations within science" (Gieryn, 1983, p. 792); he, however, does not develop this insight or apply it much analytically (see also: Swedlow, 2007).

CHAPTER 3

3   Between Consensus and Controversy

3.1   Introduction: The Bifurcation of a Paradigm

If scientific truth develops through consensuses and controversies within a scientific community, the history of survey research and sampling is no exception. The primary objective of this chapter is to bring to the fore the historical developments that have occurred within the realm of survey research and polling, introducing the epistemic context in which those activities take place today.

Drawing on the concept of 'boundary work', as developed by Gieryn (1983) and Kuhn's idea of paradigms (1996), as well as ideas from of Fleck (1981) and Lakatos (1978), this and the following chapter explore and delineate two different historical threads: the emergence of the concepts of "population" and "sample" as features of reality, and the emergence of a particular concept of public opinion as a result of these developments. The first thread characterises the idea of a sample as a way to conceive of the whole, a countermovement to the primacy of full enumeration that marked the epistemic context of the 19th century. The second thread follows up on this development and looks into the emergence of public opinion as a result of this very movement. Public opinion polling was not only crucially influenced by developments in the field of sampling, but also played a crucial role in stabilising and questioning sampling paradigms itself. In exploring those developments, those two chapters serve as a historical introduction to the topic of sampling in survey research and polling, analysing the changing landscape of statistical observation and data collection as a history of consensuses and controversies.

The core aim of this chapter is to 1) introduce my theoretical framing in more detail, and 2) offer a historical analysis of major developments in 19th and early 20th century survey sampling. To do so, this chapter summarizes the ideas of a selection of key thinkers central to the

development of sampling, and then turns to a re-examination of historical literature on key shifts in statistical thought, paying particular attention to the period of 1895-1934. I describe the invention and development of sampling as an accepted technique, as well as the establishment of probability sampling as the superior approach. In a nutshell, the following analysis can be periodised as follows: While during the 17th and 18th centuries, early attempts to estimate population characteristics based on a sample were made, 19th-century statisticians generally rejected those attempts and instead, prioritised full enumeration. The end of the 19th and early 20th century, however, marks a time of crisis for the dominant paradigm of full enumeration in that sampling was seriously proposed again and eventually developed to become the consensus among the scientific community. The bifurcation of the field into proponents of probability and non-probability sampling already emerged during this period.

3.2   Sampling as an Everyday Practice: Introducing Central Terms

To add some more explanatory detail, it is important to note that sampling does not only apply to scientific methods but is also a daily and intuitive activity. If I want to make sure my soup is seasoned well, I do not need to try the whole soup, it is enough to take a spoonful, randomly chosen, with which I can judge the whole pot. This is intuitive because the soup, taken that it is stirred well, is homogenous in its characteristics – a small part can stand for the whole. Similar statements can be made regarding other substances, such as blood or urine samples. A common phrase in the field is that "if you don't believe in random sampling, go to the doctor for a blood test, have him take it all", a sentence attributed to Arthur Nielsen Jr., the son of Arthur Nielsen, founder of the A.C. Nielsen Company, known for developing and monitoring the Nielsen rating system, which focuses on television viewership. While sampling seems intuitive when it comes to trying a soup or giving a blood sample, other situations turn out to be more difficult. Trying a salad, for instance, might turn out to be a lot trickier. If I want to try a salad before serving, I cannot just randomly select some ingredients, since it might just miss important ingredients and

does not form a miniature of that salad. While the soup constitutes, metaphorically speaking, a homogenous population, the salad does not, which is why one would need an even bigger salad and stratify, meaning dividing it, into different levels, such as lettuce, cucumber, olives, and randomly select from those.

These seemingly odd examples lead us to two central terms that will be crucial for understanding the subsequent sections: *Complete enumeration* or *census* and *partial enumeration* or *sampling*, designating the approach to count everything, a task that can be very expensive and time-consuming, or only a part. The second distinction is between *purposive* and *random sampling*, designating the approach to choose a subset of a population purposively, e.g., according to particular characteristics, or by chance. The examples of the soup and the salad can help to illuminate the difference: We don't need to eat the whole soup or the whole salad to assess its taste; it is sufficient to try a part of it. In more contemporary discussions in survey research, the notions of purposive and random sampling are usually replaced by the notions of non-probability and probability sampling. While the meaning remains largely the same, the level of sophistication has increased greatly over the decades. According to Bethlehem (2018, p. 4), one of the earliest formulations of this principle is from Miguel de Cervantes' Don Quixote, constituting an early articulation of the principle of sampling: "By a small sample we may judge the whole piece".

The origin of the notion of sampling in other languages makes its meaning more obvious than the English term. The German word "Stichprobe" for example, is composed of "Stich", which means to dig or to stab and "Probe", which translates to try or to test. As Bethlehem (2009, p. 6) states, the term "Stichprobe" was first mentioned in 1583 as a mining technique, where a small amount from a melted substance is used to determine the amount of metal contained in it.

Turning to the notion of sampling in statistics, we can see that sampling is not as old and has not been taken up with the same intuitiveness as was the case in other areas of life. However,

while its use as a method faced significant challenges at the beginning, it is now generally accepted in contemporary research. Importantly, controversies surrounding sampling persist and echo those encountered in the history of the field.

3.3 From Incomplete to Complete Observation: The Early History of Sampling

3.3.1 The Birth of Sampling in the 17th and 18th Century

Survey methodology was already known before the 19th century, but it did not rise to become an accepted approach in statistics and its validity has been contested throughout history. Early attempts to infer from a part to a population already existed in the Middle Ages, shown, for instance, by Droesbeke, Fichet and Tassi (1987). Accordingly, the most famous one took place in France, known under the name "état des paroisses et des feux des baillages et sénéchaussées de France" drawn up in 1328. The notion of 'fire' [feux] refers to a household, a family or a dwelling and was an essential element in estimating a population (Droesbeke, Fichet and Tassi, 1987, p. 4). Classical studies on the history of statistics usually draw upon the work of John Graunt (1620-1674) and William Petty (1620-1683), as well as the works of Pierre Simon Laplace (1749-1827) as early attempts to sampling. Their early endeavours in sampling constituted significant efforts towards formulating statistical assertions concerning a population based on a subset. In the context of the present chapter, their works hold particular significance due to their implications for comprehending how statisticians of the 19th century perceived the methodologies employed during those times. Exploring the reasons behind this era being regarded as a period of "speculation" provides valuable insights into the historical backdrop that facilitated the emergence of sampling as an accepted scientific method.

Presented to the Royal Society in 1662, the English merchant John Graunt (1977) described a method with which to estimate the population of London based on partial information. In his *Natural and Political Observations Made upon the Bills of Mortality*, Graunt examined a sample

of parishes with well-kept, yet problematic, registers, he found out that there were three burials per year among 11 families. Under the assumption that this ratio was similar across all districts and considering the total annual number of burials, which was around 13,000, Graunt calculated that there were around 48.000 families living in London. Multiplied by an average family size of eight, he estimated the total population of London to be 384,000. As Bethlehem (2009, p. 6) adds, despite recognizing that averages, such as the number of burials per family, fluctuated across location and time, John Graunt did not account for these variations in his method. Furthermore, since his approach lacked a sound scientific foundation, he was unable to comment on the precision of his approach. According to Bethlehem, there are two findings Graunt made, which will become of crucial importance to survey sampling: 1) He observed that some social and demographic indicators, such as the proportion of born boys and girls, remains stable over time and space and 2) he used averages to calculate total values.

It took more than a century until another survey similar to Graunt's was conducted. In 1785, Pierre Simon Laplace presented a method to the Academy of Sciences in France to estimate the population of the country (Laplace, 1786). He suggested extrapolating this number from birth registers from a sample of 30 departments over the area of France. He selected those departments based on two criteria: first all types of climates ought to be represented, second the respective communes ought to be able to provide accurate information. Laplace aimed to calculate the ratio of the number of inhabitants to the number of births in those sample regions, then multiplying it by the number of births in the whole country, which could precisely be obtained for the whole population. These examples of early forms of sampling already hinted at much of what was seen as the mode of normal science. Throughout the 19th century, Laplace' and Graunt's work was almost completely forgotten and full enumeration became the dominant way to understanding population characteristics. As we shall see now, a crucial figure in this development was the Belgian statistician Adolphe Quételet.

3.3.2 The Paradigm of Full Enumeration

Whereas there have been suggestions and attempts to gain knowledge about a population through sampling, the paradigm of normal science that developed throughout the 19th century was one of full enumeration. A crucial figure when it comes to understanding this development was Adolphe Quételet (1796 – 1874), who primarily pursued his career in Belgium as an astronomer and meteorologist at the Royal Observatory in Brussels, yet he gained global recognition for his roles as a statistician and sociologist, notably in social physics, the application of methods inspired from physics to social phenomena. He is also known for his concept of the 'average man', a statistical representation of the physical and moral characteristics of a population.

Quételet knew Laplace and his work and was firstly impressed by his suggestion of using partial data to make statements about a population. He, made, however, as Stigler points out, an abrupt turnaround, apparently due to an argument made in a note by Baron de Keverberg, which Quetelet should later add to his memoir. Baron de Keverberg, a senior state administer, states that he knew about Laplace's methods through Quételet, which consists in his words, of "a precise census of population, but only at a few given places in a country, and of then comparing the results so obtained with the mean of the numbers of births and deaths for those places" (Keverberg, 1827, p. 176; Translation, Stigler, 1986, p. 165). In doing so, he saw, however, a difficulty, which he considered to be almost impossible to tackle. The problem he identified was that the laws regulating mortality and births are composed of so many elements that it is nearly impossible to "determine in advance with any precision, based on incomplete and speculative knowledge, the combination of all of these elements that in fact exists" (Keverberg, 1827, p. 176, Translation, Stigler). There would be an "infinite variety in the nature, the number, the degree of intensity, and the relative proportion of these elements", that even if such a "division of the kingdom could be accomplished […] it is likely that it would consist of such a large number of parts that there would be little advantage in terms of work saved" (Keverberg, 1827, p. 177,

Translation, Stigler). According to Keverberg, it is due to the heterogeneity of birth and death rates, which would necessitate dividing the country into almost as many units as people, which would undermine any benefit of sampling.

Be it due to Keverberg's warning or his own experience, Quételet can be seen as a "convert" (Stigler, 1986, p. 166) from a proponent of partial to complete observation. The figure of the convert will reappear later again when discussing two contemporary stories of conversion. In a letter to the Duke of Saxe-Coburg Gotha, he wrote that Laplace's "indirect method must be avoided as much as possible, although it may be useful in some cases, where the administration would have to proceed quickly" (Quetelet, 1846, p. 293; Translation, Tillé, 2020, p. 3). The main reason why he rejected partial data was because he feared its lack of accuracy: "Statistics is valuable only by its accuracy; without this essential quality, it becomes null, dangerous even, since it leads to error" (Quetelet, 1846, p. 293). He further exemplifies this in giving examples of measurements where the arithmetic means has been calculated with the appearance of a particularly high accuracy, expressed in decimals places that can, in fact, not be measured by the instruments at hand.

The person of Quetelet incarnates the complex relationship between sampling and full enumeration at this time. Drawing on Gieryn's conceptualisation, Quételet's rhetoric can be seen as a form of boundary work in that it ascribes the characteristics of accuracy only to methods of full and direct enumeration. Even more, describing approaches that diverge from this as dangerous, is a way of establishing boundaries that define the terrain of what counts as legitimate knowledge. This strong focus on accuracy became the epistemic context of 19th century statistics, in which complete enumeration became the dominant paradigm when it comes to acquiring demographic data of the population. Full enumeration became the "scientific imperative of the day" (Lusinchi, 2021, p. 4) and was seen as an improvement over mathematical "speculation" associated with the work of scholars like Laplace.

The view that imperfect knowledge leads to uncertainty and can thus be seen as speculative was widely shared among 19th century statistics. According to Porter (1986, p. 40), statisticians at this time found great value in government records like census data, allowing them to move away from speculation and guesswork that was common during the previous century. The key strength of statistics during this time was its emphasis on precise and comprehensive counting, rejecting the use of estimates and approximations. While "direct methods of complete enumeration […] gradually became popular" (Westergaard, 1916, p. 234), advances and developments in the field of sampling and representative statistics had remained particularly limited. Since Laplace, not much work has been carried out in this area.

This "cognitive frame of mind" (Lusinchi, 2021, p. 4) can paradigmatically be traced to the early debates within the International Statistical Institute (ISI), which will be the focus of the following section. The founder of the ISI, Frederic J. Mouat (1816 - 1897), wrote in his essay on the history of the Statistical Society of London that "[s]tatistical inquiry has introduced order, method, and precision, in the place of speculation, conjecture, and uncertainty" (Mouat, 1885, p. 51). Here Mouat contrasts the sample surveys that burgeoned in the 17th and 18th centuries with the primacy of full enumeration that prevailed in the 19th century. Such characterisations can also be read as a form of boundary work: Complete surveys, characterised by order, method and precision are separated from sample surveys, characterised by speculation, conjecture and uncertainty. Here we see how sample surveys did not belong to the scientific terrain and were not considered to be a legitimate source of knowledge. Within the ISI founded by Mouat, however, a change took place in the late 19th and early 20th century, which can be described as the deconstruction of a boundary and the construction of a new one. This period of time brought about a paradigm shift that cemented sampling techniques until the present day.

3.3.3 The International Statistics Institute

Adolphe Quetelet was also the founder of the International Statistical Congress (ISC), the predecessor organisation to the International Statistical Institute.[7] Between the years of 1853 and 1876, the ISC held a number of nine meetings in different European cities with the aim of standardising the topics and methods employed in national statistics. This goal was not reached and the general endeavour that the ISC set out to reach failed, the effect of which was that the congress did not continue to take place. There is a number of potential reasons as to why the ISC failed. Lusinchi (2021), for instance, lists several, such as the participation of too many lay people or the frequent rotation of individuals within the governing body. To Randeraad, "[m]uch of this failure […] can be explained by the difficulties in realizing effective knowledge transfers, in other words effective communication, in an age that was not fully prepared for truly international activities" (Randeraad, 2011). Due to the demise of the ISC, many statisticians missed an international platform to exchange ideas with other fellow statisticians, leading to the foundation of the Institut International de Statistique, the International Statistical Institute (ISI) with the goal of promoting the development and dissemination of statistics throughout the world. The ISI was formally founded in 1885, during a meeting to celebrate the anniversary of the London Statistical Society. The organization was established through the efforts of a group of prominent statisticians, who recognized the need for an international body to promote the exchange of ideas and cooperation among statisticians from different countries. The International Statistical Institute (ISI) was established to provide statisticians with a global platform for sharing ideas and knowledge. Unlike the ISC, which allowed members of the general public to attend its conferences, the ISI had stricter guidelines as to who was eligible to participate. The

---

[7] The International Statistics Institute has become subject of several recent scholarship, such as, besides the ones mentioned in this section, a special issue on "Exploring the International Statistical Institute, 1885-1938" in the Revue européenne d'histoire (Bemmann, 2023).

ISI functions as a global authority and forum for the discussion and establishment of policies and standards related to conventions in statistics. It thus played a role in the cognitive and political construction of conventions, underlying the practice of doing statistics. The organization was an exclusive, all-male society with selective recruitment policies. In this sense, the ISI can be described as a cohesive community, which "can be an unwelcoming environment to a 'deviate opinion'" (Lusinchi, 2021, p. 7), as the Norwegian Anders Nicolai Kiær had experienced.

3.3.4 Anders Nicolai Kiær: Speaking up against Full Enumeration

In the context of nineteenth century statistics, in which full enumeration was the standard, the year 1895 can be understood as the year in which the foundations of this paradigm started to be shaken. It is birth year of sampling as a recognised statistical method, a development in which the Norwegian Anders Nicolai Kiær (1838-1919) is considered the central figure. Working as a civil servant for Norway's statistical agency, Kiær assumed the role of leading the Statistical Division of the Norwegian Ministry of the Interior in 1867 and understood early on that an independent and centralised institution for the gathering and analysis of data is necessary. In 1876, the Statistisk sentralbyrå, the Central Bureau of Statistics, was established after he advocated for its formation, and he became its inaugural director, a post he retained until his retirement in 1913. Kiær also introduced the newly developed electronic punch card tabulator in 1894, after having seen it at a great exhibition accompanying the 1893 ISI conference in Chicago (see Heide, 2009). Most important, however, were Kiær's pioneering efforts in the development of sample surveys, through which he prepared the ground for following theoretical developments up to the current time. In his early sample surveys, Thomsen (in: Alterman, 2001, p. 210) states that he already intuitively used design methods that are still used today, such as "stratification, selection in several stages, area sampling and unequal selection probabilities". Kiær called his approach 'representative method', and first presented it during a meeting of the ISI in Bern in

1895 and was to present and further develop it at other ISI meetings in the following years. During those scholarly conferences, Kiær caused a crisis when he questioned the dominant paradigm of the time: full enumeration. Bethlehem (2009), drawing on Kuhn, describes this point in time as an "intellectually violent revolution", forcefully introducing a new paradigm in statistical thought. Others, such as Kuusela (2011, p. 30) argue that since the representative method coexisted alongside censuses, it did not constitute a paradigm in a Kuhnian sense.

In examining the developments within the ISI, the following sections also contribute to the field of social epistemology. Social epistemology is a branch of epistemology, challenging the idea of solitary knowledge acquisition and instead arguing that knowledge is generated with the participation of others. The Royal Society, for instance, was founded in 1660 with the purpose to highlight the importance of multiple observers in establishing recognized facts. In a similar way, the ISI can be seen as an instance of how scientific consensus on the controversial issue of sampling was socially constituted. The following analysis details the progression of the concept of a sample representative of a population through various stages characterized by the formation and disruption of paradigms through the process of boundary work within the context of several ISI meetings.

3.4   From Complete Enumeration to Sampling

3.4.1   The Representative Method

At the time of the 1895 ISI meeting, Kiær had already been conducting sample studies in Norway for more than 15 years and was convinced that full enumeration was not always necessary. While full enumeration had the status of the gold standard by this time, the introduction of sampling and thereby of calculation rather than of mere counting was considered a disruption of the then conception of normal science. The motivation to conduct surveys based on a sample was, as Bellhouse concludes, "to increase the scope of the large-scale statistical investigation he was

carrying out" (Bellhouse, 1988, p. 3), something that could not be managed when relying on full enumeration. The central idea of what he called the "representative method" was to select a sample that approximates the general population one wants to further know. Kiær (1899) described his representative method as follows:

> By representative investigation I understand a partial exploration with observations on a large number of scattered localities, distributed over the whole territory so that they form a miniature of that whole. The localities are not chosen arbitrarily, but according to a rational grouping based on census results; and the results should be controlled by comparison with those censuses (Kiær 1899 in: Kruskal and Mosteller, 1980, p. 176).

The representative method was thus a way to construct an image of a population that resembles it in important characteristics. Importantly, certain knowledge about those characteristics had to be in place, meaning that it was necessary to have valid data through censuses at hand. Kiær's suggestion constitutes a more sophisticated progression of the ideas already laid out by people like Laplace. A useful description of Kiær's representative method can also be found in the work of Kruskal and Mosteller (1980):

> First, he thought of social and economic surveys in which one could begin by choosing districts, towns, parts of cities, streets, etc., to be followed by systematic, rather than probabilistic, choice of units (houses, families, individuals). Second, he insisted on substantial sample sizes at all levels of such a selection process. Third, he emphasized the need for spreading out the sample in a variety of ways, primarily geographically, but in other ways as well. For example, if a sample had a deficiency of cattle farmers, he would add more of them (Kruskal and Mosteller, 1980, p. 175).

As per Kiær's description and Kruskal and Mosteller's explanation, the sample should therefore be chosen to *represent* the population, hence the name representative method. It is a so-called purposive sample, as it was purposively selected according to certain selection criteria and can be distinguished from a random sample, a distinction that can be made in retrospectively looking at its development. He aimed to identify "rational selection procedures" to create "miniature populations" and draw inferences from them. Important to note is that Kiær justified his approach empirically but could not provide a theorical justification for why it worked (Seng,

1951), a task that should be later taken up by fellow statisticians. As we shall see now, from the first presentation at the ISI conference in Bern in 1895, Kiær's introduction of sampling methods was met with little approval within the scientific community, which was largely dismissive of his suggestions and argued for maintaining the paradigm of full enumeration.

3.4.2   Claiming the Boundaries of Epistemic Authority

Kiær's proposal, therefore led to controversy and rejection within the discipline. One of his main critics was the German statistician Georg von Mayr (1841-1925), director of the Bavarian Statistical Bureau and involved in Bismarck's imperial state. Georg von Mayr also was a member of the International Statistical Institute from its beginning on in 1885 and served as its vice-president from 1911 until 1923. His research was founded on the conception of statistics as an autonomous discipline, disentangled from disciplines such as political economy or geography. He defended an understanding of descriptive statistics as an "exact science of mass social phenomena" (Hertz in: Alterman, 2001, p. 211) and refused all attempts in mathematical statistics. Von Mayr also vehemently opposed Quételet's views on the 'average man' throughout his life, distrustful of statistical methods based on a priori mathematical arguments.

Von Mayr's position towards Kiær's suggestion of the representative method can also be viewed in light of those contextualizations. His position is nicely expressed in the following statement, which he made in the discussion to Kiær's suggestion, that also stands for the then existing belief that observation, thus full enumeration (rather than calculation), allows for the sole legitimate access to representing reality:

> I regard as most dangerous the point of view found in his [Kiær 's] work. I understand that representative samples can have some value, but it is a value restricted to terrain already illuminated by full coverage. One cannot replace by calculation the real observation of facts. A sample provides statistics for the units actually observed, but not true statistics for the entire terrain. It is especially dangerous to propose representative sampling in the midst of an assembly of statisticians. Perhaps for legislative or administrative goals sampling may have uses - but one must never forget that it cannot replace a complete survey. It is necessary to add that there is among us these days a current in the minds of mathematicians that would,

in many ways, have us calculate rather than observe. We must remain firm and say: no calculations when observations can be made (von Mayr in: IIS, 1895, p. xciv; Translation, Kruskal and Mosteller, 1980, p. 174,175).

Mayr's argument against the representative method resonates with earlier critiques of sampling, such as the way Baron de Keverberg argued against Laplace's suggestion of using partial data. Another objection during the same ISI meeting was raised by the Swiss Edmund Wilhelm Milliet (1857 - 1931), who served as the director of the Federal Bureau of Alcohol in Bern, Switzerland. In opposing Kiær, he stated the following:

> I believe that it is not right to give a congressional voice to the representative method (which can only be an expedient) an importance that serious statistics will never recognize. No doubt, statistics made with this method, or, as I might call it, statistics, *pars pro toto*, has given us here and there interesting information; but its principle is so much in contradiction with the demands of the statistical method that as statisticians, we should not grant to imperfect things the same right of bourgeoisie, so to speak, that we accord to the ideal that scientifically we propose to reach (Milliet in: IIS, 1895, p. xcv; Translation, Tillé, 2020, p. 4).

The reason for the long rejection of Kiær's approach, is, according to Porter (1986), that due to the lack of reliable information on the characteristics of the total population, census surveys were consistently preferred until the end of the 19th century. Another reason was highlighted by Adolph Jensen (1866 –1948), a Danish economist and statistician who served as the head of the Statistics Department of the Danish Ministry of Finance from 1913 to 1936. The reason he mentioned was one of trust. He argues that the trust built between the official statistical service and the population is crucial for the success of statistics. In his 1925 report on the representative method, he stated that while official statistics must be cautious of their reputation, they should not prevent justifiable development solely for the sake of prestige, and instead build trust by guaranteeing the reliability of their work.

> This objection contains the real kernel, that the greatest importance must be attached to the existence of a state of mutual trust between the official statistical service and the population which both supplies the material for the statistics and for whose sake all this work is done. The official statistics ought of course to be exceedingly cautious of its reputation – 'it is not sufficient that Caesar's wife is virtuous, all the world must be convinced of her virtue.' But

it would hardly be warrantable purely out of regard to prestige, to prevent a development which in itself is acknowledged to be really justified. One does not omit to build a bridge with a special design because the public, in its ignorance, distrusts the design; one builds the bridge when the engineer can guarantee its bearing strength, - people will then use the bridge in due course and rely on its solidity (Jensen, 1925 in Desrosières, 1998, p. 232).

Although Kiær was met with fierce resistance, he did not abandon his method, but developed it further and presented it at subsequent ISI events. In the development of sampling and thus of the representative method, Desrosières (1998) highlights particularly two periods. First, the time between 1895 and 1903, in which the fundamental questions as to whether "one could legitimately replace the whole by a part" (Desrosières, 1998, p. 211) and whether this way was superior to census surveys were discussed. Second, the time between 1925 and 1934 can be characterised by a debate on the choice between 'random sampling' and 'purposive sampling'. While the committee at the ISI meeting in Berlin in 1903 concluded that if a sample is carefully selected, the method could be recommended, the question discussed in the 1925 meeting in Rome was on how to select it. From the view of social epistemology, this is particularly interesting in that the verdict as to whether sampling can be seen as a trustworthy endeavour was made by a committee within the ISI.

In the subsequent ISI meetings in 1901 in Budapest and in 1903 in Berlin, Kiær continued to advocate for his method and started to receive some support. During the 1903 meeting Lucien March (1859 - 1933) supported and later suggested a compromise amidst the discussions of the representative method. March was Director of the Statistique Generale de la France, the French Statistical Bureau between 1896 and 1920 and introduced concepts akin to today's techniques of simple random sampling without replacement and simple random cluster sampling (Bellhouse, 1988, p. 4). In his discussion of the representative method, March paid attention to the fact that the representative method does have a predecessor in Laplace and the 18th century French population statistics, as previously elucidated in this chapter, but also to more recent developments. For instance, he referred to the 1891-93 salary study of French workers, wherein

a sample of one-fifth of the populace was utilized, and which precision was ascertained by comparing it with older samplings and administrative records (Kruskal and Mosteller, 1980, pp. 177–178). By the time of the 1903 ISI meeting, a general consensus arose that sampling is a valuable means to gain knowledge about populations and "the time had come for amplification and refinement" (Lusinchi, 2021, p. 13). In this context, March suggested that randomisation could provide a way of giving a scientific basis to the sampling procedure. Interestingly, while Kiær did mention the use of randomization, stating that a sample could be "selected through the drawing of lots" (Kiær in: Bellhouse, 1988, p. 3), he did not explore this possibility further.

3.4.3   Taking the idea further

This idea was eventually taken further by Arthur Lyon Bowley (1869-1957), a pioneer in creating statistical instruments for examining issues related to economic and social concerns. Bowley, who held the first chair in statistics at the London School of Economics in 1919, is mainly remembered for his development of sampling techniques. He not only played a central role in convincing the ISI of Kiær's method but also took important steps to emphasise the importance of randomly selecting participants. This also stems from the fact that during the 1890s, Bowley was already fascinated by the work around the "representative method" conducted under Kiær's guidance. What is more, Bowley was long sceptical of the epistemic climate that favoured full enumeration, an issue he famously mentioned in his presidential address to the British Association for the Advancement of Science in 1906, where he criticised the widely held conviction that full enumeration was the only feasible way to obtain accurate demographic studies:

> The method of sampling is […] so persistently neglected, and even when it is used the test of precision is ignored. We are thus throwing aside a very powerful weapon of research. It is frequently impossible to cover a whole area, as the census does […] but it is not necessary. We can obtain as good results as we please by sampling, and very often quite small samples are enough; the only difficulty is to ensure that every person or thing has the same chance of inclusion in the investigation (Bowley, 1906, p. 553).

In 1912, Bowley conducted a random-based survey in Reading, which was followed by further similar studies under his supervision in Northampton, Warrington, and Stanley. It can be said that it was Bowley who connected the statistical theory of sampling with social research, he "made the decisive methodological breakthroughs for the social surveys as we know it today" (Marsh, 1982, p. 25).

Even though there was no discussion on Kiær's method during the ISI meetings between 1903 and 1925, since the 'representative method' was generally accepted, the example of Bowley shows that a gradual shift took place in this time, leading to a consolidation of sampling methods and developments of Kiær's original suggestion.[8] At this time, there were two different approaches to sampling, purposive and random sampling. Whereas in the case of the former, Kiaer's 'representative method', knowledge about the population needed to be in place in order to purposively chose the participants under study, the latter does not rely on prior information, since the random selection guarantees representativity. Due to his fierce defence of the representative method, Bowley became a member of an ISI committee in 1924 (together with Corrado Gini, Adolph Jensen, Lucien March, Verrijn Stuart, and Frantz Zizek) tasked to assess the practice of sampling. The commission developed a report on the "Measurement of the precision attained in sampling", which was presented at the 1925 ISI Congress in Rome, arguing for accepting the principle of sampling under the condition that the methodology is respected. This set the stage for the modern theory of sample surveys, 30 years after Kiær's initial suggestion and six years after his death. Since 1925, Kiær's purposive selection and Bowley's random selection were both recognised by the ISI as two accepted approaches to sampling, both

---

[8]   The reasons why there was no discussion of the representative method is difficult to respond to. Kruskal and Mosteller (1980) state that it is "mysterious". They argue that World War I does not solely explain it; additionally, Kiaer worked and published on the topic until his death in 1919. Lusinchi (2021) argues that while the War put a hiatus to the ISI meetings, by 1903 the representative method had been accepted as legitimate and statisticians had started to refine and improve the method.

opposing methods of full two enumeration. These approaches to sampling correspond to Kiær's conception of a good sample, as drawn through "rational procedures" and as a "miniature population". Both approaches, however, relate to two essentially different scientific approaches. While the validation of purposive selection can only be reached "through experimentation by comparing the obtained estimations to census results" (Tillé, 2020, p. 5), the validation in the case of random selection is "based on the calculation of probabilities that allows confidence intervals to be built for certain parameters" (Tillé, 2020, p. 5). Random methods are validated based on mathematical principles; purposive methods based on experimental approaches, a continuum that still shapes contemporary discussions.

The year 1925 thus marks the beginning of a scientific consensus regarding the validity of sampling techniques. The change that took place within the ISI has nicely been summarized by Adolph Jensen, who in 1926 stated that while the debates from 1903 were mainly about the "recognition of the method in principle": there are now barely any statisticians "who in principle will contest the legitimacy of the representative method" (Jensen, Saenger and Bowley, 1926, p. 59). The subsequent ten years were then marked by another controversy over the question as to which sampling procedure was superior.

3.4.4   An Ostensible Triumph

3.4.4.1   Some Notes on Closure

Before further exploring the development of how sampling turned from a widely rejected outsider approach to becoming the central tenet of survey research and polling, I would like to introduce the notion of closure. Theories of closure were developed by Pinch and Bijker (1984), who draw a picture of technology that is almost exclusively shaped by social processes. Closure involves the integration of various components, such as technical features, market requirements, and social norms, into a coherent and durable configuration. The concept of closure highlights the social and political dimensions of technological change, emphasizing that the adoption and

use of technology is not simply a matter of technical efficiency, but rather shaped by a complex set of social and cultural factors. The basic underlying assumption here is that stabilized technologies are always the result of long and complex social processes and negotiations. Following this line of thought, both science and technology can be investigated in terms of their socially constructed nature and the social patterns and mechanisms that are effective in the process, an approach that became known as SCOT (Social Construction of Technology).

Pinch and Bijker (1984) suggest two forms of closure: rhetorical closure, which, in the area of science relates to "some 'crucial' experimental result, 'definitive' proof or 'knockdown' argument which has the effect of closing the debate on some controversial issue" (Pinch and Bijker, 1984, p. 425). Those results may not convince the scientist from the "core-set", but rather the wider community. The second form is closure by redefinition of the problem. In this case, a controversy can be stabilized when the technology in question is being used to solve a different problem. While purposive sampling may not be the superior way to conduct surveys, it might still serve as a cheaper way if quality standards are deemed less important.

In slight contrast to SCOT, a perspective informed by Actor-Network Theory (ANT) would highlight a process-oriented understanding, focusing not only on the construction and implementation of a scientific approach or a technology but also on its usage. In this sense, the work that needs to be done in order to stabilize a technology becomes the crucial locus of sociological investigation. Despite it already being stable, its stability needs to be maintained. Similar to SCOT, ANT also describes a form of social constructivism, understanding scientific knowledge as an effect of established relations between heterogeneous entities involved in scientific practices and forming a network. An established network implies, as highlighted by Detel (2001, p. 14265), implies a form of closure in the sense that it limits the entry of other actors and relations into the established network, a prerequisite for the accumulation of scientific knowledge.

3.4.4.2   Closed, once and for all?

What those approaches to sampling lacked was a theoretical justification as to why they worked. The next step in the process was thus to construct a statistical theory of survey sampling, a decisive role in the development of which was played by Jerzy Neyman (1894-1981). Neyman was born in Poland and studied mathematics at the University of Kharkov, where he also became a lecturer. After obtaining his doctorate degree from the University of Warsaw, he visited University College London, where he worked in Karl Pearson's laboratory, where a long-lasting collaboration with Egon Pearson, Karl Pearson's son, was first formed. This collaboration led to the theory of hypothesis testing and Neyman's invention of confidence intervals.

In his ground-breaking paper, Neyman (1934) not only developed the theory of confidence intervals but also showed that Kiær's representative method led to incomplete results. Confidence intervals indicate a range within which the true population value is likely to fall with a certain level of probability, they give us an idea of where the real population value probably is. Such statements could not be made in the case of Kiær's representative method. Neyman states that if one is interested in a "collective character X of a population π and use methods of sampling and of estimation" that allow ascribing confidence intervals, "*whatever the unknown properties of the population*, I should call the method of sampling representative and the method of estimation consistent. We have seen that the method of random sampling allows a consistent estimate of the average X whatever the properties of the population" (Neyman, 1934: 585-586, emphasis original). The emphasis "whatever the unknown properties of the population" is crucial to understanding the value of this framework. In the case of purposive methods, properties of the population need to be known to draw representative inferences. The sample itself does not matter in this case; it is the sampling design itself that justifies the inference. Neyman's 1934 paper played, as Hansen *et al.* put it, "a paramount role in promoting theoretical research, methodological developments, and applications of what is now known as probability sampling"

(Hansen, Dalenius and Tepping, 1985, p. 330). It now counts as one of the founding texts of sampling theory and has become highly influential and important. Since this time, probability sampling has become the preferred approach to sampling.

One very interesting and important aspect of Neyman's paper is that it can be understood as an attack on the work of Gini and Galvani. Their goal was to choose a subset of responses from the Italian census as they needed to make space for storing the files related to the 1931 census. They thus decided on a purposive selection method based on seven variables, leading to a selection of forms from 29 out of 214 administrative districts from the 1921 Italian Census, which they considered to reflect the Italian population.

Neyman's paper can be seen as a type of rhetorical closure, heralding a turning point when it comes to the use of purposive sampling. Neyman, to some extent, closed the controversy over which sampling procedure was better, setting the stage for random sampling, to be elevated to the gold standard. As Lusinchi (2018) puts it, Neyman "de-legitimized the purposive/quota method", so that from this time on, both approaches became two "competing norms within the community of polling practitioners" (Lusinchi, 2018, p. 15).

3.4.5 Agriculture and Democracy: Representative Surveys Beyond the ISI

The narrative of the development of "representative surveys", as outlined in this chapter, gives great credit to Kiær, highlighting how his concepts evolved through discussions at the International Statistics Institute. Before concluding this chapter, it is, however, important to also explore a separate tradition. This tradition, which developed independently of the discussions at the ISI, similarly introduced the idea of a representative sample. Although it had a relatively minor impact on the established theory of surveys, its contributions are still relevant and deserving of attention. In *Sampling and Democracy: Representativeness in the First United States Surveys*, Didier (2002) shows how the emergence of statistical representativeness in the US can be traced back to the Department of Agriculture's efforts since the 1860. Didier shows

how agricultural statisticians focused on greating representative groups of farmers to gather data on crop production, suggesting to look to democratic theory to identify justifications for their representativeness.

Initially, representativeness in agricultural surveys was pursued by selecting "voluntary crop reporters" (Didier, 2002, p. 428) who were were chosen for "their known intelligence and judgment" (Taylor and Taylor 1952, 188). This selection was intended to ensure that the statistical samples represented broader agricultural trends rather than individual discrepancies. However, a notable shift occurred, following a significant error in 1921 when the reported data did not accurately reflect actual conditions. The reason was that farmers reported to follow a policy change, which they in fact did not. This incident led to a reevaluation of the methods used to gather data, emphasizing the need for more objective data collection strategies. Post-1921, the strategy shifted towards using mechanical tools like the "crop meter" and more stringent criteria for selecting respondents, who were now required to be actual farmers, thus expected to provide more accurate and direct observations of their agricultural output. While before, crop reporters measured an area of crop in their "locality", post-1921, they reported only on their farm, introducing a different nature of reported space. The farm was fully under control of the farmer and led to a more individualistic approach focused solely on the conditions within a single farm's boundaries. The single farmer now served as an example, best illustrated by the introduction of a new term, that of a sample.

While this new approach promosed more objectivity, their non-response rates were quite high, leading the Department of Agriculture to turn to another actor: the Post Office Department. The ways of distribution within the post office allowed the Department of Agriculture to chose as many farms as necessary to be able to receive an accurate picture of the area in question. Thus, while previously, the system depended on the loyalty between farmers and the statistical system, the new method, involving more farmers and reducing the need for established relationships,

responses could be efficiently gathered on a larger scale. Mail carriers facilitated this process by directly questioning farmers and completing the questionnaires themselves.

Didier draws a parallel between the principles of representative democracy and the statistical methods employed. In both systems, a small subset is selected to represent the larger whole. Just as in a democratic election where representatives are chosen to speak on behalf of a larger population, statistical sampling involves selecting a portion of the population to infer the characteristics of the entire group. For instance, the concept that representatives (like crop reporters) should embody the broader characteristics and interests of the communities they represent, is similar to the way statistical samples are expected to accurately reflect the broader population's traits. This includes the idea that the selection process, whether for political representatives or statistical samples, must aim to construct a microcosm that is genuinely representative of the larger system. Didier thus argues that the methodologies developed for statistical sampling in agricultural surveys share foundational principles with the processes used to elect representatives in a democracy.What we have just seen in turning to Didier's (2002) on the history of agricultural statistics in the US, we can see that in different domains, similar ideas of representativeness and sampling evolved outside the realm of the ISI. But, as Didier also notes, Kiaer and the developments at the ISI still had the greatest influence on the invention of representative surveys, "for it was indeed his speeches that had the strongest influence on the theory of surveys as we know it today" (Didier, 2002, 427).

3.5   Conclusion: The Cycle of Representation

The development described in this chapter, particularly the origins of random sampling, laid the foundation for what should guide the field of survey research and polling throughout the second half of the 20th century until our time. Despite random sampling having been well developed and established in the mid-1930s, it was, rather poorly implemented in practice. As we shall see in the next chapter, most pollsters continued to work with quota samples - a form of purposive

sampling - until its shortcomings became apparent in the light of the 1948 US presidential election. It was really only in the aftermath of this famous miscall that random sampling was elevated to become the gold-standard in polling. Before this, polling already reached *predictive stability*, referring to the "socially shared expectation that predictive claims point toward the right direction" (Griessl, 2022, p. 116). Its stability was, however, only due to the pollsters predicting the right outcome, not due to the way how they reached this particular outcome. Stability is here used as an alternative to the notion of closure, highlighting the establishment of stable networks between sampling instruments, stakeholders, respondents, the public, polling companies and others, rather than rhetorical closure in the sense of scientific arguments with the expectation that everybody will adapt accordingly.

This chapter has explicated the evolution of sampling within the context of the social history of its disciplinary context. This development has been characterized by a dynamic interplay of consensus and controversy, forming what can be termed as the "cycle of representation." Until today, the discipline continues to remain within this cycle, oscillating between two dominant approaches: purposive sampling and random sampling. Consequently, the impact of this ongoing cycle endures, exerting a significant influence on the field. Moreover, other aspects that we came across in this chapter, such as the practices of boundary work, consistently shape and define the field. The next chapter will further explore this history, paying particular attention to the implementation of sampling in the field of public opinion polling and how developments in sampling brought about a certain conceptualization of public opinion.

CHAPTER 4

4 Sampling, Polling and Public Opinion

4.1 Introduction

From today's perspective, sampling is no longer a controversial issue, and it might seem odd to some to think that it was highly controversial only just over 100 years ago. While it was a marginal instrument at the beginning of the 20th century, it became a fundamental instrument to gain knowledge about populations. In his 1948 contribution to the *Symposium on Research Frontiers in Human Relations,* Rensis Likert, known for the development of the Likert scale, described the sample survey as "one of the most valuable and powerful research instruments of the social sciences" (Likert, 1948a, p. 341). One application that evolved from the developments of the sample survey was the public opinion poll, a technique that would become of crucial importance to people's lives, the way governments function and society understands itself. A key figure in this development was George Gallup (1901 – 1984), whom I will introduce in more detail in due course. First published in 1940, George Gallup and Saul F. Rae's book *The Pulse of Democracy: The Public Opinion Poll and How it Works* starts with a series of assumptions, which help to better understand the development of public opinion polling as a result of developments in sampling.

> What is the common man thinking? The life history of democracy can be traced as an unceasing search for an answer to this vital question. The following pages provide a modern answer on the basis, not of guesswork, but of facts. They tell the story of a new instrument – the public opinion poll – and describe how it works to provide a continuous chart of the opinions of the man in the street. […] The application of sampling has made it possible to predict the divisions and trends of public sentiment with a high degree of accuracy (Gallup and Rae, 1968, p. v).

What we can take from this is the idea that (1) public opinion exists, (2) can be accurately measured through (3) the use of sampling techniques and is a crucial aspect in (4) democratic societies. Gallup and Rae thus claim that the public opinion poll is a way to scientifically know what the people think on almost all matters of public concern. Throughout this chapter, I will show how the development of survey sampling and its application in the public opinion poll co-produced a particular understanding of public opinion, rendering public opinion the aggregation of individual opinions. The notion of co-production refers to the idea that social and technical elements are intertwined and shape each other during the process of knowledge production (see Jasanoff, 2010).

Whereas the previous chapter has drawn on the development of a particular technique, this chapter looks into the way this technique brought a particular reality, that is, public opinion, into being. It thus explores how public opinion evolved as a result of advances in sampling techniques, showing how those who cultivated and owned those techniques were able to monopolize their access to public opinion. In defining public opinion as the outcome of what public opinion polls measure and in promoting the idea of polling as the best and most accurate way to gain knowledge about people's opinions, pollsters and polling agencies became the most trusted entities, endowed with the power to know, define and delineate what the people think. Pollsters and polling agencies have thus made themselves and the survey sample essential for everyone who wants to know public opinion.

This chapter will proceed as follows. It will draw on the previous chapter in so far as it connects the history of survey sampling and the history of polling. It will be shown how despite major success in the field of survey sampling in the 1930s, it took quite a while until those developments became fruitful in the field of election forecasting and opinion polling. It was only after two famous miscalls in the context of the 1936 and 1948 US presidential elections that random sampling evolved to become the dominant way of conducting polls, a development that

played a crucial role in producing the view of public opinion that we still hold on to today. In order to elucidate this continuation of the cycle of representation in the progression of public opinion polling, this chapter elaborates on the history of polling, followed by a presentation of how certain forms of sampling found its way into the practice of polling, culminating in a general acceptance of random sampling as the gold standard. In doing so, I will also present sociological critique towards this development and will end with an account of how pollsters were able to monopolize access to and representation of public opinion in their hands.

4.2 The Embarrassing Roots of the Emergence of a Research Programme

4.2.1 The Pre-History of Polling

The history of polling for election forecasting dates back at least to the year 1824 and the time of the run-up to the American presidential elections between John Quincy Adams, Andrew Jackson, William H. Crawford, and Henry Clay. Forecasts produced in this time are to Gallup and Rae the "earliest counterpart of modern opinion surveys" (Gallup and Rae, 1968, pp. 34–35) and took place in a particular historical and political constellation. The 1824 US presidential election can be regarded as a realignment in American politics. Since the 1790s, the American party system was shared by Federalists and Democratic-Republicans, whereas the former ceased to be a relevant political power prior to the 1824 election, leaving the Democratic-Republican Party as the sole party. Lacking an opposition, the party split and four of its candidates vied for the presidency. The outcome of the election was thus highly uncertain due to the multiple Democratic-Republican candidates, making past voting behaviour an unreliable predictor (cf. Smith, 1990, p. 23).[9] These uncertain circumstances thus led to attempts by politicians, newspapers, and others to try to predict the outcome of those elections. Some drew on very

---

[9]   Smith (1990) mentions other factors, such as the fact that multiple candidates were running without a party label and more fundamental change taking place in the political system. The right to vote was extended to all white males and the direct election of electors was introduced.

interesting proxies to acquire a glimpse into the preferences of the population. For example, the number of toasts made to the candidates during the Fourth of July celebrations in Pennsylvania was seen as an indicator of support. Despite drawing on such proxies, people began to conduct, what was later characterized as straw polls during public meetings, such as militia musters or tax gatherings. Those first straw polls "emerged out of a desire of people both to know public opinion on the presidential candidates and to express their own opinions" (Smith, 1990, p. 30). In this sense, they already served similar goals as today, in that polls were viewed as a way of political participation and a tool to learn about public sentiments. Results of those early polls were communicated to newspapers that were highly interested in those numbers and disseminated them to the public. Some newspapers already summed up some of those results in order to provide their readers with impressive election forecasts. Importantly, those newspapers did not conduct or sponsor those polls themselves; rather, they aimed, as Smith (1990) notes, to forge public opinion in favour of the candidate that they supported.

According to Smith, the advent of early straw polls in the United States was shaped by three significant trends in American history: democratization, centralization, and quantification. The desire to anticipate the outcome of democratic elections has been a persistent concern, especially during moments of unprecedented political circumstances, such as in 1824. Moreover, achieving a national perspective requires the aggregation of local reports, which reflects the increasing centralization of political power in the United States. Furthermore, the reliance on numerical data in elections was in line with the development of sample counting as a method for estimating voting outcomes, thereby underscoring the growing importance of quantification in American political culture.

Those early polls are usually termed 'straw polls', which is important to consider in light of later developments. There are different interpretations regarding the origin of the term 'straw' poll. One interpretation highlights the use of a straw poll as a way to gauge public opinion in a

quick and informal manner. This is depicted through the metaphor of using a straw, much like holding a straw up to see which way the wind is blowing. This metaphor underlines the idea that a straw poll is not a conclusive measure of public opinion but rather a rough estimate. Smith (1990) also highlights that those early straw polls were initiated by partisan operatives who were interested in gauging the political climate and, if their candidate was leading in the polls, they would use the results to influence public opinion either by highlighting the popularity of their candidate or by portraying the other candidates as hopeless. Another interpretation of the term is rooted in historical practices, where straws were used as a tool for voting. During in-person voting in public, voters would drop a straw into a container to indicate their choice. The tall straws indicated a vote for one candidate, and short straws for another - the candidate with the tallest straws was considered the winner. Over time, the term "straw poll" has evolved to encompass any informal, unscientific poll.

In light of this, these early polls were often biased and faced criticism for their representativeness early on. In addition to the lack of knowledge and understanding of statistics and probability theory, Newsome notes, for instance, that in the context of the 1824 US elections, Jackson supporters influenced military musters in that officers "would treat their men, make them drunk, and then raise the war whoop for General Jackson […] [so that] the result was always in favor of Jackson" (Newsome, 1939, p. 139). Another issue that arose was that people not eligible to vote were counted in those straw polls, raising further representation issues. Some of those claims were challenged at the time. Smith, for instance, quotes a statement made by the *Carolina Observer*, suggesting that while they may not have information from all districts, the opinions expressed in those they do, should, however, correspond to the sentiment in other districts.

While those early straw polls might have satisfied a desire for insights into the future, their predictive stability was very little. This would change as later researchers and newspapers focused on refining the reach of the samples, as I discuss below.

4.2.2   Quantity does not mean Quality: The Literary Digest

As those polls gained in popularity, a particular practice emerged, namely that newspapers printed questions on current political issues, which readers could fill out and send back to the editorial offices of the newspapers. Extending the participants from the readership to a larger public, some polls started to also use lists, such as from telephone or car owners as a basis to contact potential respondents. Newspaper-run polls were not only concerned with election forecasts but covered all kinds of topics. A prevalent topic at this time was, for instance, the question of whether the United States should be involved in World War I, a question that should also become subject to several polls (see e.g., Hicks, 1949). The interest in those methods increased a lot, such that within ten years, from 1916 until 1926, over 60 polling institutes were created (Keller, 2001, p. 33). Among those, the American weekly magazine the Literary Digest stood out: Already in 1895, the year in which Kiær first proposed the representative method in front of the ISI, its file contained over 350,000 addresses and grew to 32 million by 1932. For example, to predict the outcome of the 1928 election between Herbert Hoover and Al Smith, the Literary Digest distributed 18 million questionnaires and could accurately forecast Hoover's win (cf. Keller, 2001, p. 33p). Considering the history of sampling, as presented in the previous chapter, it becomes clear that while sampling methods were already much more advanced by this time, as compared to the early straw polls discussed in the previous section, its application and implementation in practice was very limited. The polling world, so it seems, did not take much notice of the important and path-breaking developments in sampling. It took, as we shall see now, two major miscalls in election polling until those approaches were applied in practice.

Despite its lack of scientific rigor, the Literary Digest's election forecasts were widely trusted due to their past accuracy. From 1916 to the mid-1930s, they successfully predicted the outcome of five presidential elections, despite not using sophisticated sampling methods. It was precisely due to their success that their sampling approach went unchallenged.[10] Because of the Digest's successful track record, its lack of sophistication went unnoticed. This situation would, however, change in the aftermath of a major US presidential election. In 1936, the Digest unintentionally tarnished its reputation for polling accuracy in a lasting way when it made a crucial error by predicting a landslide victory of the Republican Alf Landon over the Democrat Franklin D. Roosevelt. The Digest's poll had a sample size of about ten million voters and a response rate of about 25 per cent, which was the basis for the prediction in its 31 October 1936 issue, stating that Landon would win 370 out of 531 electoral votes. After the election, however, the picture was completely different: Landon won only eight electoral votes, while Roosevelt won 523. In complete contrast to the Literary Digest's prediction, Roosevelt achieved an unprecedented electoral victory by securing an astounding 98.5% of the available electoral votes, marking, apart from the unanimous elections of George Washington and James Monroe, the most significant electoral landslide in the history of the current two-party system established in the 1850s. Roosevelt's dominance extended across all states, with the exception of Maine and Vermont.

The reason for the Digest's incorrect prediction is usually believed to be coverage bias: respondents were recruited primarily from the telephone directory and a register of car owners, resulting in more wealthy people being sampled than less wealthy. While this is the usual explanation about the misprediction, Lusinchi challenges this "conventional explanation" (Erikson and Tedin, 1981, p. 953), which has also been repeatedly stated by early pioneers such

---

[10]   In Griessl (2022), I explore the question as to why the lack of sophisticated sampling methods was not a problem at this time, suggesting that due to its successful predictions, the Literary Digest already reached *predictive stability,* which is a feature of the networks of social expectations in which polling procedures were situated, rather than a feature of its methodological sophistication.

George Gallup: "The failure of the Literary Digest's polling approach can be explained rather simply. The Digest's sample of voters was drawn from lists of automobile and telephone owners" (Gallup, 1976, p. 147). In drawing on a 1937 Gallup poll, Lusinchi argues that it could be said that telephone and car owners supported Roosevelt and that it was not them who caused the inaccurate prediction, but rather those who did not participate in the poll - who happened to be mainly supporters of Roosevelt. Without going deeper into the reasons of the Digest's misprediction, it is important to note the consequences for the magazine itself. As put by Keller (2001, p. 34), many lampooned the Digest as a false "oracle" with unearned authority as a source of reliable prediction. Its failure in 1936 exposed the Digest to widespread ridicule, with the effect that its credibility was severely damaged by this failure. The magazine went out of business two years later, and it was no longer able to maintain predictive stability since the long-lasting success was diminished. Another person, however, could claim this as an episode of success and as the starting point for a different research programme in the field of polling, which should, for a certain period of time at least, become very successful in assembling predictive stability. What was needed was the implementation of sampling methods into the practice of polling.

4.2.3   George Gallup and the New Charisma of Polling

The adoption of a new research programme involved a series of triumphant election calls and less triumphant ones involving George Gallup, which illuminated the real-world electoral stakes of sampling methods. Gallup was a pioneer and one of the leading figures in the implementation of sampling methods for opinion polling and was recognised as one of the 100 most influential Americans of all time by the Atlantic in 2006. He was the founder of the American Institute of Public Opinion in 1935 and the first to apply probability theory to opinion polls and election forecasts. Gallup became particularly prominent during the course of the 1936 US presidential elections and was able to become a trusted brand name in the area of market and opinion research. While the Digest embarrassingly predicted Landon to win, Gallup correctly, even

though he overestimated Roosevelt's vote share, predicted Roosevelt's victory based on a quota sample involving about 50,000 people. In the aftermath of this triumphant prediction, Gallup engaged in a form of boundary work to promote his own research as scientific and denouncing this label to the work conducted by the Literary Digest: "When you do that, you are operating a scientific poll; when you do not do it, you are conducting a straw poll" (Gallup, 1936, p. 371). Even more interesting, as Lusinchi (2017b) has analyzed in great detail, Gallup strategically confounded quota and random sampling in interviews, articles and more, to make it look like he was already using the superior and scientifically proven approach of sampling, while he actually drew on quotas: "[A]s a symbol of science, random sampling was used to enhance the legitimacy of the pollsters' brand of polling, even when their actual practice did not conform to the theory" (Lusinchi, 2017b, p. 123). In doing so, Gallup could elevate his own work to promote it as a methodologically and scientifically sound practice in comparison to the Digest.

Another interesting episode involving Gallup and the way he promoted his work as superior to the Digest started shortly before the 1936 elections. During the 1936 election campaign between Landon and Roosevelt, Gallup announced a bet, namely that he could predict the election results better than the established and widely read Literary Digest. On July 12, 1936, Gallup featured the forthcoming Literary Digest poll as the subject of his *America Speaks!* column: "If the Literary Digest were conducting its poll at the present time, following its usual procedure, Landon would be shown in the lead. The actual figure would be in the neighborhood of 44 per cent for Roosevelt, 56 per cent for Landon" (Gallup in: Lusinchi, 2017b, p. 122). And indeed, the Literary Digest predicted 41% for Roosevelt and 54% for Landon in the final results of their presidential poll on October 31, 1936. The idea that Gallup could predict the outcome of the Literary Digests has been promoted by Gallup himself and his fellow pollsters and has become part of many works and textbooks looking at this episode (For an example see Freedman, Pisani and Purves, 2007, pp. 334–335). The reality is, however, that Gallup never predicted the

Literary Digest's poll, since his estimate was based on the premise that the Digest would be "conducting its poll at the present time" and not four months later. Gallup thus spun this publicity stunt that he was able to predict the Digest's forecast after the election but made it look like he had predicted the results already in July.[11] This episode played a crucial role in how the Gallup Poll emerged as an indispensable element in subsequent presidential elections and continues to stand as one of the foremost and influential organizations in the field of election polling.

The problem with the Literary Digest's poll was, as Katz and Cantril (1937) stated back then, that "[t]he selective error in the sampling technique of the Literary Digest was logically apparent long before it became empirically important. Merely because a method works fairly well on one or more occasions is no guarantee of its reliability" (Katz and Cantril, 1937, p. 176). Gallup could successfully capitalize on this failure of the Digest and rhetorically draw a boundary between the old pollsters' fragile and unsound methodology and his science-based way of conducting polls. Those remarks by Katz and Cantril, however, already foreshadow what later happened to Gallup himself, highlighting that rhetorical closure does not equate to widespread implementation in practice.

4.2.4 Cementing the Random Sampling Research Programme

12 years after his famous election prediction, however, during the 1948 presidential election, a major polling disaster happened to Gallup himself. Gallup, Roper and Crossley, who emerged as the "scientific pollsters" since 1936, predicted a victory for the Republican Thomas E. Dewey over the Democrat Harry S. Truman. During election night, Gallup was on ABC Radio when he started to realise that his forecasts did not quite match the known results so far. Gallup confidently predicted Dewey's victory by a margin of five percentage points in the final poll

---

[11] The view is widely shared in textbooks and historical literature on the development of polling. I also misrepresented this episode in my contribution to the Conference Proceedings of the STS Conference Graz 2022 (Griessl, 2022).

published before the election. Off air at 8 o'clock, as Dewey biographer Richard Norton Smith writes, Gallup "turns to his co-workers with the first distress signal of the day. 'Boys', he says, 'I think we're in trouble'" (Smith, 1982, p. 42). The result of the election was that Truman won with 303 electoral votes over Dewey with 189 and J. Strom Thurmond with 39. By this time, Gallup's election prediction enjoyed so much confidence – his poll was nationally and internationally very successful in assembling predictive stability – that the German Newspaper *Münchner Merkur* ran with the headline "Thomas E. Dewey America's New President". Due to the time difference between Germany and the US, the election results were to be expected after the editorial deadline, but Felix Buttersack, then-editor of the *Münchner Merkur,* aimed to be the first to print the news in his midday publication. The response Buttersack gave in the aftermath of this incident speaks in interesting ways to the "trust in numbers" (Porter, 1995) that was given to the pollsters by this time. Buttersack is quoted in a New York Post article by Ernest Leiser as follows: "All your American experts – your opinion poll takers, your political analysts—have been reporting a Dewey victory certain for months, so I thought it would be perfectly safe to go ahead as soon as the polls closed with the story I prepared in advance. How could I have known this would happen?" (Buttersack in: Leiser, 1948). This statement highlights in interesting ways how much trust pollsters were granted as an authority endowed with the power to speak about future states of the world. Buttersack was at risk of losing his licence, due to this violation of press ethics, as current publisher of the newspaper Dirk Ippen recalled in a talk at the Protestant Academy of Tutzing (Ippen, 2016) in 2016. What helped him after all was the more famous and iconic incident of the Chicago Tribune, which headlined "Dewey Defeats Truman" the day after the election.

Those two headlines stand for the enormous trust in numbers as an objective disinterested technique that was given to the predictions made by pollsters at this time. And people like Gallup played a major part in promoting this trust in their techniques, touting their own approaches as

science-based and upholding themselves as the sole legitimate spokespeople of public opinion. What would later become apparent was that the reason for Gallup's 1948 misprediction was that the interviewers were allowed to choose whom to interview, given certain quotas. In each of the fixed categories (including gender, age, and economic status), Republicans were apparently easier to reach. This episode did not lead to the demise of Gallup's polling company, as it happened to the Literary Digest, but the celebrated and apparently scientific method of quota sampling got under scrutiny. In paraphrasing Katz and Cantril (1937) point about the 1936 US presidential election, also here, the selective error in the sampling technique of the Gallup poll was logically apparent before it became empirically important.

The election miscall was met with strong accusations by the press. Stuart Chase, a writer for the New York Herald Tribune referred to the case as the "complete collapse of the oracular smug-masters who have pretended to test public opinion" (Chase, 1948). This situation was particularly difficult if considering the broader context of the time. The social sciences were, as Lusinchi shows, in a "post-war struggle to be accepted as legitimate scientific endeavours by the extant scientific community and by the politicians in Congress" (Lusinchi, 2018, p. 2). The disaster thus put pollsters in a legitimization crisis. They had to find a way out of it in order to become a recognized discipline among their neighbors from the natural science. Many natural scientists and politicians feared an overall politicization of science when giving social sciences the same recognition as the natural sciences. Besides other reasons, it was also the fear that social scientists were, as Lusinchi put it, "a bunch of busy-bodies involved in a normative endeavour and promoting possibly dangerous doctrines, such as socialism" (Lusinchi, 2018, p. 8).

One of the results of the investigations about what went wrong in the 1948 elections was a critique of quota sampling (cf. Likert, 1948b) and Gallup also drew this lesson from it and subsequently began to use random sampling as the basis for its surveys. There was already definite evidence about the superiority of random sampling, Neyman already provided the

knockdown argument in the sense of Pinch and Bijker's (1984) notion of rhetorical closure in the mid-1930s. Furthermore, it was already used in official statistics and yet, pollsters started to apply it only after the problems regarding quota sampling became apparent. This event can, however, be seen as the beginning of a consensus, not only within the scientific community but also among polling organizations, on the legitimacy of sampling and the superiority of random methods over purposive methods. This consensus on the superiority of probabilistic sampling methods became the characteristic element of survey research and the polling industry.

As we have seen in this section, the wide adoption of random sampling was not employed after 'rhetorical closure' had been achieved; rather, it happened when the *predictive stability* of other forms of sampling became fragile and when a stable network dissolved. New sampling methods were not widely adopted by pollsters after their superiority has been theoretically shown by survey methodologists; new sampling methods were instead adopted after shortcomings became evident in the aftermath of election forecasts. In the aftermath of the 1948 elections, new actors had to be mobilised in order to reassemble stable networks. It was the beginning of a successful story in which public opinion polling established itself as a legitimate producer of knowledge about public opinion. In this context, public opinion, viewed as the outcome of a specific technology, gained widespread acceptance not only as something measurable but also as something that is based on science and can be accurately measured.

4.3   The Co-Construction of Public Opinion and Sampling

4.3.1   Two Views of Public Opinion

Many notions and concepts we take for granted are ones on which there is no or little consensus as to what they actually mean. One of those is the notion of public opinion, which, as Herbst (2011) puts it, always struggles between being seen "as an aggregation of individual opinions" and "as a nonquantified but powerful conversation" (Herbst, 2011, p. 88). Probably the best

illustration of those two different understandings of public opinion has been formulated by Bavarian politician Franz Xaver Unertl (1911-1970), who once said: "I don't need any poll. When I want to know what the people are thinking, I go to the toilet during the intermission at a meeting and listen to what they say" (quoted in: Pack, 2022, p. 217). Obviously, Unertl's approach to tapping into public opinion did not become the standard when it comes to assessing the interest and the will of the public. It does, however, illuminate the difference between a concept of public opinion as a conversation between people and one as a conversation between individuals and a questionnaire. Whereas to Habermas (1999), for instance, public opinion is positioned in the salons and coffee shops, pollsters promoted the idea of polling and thus of public opinion as the statistical aggregation of individual opinions. This was especially throughout the first half of the 20th century, when pollsters, such as Gallup, Roper or Crossley promoted polling as the best way to understand what the people think. It was especially through the work of those early pioneers in polling, who worked hard to sell the techniques of opinion polling and survey sampling as the "greatest contribution to democracy since the introduction of the secret ballot" (Roper in: Igo, 2007, p. 121) and as a way to protect the common man or woman from the "tyranny of the majority" (Gallup and Rae, 1968, p. 268). Public opinion, for Gallup, is defined as "the aggregate of the views men hold regarding matters that affect or interest the community" (Gallup, 1957, p. 23) and pollsters are best placed to gauge it. Through the development of the random sample and its implantation in the practice of polling – the meeting point of the history of sampling and the history of polling – public opinion as a statistical aggregate, entered, as Osborne and Rose called it, "into the true" (Osborne and Rose, 1999, p. 367). Public opinion became real because there was a consensus that the sample survey is a stable way to measure it. Or, put differently, it became real because standards about how to quantify public opinion were established, contributing to a "conventional space of equivalence" (Desrosières, 2005, p. 6). In extending their endeavours to understand public opinion on more

and more issues through surveying, a view of public opinion as a statistical aggregate gained ground, a view that holds until now.

4.3.2   An Egalitarian View

4.3.2.1   The Preferred Public Repertoire

The perspective of public opinion as a statistical aggregate is built on a particular egalitarian view, inherent in the idea of the random sample. When studying public opinion, researchers might also analyse other forms of what Tilly terms public "repertoires", which are "alternative means of acting together on shared interests" (Tilly, 1983, p. 463). In the repertoire of most people to express their opinion are social activities, such as letter writing, protesting, donating, striking or participating in electoral campaigns. In defining public opinion as the aggregation of individual opinions, public opinion research, however, limits the study of public opinion to individual responses to predefined questions. This privileging of "what people say over what people do" (Krippendorff, 2005, p. 133) is rooted in an apparently egalitarian view implicit in the practice of random sampling. Randomisation for surveys and polls is based on a form of equality in the sense that, as noted, every person that belongs to the universe for which the results are to be generalised must have an equal or calculable probability of entering the sample. Public opinion as the result of polls thus points to an apparently fully egalitarian view since it also represents those who do not respond to polls. In this context, Gallup could claim that "in many situations - particularly those in which a substantial portion of the population fails to take the trouble to vote - the poll results might be even more accurate as a measure of public sentiment than the official returns" (Gallup, 1955, pp. 20–21).

4.3.2.2   Equality Before the Polls

I would like to exemplify this notion of equality by drawing on Sidney Verba (1932 – 2019). Verba was an American political scientist, known, besides other things, for his work on political

participation and inequality, including the well-cited book, with Gabriel Almond, "The Civic Culture: Political Attitudes and Democracy in Five Nations" (Almond and Verba, 1972 [1962]), in which they surveyed individuals from the United States, United Kingdom, Germany, Italy, and Mexico to gather information about political attitudes and democracy. In his 1995 presidential address before the American Political Science Association, Verba characterises the sample survey as a "major social science tool [,] […] a technology with an important influence on representative democracy" (Verba, 1996, p. 1). Verba contrasts surveys with other "repertoires" (Tilly, 1983) of political participation, and states that while those "ordinary modes of citizen activity […] allow quiescence […] [,] [s]urveys do not let people be quiescent; they chase them down and ask them questions" (Verba, 1996, p. 4). Surveys thus bring together social science technology, the political theory of representation and real issues in politics. They allow for equal responsiveness by governing elites towards the citizenry in the sense that this responsiveness requires "the capacity to provide equal consideration' and 'equal information about the needs and preferences of all citizens" (Verba, 1996, pp. 1–2). Verba acknowledges the inevitability that in any given constituency there will be active and quiescent people; however, he points out that "it makes a big difference whether the quiescence is due to preference or resources – to not wanting to act or to being unable to act" (Verba, 1996, p. 2). In this regard, the epistemology of survey sampling has a special character, as participation in a survey does not hinge on whether people have the resources or the motivation to do so: everybody participates in that their responses are calculated based on other responses, whilst the chance to provide a response is equal or known.

Despite polls never being perfectly representative, Verba argues that they, nevertheless, offer a better representation of the populace than almost all other modes of citizen activity. They achieve a "relatively unbiased view of the public by combining science and representativeness, indeed, by achieving representativeness through science" (Verba, 1996, p. 4). Without surveys,

political leaders would still "sway with the wind of opinion", "[t]he wind would just blow from different quarters, more likely from the better parts of town" (Verba, 1996, p. 6). It is in this sense that "[s]urveys produce just what democracy is supposed to produce – equal representation of all citizens" (Verba, 1996, p. 3). According to this perspective, the sample survey is fundamentally egalitarian as it provides every citizen with an equal opportunity to participate and ensures an equal voice when participating. In this understanding, polls become a crucial instrument for democracy as they allow for equal representation and responsiveness; everybody, it seems, is equal before the polls and represented equally through the polls. Opinions of those who do not participate are, as Hannah puts it in a discussion of the US census, "either estimated or 'imputed' (for incomplete returns) on the basis of the characteristics of neighbors" (Hannah, 2001, p. 526). In the words of political theorist Hannah Pitkin (1967), statistical generalisation through polls and surveys forms a "standing for" relation between a political leadership and its constituency. The selected sample thereby stands for the whole constituency and a political leadership can respond to the whole through a part.

4.3.2.3   Challenging the Egalitarian View

The egalitarian view, as I just exemplified in drawing on Verba, has also been challenged several times throughout the 20th century. Herbert Blumer, for instance, stated in his 1948 article *Public Opinion and Public Opinion Polling* that "polling gives an inaccurate and unrealistic picture of public opinion because of the failure to catch opinions as they are organised and as they operate in a functioning society" (Blumer, 1948, p. 457). In the case of polling for election forecasting, "a ballot cast by one individual has exactly the same weight as a ballot by another individual", which means that voters are a population in which each individual "has equal weight" (Blumer, 1948, p. 457). This changes, however, when moving from election forecasting to issue polling. The formation and expression of public opinion "is not an action of a population of disparate individuals having equal weight but is a function of a structured society, differentiated into a

network of different kinds of groups and individuals having differential weight and influence and occupying different strategic positions" (Blumer, 1948, p. 457). Treating everyone equal thus does not consider the actual power dynamics and social factors out in the real world.

This argument has also been echoed by Pierre Bourdieu, according to whom society consists of fields, in which people have different social, cultural, or economic capital, which is why people will have different backgrounds on which they form opinions. In his article, *Public Opinion Does Not Exist,* Bourdieu (1972) argues that public opinion "does not exist in the form which some people, whose existence depends on this illusion, would have us believe" (Bourdieu, 1972, p. 129). In order to arrive at this conclusion, he emphasizes three assumptions that underlie all opinion polls. First, on the basis of every opinion poll is the supposition that everyone can have an opinion, "that the production of an opinion is within everyone's range of possibility" (Bourdieu, 1972, p. 124). Second, it is assumed that "all opinions have the same value" (ibid.) and third, that if one asks the same question to different people that "there is a consensus about the problem" (ibid.), that everyone understands the questions in the same way.

The most important function of opinion polls is "to impose the illusion that a public opinion exists, and that it is simply the sum of a number of individual opinions" (Bourdieu, 1972, p. 125). Opinion polls are, however, much more an artefact, concealing the actual state of opinion as "a system of forces, of tensions, and that there is nothing more inadequate than a percentage to represent the state of opinion" (Bourdieu, 1972, p. 125). Bourdieu states that pollsters and survey researchers cannot expect that everyone has an equally well-informed opinion or an opinion at all. In the same sense as polling presupposes that everybody has the same or a known chance to enter a sample, polling presupposes that everybody is equally placed to form an opinion. Without considering the power of those who hold an opinion, it is not a strength of polls to treat all opinions equally. Distinct from election forecasts, the reality of the world makes it impossible to assume that a given opinion means the same for every participant.

4.3.2.4 Reifying Elite Opinion

Another argument that has been put forth against the egalitarian view is that rather than seeing polls as the aggregation of individual opinions, they are, as Herbst puts it in relation to Walter Lippman, "simply a projection of what political élites and journalists think about" (Herbst, 1992, p. 222). This brings up the question of whether the genuine interests of ordinary citizens corresponds with the issues that pollsters, journalists and those who commission polls are preoccupied with? The problematics that polling institutes put forth derive from a "specific kind of demand", meaning that one must ask "who can afford to pay for an opinion poll" (Bourdieu, 1972, p. 124). The problems posed by polls are political problems, which are "directly linked to the political preoccupations of the 'ruling power'" (Bourdieu, 1972, p. 125), and thus relate to particular interests. The 'dominant problematic', the questions that have been asked on an ongoing basis, is "the problematic which essentially interests the people who hold power and who consider themselves to be well informed about the means of organizing their political action" (Bourdieu, 1972, p. 127).

Continuing along this line of thought, opinion polling plays a crucial role in forging and constructing public opinion. Some of the interviews that I conducted also echoed this view that polls do not necessarily reflect the genuine concerns and interests of ordinary people but rather serve the demand to elevate certain issues to become matters of public concern. In one of my interviews, I asked a question regarding a poll published in the German Newspaper Münchner Merkur, which suggests that Markus Söder, the current Bavarian prime minister (CSU) would be the best candidate for the upcoming federal elections in 2021.[12] The response of Ulrich

---

[12]   This question was of crucial importance to Markus Söder, who, in the years prior to the election emphasized on several occasions that his place is in Bavaria and that he has no intention of becoming a candidate for the federal elections. Whilst it was always clear to many observers, he later declared his willingness to become candidate, the CDU/CSU, however, decided in favor of Armin Laschet, who should later lose against Olaf Scholz (SPD).

Rendtel was revealing. He stated, "of course, those who publish such a survey in a very public newspaper, have an interest in that this candidate is being discussed. And yes, this is the business of politics" (Rendtel, Interview, Translation). This illustrates how polls can be utilized to bring specific topics into the public debate and create them as issues of public concern. Furthermore, and already foreshadowing what I will discuss more in the final chapter of this dissertation, the mere conducting of a poll does not make a big difference. What is important is when, how and by whom those polls are disseminated, how they are implemented in networks in and through which those representations gain stability.

4.3.3 Discovered or Enacted?

Through the development of the representative sample, to which the early pioneers in polling played an important part, public opinion in the form of a statistical aggregate became a knowable and measurable phenomenon. A sample that resembles the population in important ways thus serves as an illustration in which the alleged whole can be seen; it apparently holds up the mirror through which the population can observe itself. What we can take from these considerations is that public opinion, as a statistical aggregate, does not exist independently of the statistical practices that are used to know and quantify it. This points to a performative view of social science methodologies, resonating with other scholarship from the last decades, showing an increasing interest in the social studies of social science and thereby in the way how social science methods play a role in creating the world they purport to describe (e.g., Grommé and Scheel, 2020; Law and Urry, 2004; Osborne and Rose, 1999). What this line of scholarship suggests is that in the same way as the natural sciences are capable of producing phenomena that previously did not exist, the social sciences are able to produce phenomena that did not pre-exist their own construction (Hacking, 1983). Public opinion becomes, so to speak, more or less equated with the aggregated response to polling questions. In this sense, methodological

decisions have implications on the results of polls and thus on the reliability of our estimates of public opinion.

A view holding on to see statistical facts as a pure and unfiltered representation of social reality cannot be maintained when following this line of thought. Instead, statistical facts are viewed as stemming from the interplay of conventions, actors, technologies and power. Through the application of different methods and conventions, statistical techniques bring different populations into being. It is especially in times of methodological disruptions, when certain methods are replaced by others, that the battle over the hegemony regarding the conventions of how to measure public opinion is fought. In upcoming sections, I explore this by drawing on current methodological controversies in the field, a debate that structurally resembles very much the methodological discussions over the use of purposive or random samples throughout the first half of the 20th century, as presented in this and the previous chapter.

What those developments brought about was not only the manifestation of random sampling as the preferred way of conducting surveys and polls, it also marks the development of a monopoly over public opinion in the hands of pollsters. This aspect concerns the particular position pollsters and polling companies have managed to occupy. Over the 20th-century, pollsters and polling companies have consolidated their authority to speak in the name of the public and public opinion. McGoey argues that "oracular power" encapsulates the power to monopolize perceptions of which groups are "trusted authorities" on knowledge and ignorance (McGoey, 2019, pp. 61, 64). Despite ongoing tensions over their methods, polling companies have come to dominate this very capacity to delineate what counts as legitimate knowledge about public opinion.

4.3.4   Monopolizing Public Opinion

4.3.4.1   Making Oneself Indispensable

As we have seen in the previous sections, throughout the first half of the 20th century, a view of public opinion as a statistical aggregate gained ground, a view of public opinion that still holds until today. In developing and defending the practice of polling as the dominant means to represent public opinion, the polling industry thus created a monopoly over the sovereignty to represent and define public opinion. Much of this happened through rhetoric forms of boundary work, especially in the way that Gallup promoted his approach to sampling as the sole legitimate way to measure public opinion. In his 1996 published *On Television*, Bourdieu argues that polls have reached such a dominance that they are seen to provide an unmediated relationship between voters and the political field that it "takes away from all self-styled spokemen [sic] and delegates the claim (made in the past by all the great newspaper editors) to a monopoly on legitimate expression of 'public opinion.' […] [and] deprives them of their ability to elaborate critically (and sometimes collectively, as in legislative assemblies) their constituents' actual or assumed will" (Bourdieu, 1998, p. 93). In his 1998 published social history of public opinion polling, *La fabrique de l'opinion: une histoire sociale des sondages*, French political scientist Loïc Blondiaux asks the question of how we came to accept the equivalence between public opinion and polling results and states that pollsters "have succeeded in measuring public opinion, i.e., in imprisoning it in their tables of figures"? He goes on stating that "[b]y getting more and more social groups to accept the validity of their figures, they have succeeded in bending reality and bringing it into their laboratory, but also in making the latter an obligatory passage point for anyone who, from now on, wants to know what is going on" (Blondiaux, 1998, p. 155, translation by author). As Blondiaux states in a footnote to this statement, this way of thinking directly derives from his engagement with various work in the sociology of science, besides others, the work of Michel Callon. Callon (1984) developed the notion of an obligatory passage point,

defined as successfully established passages, through which one needs to pass in order to reach a certain goal. An obligatory passage point constitutes a point in a network where different actors are enrolled so that their interests converge, which is essential for the network to function properly. This concept helps to explain how certain actors or elements can become central in a network, and how they can exert influence over other actors or elements in the network. The polling industry has thus managed to establish itself as an obligatory passage point that defines what counts as legitimate knowledge of public opinion. They have, so to speak, become the only legitimate cartographer of the social world and their opinions.

4.3.4.2 The Force of Polling Data

One can expand the argument by not only considering how pollsters emerged as an indispensable actor when it comes to the articulation of public opinion, but also looking into the power of public opinion, as gauged by the polls, itself incorporates. Whereas Pitkin differentiated between representation as standing for and representation as acting for, polling results seem to have become an authoritative force in its own right, encompassing both elements, as Ellwanger (2017) explores in suggesting the use of polling data to be understood as "a form of rhetorical argument". He shows that "polling data do not embody an objective empirical artefact, but rather a form of deliberative argumentation that is at odds with democratic life" (Ellwanger, 2017, p. 181). Ellwanger is thereby interested in prescriptive polls, designating polls that point towards what should be done, rather than towards what will happen. In thus exploring that "governmental obligation to capitulate to the will of the statistical majority […] as a rhetorical and political problematic" (Ellwanger, 2017, p. 182), he states that modern public opinion polling re-invented the Greek notion of *doxa,* the silent 'conventional wisdom' that is evident to everyone. In scrutinizing Gallup and Rae's (1968) notion that the public opinion poll allows the "public" to "speak" in its own voice, Ellwanger argues that "polling relies on a unique notion of *doxa* [as]

in order for public opinion to serve its proper democratic function, it *cannot* remain unstated or implicit" (Ellwanger, 2017, p. 186 italics in original).

The scientific study of public opinion is thus necessary for political leaders to "fulfil their obligations to enact the will of the people" (Ellwanger, 2017, p. 186). In this sense, modern polling re-defines *doxa* from a silent and implicit belief to one that must be expressed. It becomes imperative to undertake the task of rendering the "unseen facts" comprehensible to those responsible for making decisions. Ellwanger names four conditions (stability of opinion, predictability of results, prevalence of opinion and apathy on the part of participants), that determine the poll's validity, as perceived by both the public and their representatives, which, when in place, have the power to influence political processes. In this sense, prescriptive polling become a threat to the separation of powers in democracies, as they implicitly claim that the gap "between execution and representation, between governmental action and public opinion, must be closed" (Ellwanger, 2017, p. 194). If public opinion is to be followed by the political leadership and polling data is the sole measure of it, polling is, drawing on Rancière, "not only a 'science of opinion'" but is a "science immediately accomplished as opinion" (Rancière, 1999, p. 105).

4.3.4.3 Polling and Power

Pollsters brought to the fore this "immediate unity of science and opinion" (Rancière, 1999, p. 105), through which they fostered their monopoly. The ability to monopolise access to and representation of public opinion in the hands of pollsters and polling companies leads us to a discussion of power. The power to speak in the name of a public and their opinion becomes pivotal, not only because disciplinary controversies around the conventions of how to measure it has an important influence on the outcome, but also because it might open the doors for manipulation with potential effects on policy-making and democratic governance. In recent years, we can observe several well-known and lesser known instances including the tampering

of polls to boost former Austrian chancellor Sebastian Kurz and the Austrian People's Party's standing in the lead-up to the 2017 Austrian legislative election, as well as the fabrication of a poll purporting that American musician Kid Rock was the top choice for a senate seat in Michigan These cases, which will be thoroughly examined in the final chapter of this dissertation, illustrate that it is insufficient to only and primarily focus on the conventions employed in measuring public opinion; rather, when trying to understand the relation between public opinion and polls, it is crucial to consider the broader context in which these conventions operate. Notably, polling institutes remain the obligatory entry point in the creation and dissemination of public opinion.

4.4 Conclusion

Throughout this chapter, I explored the transition in polling methodologies from early straw polls to the Literary Digest's approach up to the adoption of quota and random sampling by George Gallup. An important aspect of this chapter was the assembling of trust and credibility towards the Literary Digest's and Gallup's election forecasts. As I argued, trust has been assembled and disassembled through *predictive (in)stability*, which both, the Digest and Gallup had laboriously created and embarrassingly lost. After the 1948 US presidential elections, however, the – at first only rhetorical - implementation of random sampling methods in opinion polling gave new credibility and stability to the enterprise of polling. Importantly, the adoption of random sampling was not solely driven by theoretical arguments but rather by the empirical shortcomings of other sampling approaches. The consensus on the superiority of probabilistic sampling methods became a defining characteristic of survey research and the polling industry, a development that is based on the apparent concept of equality underlying those approaches.

These developments not only solidified certain conventions in the field of polling but also contributed to the establishment of public opinion as a statistical aggregation of individual opinions. The introduction of random sampling helped to standardize and quantify public

opinion, transforming it into a tangible and measurable entity. Furthermore, through these developments, pollsters assumed a unique position as the legitimate spokespersons for public opinion, monopolizing access to and representation of public sentiment. The significance of this position is crucial, as the creation of public opinion is intricately tied to the infrastructure of polling companies, an issue that will be explored in subsequent chapters.

It is in this sense that I understand polling practices as a *forging* force, in that the results of public opinion polling are equated with public opinion itself and are thus "entirely caught in a structure of the visible" (Rancière, 1999, p. 103), in which measured opinion comes to be seen as true public opinion. Throughout following chapters, drawing in more depth on my interview data, I thus explore the question of the *what* and the *how* of public opinion, as gauged by the polls. The next chapter will be guided by the question of how we can think of the representation of public opinion through polls?

CHAPTER 5

5 The Representative Axis: Capturing Public Opinion

5.1 Introduction

The previous two chapters explored the history of sampling and the history of public opinion polling and election forecasting to the point that when pollsters started to adopt random sampling, they not only stabilized the definition of public opinion as the statistical aggregate of individual opinion but, in so doing, they also managed to gain a monopoly over the access to and the representation of public opinion. Interspersing an analysis of my interview data with a discussion of shifting understandings of representation in political theory, the aim of this chapter is to ask the more conceptual question of what it means to *represent* public opinion through polls and which ontological commitments towards public opinion are inherent in the understandings of representation. This chapter will outline two views on representation, which I term representation as mirroring and representation as performance. While the first adheres to the view that practices of representation bring a pre-existing entity to the fore, that representations stand in a mirroring relation to the represented, the second adheres to the view that practices of representation bring the represented into being, that they perform and enact the represented. After discussing different conceptual views on representation, I will suggest a typology of what it means to represent public opinion through polls. Therefore, I will introduce what I term the *standard view*, the *deliberative view* and the *performative view*. This step is important for the overall argument of this dissertation, which consists of three steps when studying the representation of public opinion through polls: legitimation, authorisation and forging.

This chapter will explore various reasons and effects of this development, before pointing out the different ontological commitments underlying the idea of public opinion as measured by

public opinion polls. In taking up an argument made by Sismondo and Chrisman (2001) on the nature of maps, I will explore how there is no single account of representation that fully grasps the nature of polls. While the 'standard view', characterised as a view that sees the results of polls as the real representation of public opinion, is prevalent among many accounts of polls and surveys, it fails to account for the formation of public opinion through deliberative and communal processes. It has to be noted, though, that while the standard view is the ideal-typical perspective one finds among survey researcher and pollster, many of my interviewees adopted a less certain position, acknowledging the performative capacity of polls and surveys, as well as the uncertainty regarding their results. One pollster, for example, compared polls to torch lights, of which one needs many to see what's going on. The 'deliberative view', on the other hand, locates public opinion in the collective, rather than the collection of individual responses. While this view can be seen as a response to sociological critique of polling as presented in the previous chapter (Blumer, 1948; Bourdieu, 1972), this view falls short in explaining what it takes to move from the mere measurement to the creation of public opinion as a result of polls. The 'performative view', according to which polling methods play a role in constituting publics, rather than merely describing them, will form the last of those three accounts. It thus asks the question of what needs to be mobilised in order to create stable representations of the social world. Rather than being a true reflection of the genuine views and opinions of the general public, opinion polls bring into being what they purport to describe. This idea is not incompatible with the standard view, or the notion that public opinion can be quantified by attaching numbers to measurable phenomena. It does, however, add an additional layer to it. It grounds the idea of public opinion in its material manifestations, making clear the real-world stakes of polls and their democratic role and effects. As it become apparent throughout this dissertation, the performative view is also the perspective I hold when theorising about polls and surveys. Drawing on my interviews, this chapter explores how practitioners themselves perceive the stakes of onto-

epistemological debates over the uses of polling in (mis)representing public will and public opinion. Whilst this chapter mainly discusses onto-epistemological perspectives towards polls and surveys, it will also become apparent that the value of those distinctions differs regarding to the socio-political context in which polls and surveys are conducted. This specifically pertains to the existence and nature of parallel forms of representation, like elections, and the extent of prior knowledge about the queried topics among respondents.

5.2 On Representation

5.2.1 Representation as Mirroring

5.2.1.1 Standing For the Represented

As presented in the Literature Review, the classical view of representation adheres to the notion that associates representation with responsiveness. Following this classical view, political leaders and representatives should be guided by the interests of the represented. This "bedrock norm", as Disch (2011) termed it, mainly relates back to the work of political theorist Hannah Pitkin and her seminal work on *The Concept of Representation* (Pitkin, 1967). This notion can be construed as the foundational conviction that within a democratic setting, representatives ought to ground their decisions in the preferences of citizens. According to Pitkin, the origin of this unidirectional understanding of representation, where it flows solely from the represented to the representative, rather than the reverse, can be traced back to the etymological roots of representation as re-presentation. This notion serves as a crucial link between the representation on one side and the represented fact or object on the other. It establishes a relationship wherein the representation can *stand for* the represented entity. Thus, it creates an asymmetrical dynamic where the represented entity cannot simultaneously stand for its representative or the representation itself. Political representatives ought to be driven by constituent demand, which means that the represented must be logically prior to the representant. This asymmetry brings to

light notions of power dynamics inherent within conceptual discussions surrounding representation.

A notable distinction arises when comparing this framework of representation to the fascist model, wherein the followers or the party merely reflect the will of a leader (see e.g., Barker, 1967, p. 377; see also: Pitkin, 1967, pp. 107–109). In this case, the concept of representation takes on a divergent character, since the will being represented, or rather incarnated, is one that the leader inspires and that would not exist without such inspiration. While in this case, a political leader might still represent the will of the people, the notion of representation is used in a different way, as political scientist Ernest Barker remarked: "The will which he represents, or rather incarnates, is a will which he inspires, and which would not exist without his inspiration. He represents a will projected from himself and reflected back upon himself." (Barker, 1967, p. 377; see also: Pitkin, 1967, pp. 107–109). This is what, according to Pitkin, distinguishes this model of representation from the fascist model, in which the follower or the party reflects the will of a leader.

Pitkin (1967) presents four perspectives on representation, namely formalistic, symbolic, descriptive, and substantive representation. Important for our understanding of polling is the third perspective designating how a representative resembles the represented. Pitkin's understanding of the concept of representation implies a process of "making present again". Representation goes beyond the mere act of bringing something into presence; it involves making something present in a certain sense, even if it is not physically or factually present.

Rather than merely bringing something into presence, representation is about "making present *in some sense* of something which is nevertheless *not* present literally or in fact" (Pitkin, 1967, pp. 8–9 italics in original). The *sense* in which something is to be made present encompasses various manifestations, within which Pitkin introduces a distinction between two perspectives: representation as *acting for* and representation as *standing for*. Acting for entails

situations where representation involves either authorization, where the representative is granted the authority to act, or accountability, where the representative is held responsible for their actions. Conversely, representation as standing for does not primarily revolve around action itself; instead, it focuses on the characteristics and attributes of the representatives. It centers on their being rather than their actions. If survey research engages in a mirroring relation, in that its aim is that a sample resembles the population in important ways, it can be said that this is the form of representation inherent in the practice of polling. In the case of public opinion polling, poll results mirror and reflect a popular sentiment that pre-exists its measurement.

This promise that polling techniques are able to represent and quantify a pre-existing will has existed since the beginning of polling and is being maintained and stabilised in today's perception of polls. Yudin (2020) traces this model of representation back to Rousseau's notion of the general will, as developed in *The Social Contract* (Rousseau, 2002), first published in 1762. The claim is that pollsters aim at finding a technique to measure and mirror a pre-existing general will. In the categorisation provided by Pitkin, one subset of representation as *standing for* is *descriptive representation,* defined as "the making present of something absent by resemblance or reflection, as in a mirror or in art" (Pitkin, 1967, p. 11). The making of a descriptive representation involves an understanding of representation in terms of "an accurate correspondence or resemblance to what it represents, by reflecting without distortion" (Pitkin, 1967, p. 60). In referring to Gallup's ideas, Pitkin states that they can be seen as a descriptive representation, aiming to establish a "mirroring" relation between the representative and the represented. Pitkin counts random sampling as the best approach, "no doubt because of the power of that technique in scientific research and because it is familiarly linked with representation in the idea of the representative sample" (Pitkin, 1967, p. 73). The pollsters' promise can thus be summarised in the way that a representative sample contains the target population and that possible disturbances can be quantified through confidence intervals.

5.2.1.2 Samples as Representation

This view also echoes views of representation found in the literature on survey research. In a four-part series in the International Statistical Review, Kruskal and Mosteller (1979a, 1979b, 1979c, 1980) extensively explore the notion of representativity and representative sampling and derive at different ways of how the concept is being used. It is being used as (1) a seal of approval, (2) the absence of selective forces, (3) a miniature or small replica of the population, (4) typical of the population, (5) heterogeneity and wide coverage, (6) a vague term that is then made precise, (7) a specific sampling method, (8) permitting good estimation and (9) good enough for a particular purpose. While we can say that the concept of a representative sample eludes an exact definition, the use of representativity aligns to the view of representations as mirrors. For instance, in the case of the dispute regarding the Civey poll published by Focus Online, the complainants stated that Civey's sampling approaches are "fundamentally not suitable for delivering 'representative' results". In his expert opinion for the German Press Council regarding the Civey dispute, Rendtel (2018, p. 2) also refers to Kruskal and Mosteller's work, stating that the complainants argument that representative surveys allow for statistically valid statements about the population coincides with the use as "permitting good estimation" in Kruskal and Mosteller's terminology. A representative survey should, therefore, enable a reliable estimation of the true opinions within the population. The promise of pollsters is to provide such an estimation, which political leaders and governments can then respond to.

5.2.1.3 The Pollster's Promise

This view that governments should be driven by constituent demand and that polling is a way to represent this demand has also been echoed by pioneers and practitioners of surveys and polls, as already elucidated in my presentation of the "scientific pollsters", Elmo Roper, George Gallup and Archibald Crossley. It was presented as a means to grant political agency to those voices that might otherwise not be heard. The point being made here has to do with the logic of survey

sampling. As we have seen throughout the previous chapters, the principle is that it must not depend on the individual and on whether they enter the sample, but that, for each person that belongs to the population for which poll results are to be generalised, the probability of the individual entering a sample must be equal or, if unequal, must still be calculable and accounted for. To give another example, polling pioneer Archibald Crossley stated in 1937 that "[s]cientific polling makes it possible within two or three days at moderate expense for the entire nation to work hand in hand with its legislative representatives, on laws which affect our daily lives". This, he states, "is the long-sought key to 'Government by the people'" (Crossley, 1937, p. 24,25). These claims included the argument that 'democracy's auxiliary ballot box', as Elmo Roper termed it, is even more democratic than voting, since it also includes the voices of those who do not vote (Roper and Woodward, 1948), echoing Gallup's argument that polling results might be more accurate than official returns. Polling, it seems, was and is viewed by many as an endeavour that gives everyone in the populace an equal say in matters of public concern, and that represents each and everyone equally.

This promise of pollsters to present the will of the people has also been raised in many of my interviews. During an interview with me, Jon Krosnick, Stanford University, for instance, raised the question of "[w]hy do we need to live in a world where we have surveys?" His argument is that "[h]uge numbers of decision makers will be handicapped if surveys disappear, because […] the unemployment rate and lots of other economic statistics that influence investing and decision making by businesses and government come from surveys [and that] if a democratic government wants to at least consider what the public wants it to do, what the public is willing to pay for, what the public approves of and doesn't approve of, the only way I know to find out is surveys" (Krosnick, Interview). In a similar vein, Manfred Güllner, founder and head of Forsa, stated that "[t]here are still politicians here who say that they don't think much of polls, and they don't understand that we report what people tell us" (Güllner, Interview, Translation).

5.2.2   Representation as Performance

5.2.2.1   From Representation to Presentation

Whereas the classical view posits that practices of representation make a pre-existing entity present again, a diverging view can be described as merely making present, without the precondition that the represented needs to exist prior to its representation, hinting at an understanding of representation that does not contain the represented, but rather brings it about. In *Surveying Public Opinion*, survey researchers Patrick Sturgis and Patten Smith, state that "[we] do not simply reveal a pre-existing public mood but, to some extent, we serve to create it as well" (Sturgis and Smith, 2010). This major shift in theories of representation has been termed the 'constructivist turn' (Young, 2000; Saward, 2006; Urbinati, 2006; Disch, Sande and Urbinati, 2019; Disch, 2021). According to this view, representatives play a role in creating the identities and claims of the represented and not the other way around. To Saward, the central aspect of representation is not that representatives represent and advance the interests of the represented, but the focus is on the conditions under which a representative makes representative claims: "[T]he represented play a role in choosing representatives, and representatives 'choose' their constituents in the sense of portraying them or framing them in particular, contestable ways" (Saward, 2006, pp. 301–302). Against classical accounts of representation, representation does not happen as the mere result of a particular process, it posits that there are 'makers'. Saward mentions spin doctors as a very obvious example, who define the claims about subjects and objects to create audiences for those claims. Saward's 'representative claim' is thus an example of the constructivist turn, highlighting the role of the representative in constructing the represented. This view has been considerably developed by Lisa Disch, to whom the constructivist turn is a turn away from the view of representation as response, or mandate representation, towards one of representation as mobilisation.

In contrast to the standard and hegemonic understanding of representation as "mandate representation" (Sintomer, 2013), in which power is delegated from the people to a representative government, the constructivist turn in political representation challenges this view, without endorsing the fascist model, as presented in the previous section. As Disch (2021, p. 4) outlines, the constructivist turn in political representation is based on two premises: First in arguing that political representation is not a mere reflection of constituencies, but plays a role in constituting them. And second in stating that political conflict is not a mere reflection of societal divisions but plays a role in forging them. Scholars in the tradition of the constructivist turn thus reject "the commonsense model of democratic representation as a mirror or 'transmission belt' for constituency interests, instead affirming the basic premise […] that a popular constituency comes to a political agent only through acts of representation" (Disch, 2021, p. 4). The critique has been formulated against the assumption underlying Pitkin's account that re-presentation is unidirectional and that the 're' is a temporal 're'. In drawing on Derrida (1973), Disch interrogates this view and states that "it might as well be an iterative 're', the 're' of repetition" (Disch, 2011, p. 109). The constructivist turn, to use the words of Lisa Disch, thus "violates the basic assumption that 'public opinion on policy matters' ought to give democratic theory 'its starting point'" (Disch, 2019, p. 4).

5.2.2.2 Creating Publics Through Polls

Several recent scholarly accounts on the practice of polling implicitly echo this view, as put forth by Disch. In this section, I will highlight such accounts, drawing on recent work on the far right and on how polls play a role in fostering it, as well as a case from Russia, in which polls are used as a way to not only create a people, but also to legitimize political actions concerning this people.

5.2.2.2.1   The Role of Polls in the Mainstreaming of Far-Right Ideas

Aurelien Mondon explores similar developments that resonate with the constructivist turn in theories of representation. Mondon (2022) shows how "rather than following 'what the people want', elite actors play an active part in shaping and constructing public opinion and legitimising reactionary politics" (Mondon, 2022, p. 1). In doing so, Mondon highlights the frequently overlooked unequal power dynamics. By examining the emergence of the far right, he contends that its rise should not be solely attributed to a straightforward 'democratic' demand, but should rather be analyzed as a top-down phenomenon. In his example of the far right, he illustrates that the utilization of public opinion in policy discussions is driven by political decisions that carry significant political consequences, despite presenting themselves as unbiased, scientifically grounded portrayals of "the people's" thoughts and desires. To be more concrete, Mondon states that shifts towards tightening borders are often justified "by the use and misuse of public opinion surveys that claim that immigration is among the top concerns of many if not most people" (Mondon, 2022, p. 2). Mondon is thus interested in the 'symbolic value' (Herbst, 1993) of polls, which "emerges *during* the public debate" (Herbst, 1993, p. 20 emphasis in original).

In his research, he shows how the perception of what is considered to be among the most important issues, as collected e.g., by the Eurobarometer, strongly differed when asking about issues concerning the country or one personally. Thus, immigration is usually among the top issue in the first instance, whereas it fares far lower when focusing on the personal situation. In showing these differences, he aims to "interrogate in which setting people appear to be concerned with immigration, which interests it servers, and what that can tell us in terms of mediation processes and democratic legitimisation of certain issues and in this case those that directly benefit the far right" (Mondon, 2022, p. 10). Analytically, Mondon thus moves from a linear model of public opinion, implying that "'the people' think something that is translated by opinion surveys and addressed by democratic elites" to a more circular model, in which the power of 'the

people' is much more limited "in shaping the agenda as public opinion is shaped to different extents by the content of public discourse, and also by polling companies that themselves play a part in shaping said discourse" (Mondon, 2022, p. 11). The second case that I will present is that of a people to legitimize, a case that leads us to contemporary Russia.

5.2.2.2.2   The Power of Polls: Manufacturing Political Reality

To draw on a different example, I would like to emphasise this point with briefly referring to a case from contemporary Russia. In *Governing Through Polls: Politics of Representation and Presidential Support in Putin's Russia,* political scientist and sociologist Greg Yudin (2020) explores the "specific conception of democratic representation behind polling practices", to argue that "opinion polls are capable of manufacturing the political reality they represent" (Yudin, 2020, p. 1). He demonstrates how Russian authorities employ polls as a substitute for referenda and as a means to validate election outcomes, thereby revealing the mechanism of representation that transforms polls into a powerful tool for governing, sustaining hegemony, and fostering depoliticization.

Beginning with the early years of the polling industry, when people like Gallup imagined a "plebiscitarian utopia" (Yudin, 2020: 2) in which polls would be conducted on a regular base to continuously access the popular will, Yudin states that polling became an institution of public government in contemporary Russia. Yudin argues that polling is a necessary partial and performative activity, making polls "the perfect anti-political machine, for they use the authority of scientific sampling to impose the representative sample as a total representation of the people and therefore suppress the political action aimed at subverting the hegemony" (Yudin, 2020: 4-5). In referencing to the case of the annexation of Crimea in 2014, Yudin shows how this annexation was justified by polling data in the sense that polls were used to assess the will of "the people of Crimea", a people that in fact did not exist "until the poll was conducted". Furthermore, instead of a referendum about the annexation of Crimea, Russia conducted a

"Mega-Poll", which after all, served as a replacement of a referenda. According to the official outcome, 91% agreed with the accession of Crimea and only 5% disagreed, a suitable outcome to legitimise the annexation. The poll was, however, conducted in secrecy, and their methodological details have not been made available. Consequently, it is impossible to verify the official outcome.

5.3 Representing as Mirroring or Performing

Following the discussion just presented, one might understand those views in the sense that polling constitutes a technique for *describing* reality or a technique to *perform* or *enact* realities. While the former, to borrow a definition from John Law, "works on the assumption that in one way or another reality has a definite form that is substantially independent of and prior to the tools used to inquire into it", the latter "treats knowledge practices as more or less *performative*" (Law, 2009, pp. 239–240 emphasis in original). In the following, I will further elaborate on those views, in presenting three types in which the relationship between polling, representation and public opinion can be understood. I will therefore review and present what I term the standard and the deliberative view as two representants of the view that polls describe or mirror an underlying reality and will end with an elaboration on the view that polls enact and perform realities, which I term the performative view. In this sense, I seek to include what might be termed the 'hinterland' (Law, 2004, 2009) or the 'matters-of-concern' (Latour, 2004, 2014) of polls, allowing for a broader analysis not only of what polls are but of what polls do and what is done to them. In this sense, polls not only describe, but they are also productive and their productivity stems from their embeddedness in networks of production, dissemination and consumption of polls. This perspective aligns with the Foucauldian approach to understanding the invention of population as a statistical entity, whereby statistical knowledge serves as a tool to render populations observable, comprehensible, and controllable (Foucault, 2009).

5.3.1 The Standard View

The first view, which I term the standard view, can generally be characterised as a view that holds the results of polls for the real representation of public opinion. This view aligns with the view articulated by early pollsters, such as Gallup. Some months before the 1948 election, which he famously miscalled, Gallup (1948) stated that public opinion polling methodology became highly successful and reliable.

> The reliability of methods now employed to gauge public opinion has been demonstrated time and again, not only in the United States but in a dozen different nations. Polls have met successfully the test which any scientific method must meet. They have proved equally reliable when applied in completely different circumstances and by different organizations. […] Modern poll procedures make it possible to conduct a nationwide referendum or plebiscite in a matter of hours, and to report results that would differ by only a few percentage points from the results which would be obtained if the entire voting population of a nation went to the polls (Gallup, 1948, p. IX; Gallup in: Bird and Merwin, 1971, pp. 20–21).

According to this view, polls are able to represent public opinion with a high degree of accuracy. This view aligns with the egalitarian view underlying polls, positing that even though not everybody goes to the vote, sampling still maintains the principle of 'one person, one vote'. Interestingly, Bird and Merwin (1951), who republished parts of Gallup's 1948 essay in which he made the just-mentioned proclamation, deemed it necessary to add that Gallup made this statement "just a few months before he and his fellow pollsters found themselves under terrific pressure because of their failure to forecast correctly the election of President Truman in November, 1948" (Bird and Merwin, 1971, p. 20).

This view that Gallup presented posits that the aggregation of individual responses constitutes the true representation of public opinion. It takes the sample at face value and as true representations of collective opinions, without the need to adjust those estimates, it (mis)takes the aggregated responses of a sample for the population. Ontologically, this view understands public opinion as the aggregation of reported private opinions within a sample. Epistemologically, this view assumes that "answers given reflect the authentic attitudes of the

respondents, or alternatively that instability in respondents' answers will average out in the aggregate" (Perrin and McFarland, 2011, p. 96). In contrast to this slightly naïve perspective, we may also differentiate a more sophisticated perspective, that acknowledges the ontological commitments, which, however, suggests adjusting the samples using statistical techniques in order to identify possible distortions and bias. It thus acknowledges problems that come with the interview mode, sampling or question wording, but seeks for "technical approaches to ascertain the contours of underlying public sentiment" (Perrin and McFarland, 2011, p. 97). Also Gallup's epistemological perspective shifted in the context of the 1948 US elections. In the aftermath of the 1948 election, Gallup made a slightly less confident statement during a symposium on the question of whether public opinion polls should make election forecasts. He stated that polls are always subject to probabilities and also prone to fail, but that pollsters will give their best and continue to improve their methods:

> With the same certainty that we know we can be right most of the time, we know that we will be wrong some of the time. It has to be that way. We live by the law of probabilities. We will do our best to improve our methods and to do better the next time. (Gallup in: Seymour *et al.*, 1949: 141)

Most polling, one might say, happens on the assumption of this perspective in that it acknowledges the necessity to apply adjustments, such as re-weighting or stratification. The rather naïve view was more prominent in the early straw polls, as presented in the previous chapter.

When contrasting different contexts in which polls and surveys are conducted, the applicability of the standard view might change considerably. Whether public opinion is deemed to exist independently of our knowledge of it depends on whether parallel modes of representation exist, such as elections, or whether a poll or a survey is the sole mode of representing and thus enacting an opinion. In the case of polling for election forecasts, the type of question and potential answers are usually known by the respondents, that is, it can be assumed

that they hold and are able to mobilise an opinion on that matter. In the case of issue polling, respondents might give a respond on issues they are themselves ignorant about or do not yet have a stable opinion. In this case, the standard view might mistake the measured opinion for an, often illusory, actual opinion. This does not mean, however, that in the case of election forecasting, polls are necessarily correct, but that the ontological assumptions that one holds towards public opinion are different from other contexts in which polls and surveys are being conducted.

What lies underneath the standard view is that interviewees are interviewed separate from each other, through which it becomes countable and comparable. It is due to this mathematical necessity that individuals' opinions cannot be measured as expressed in conversations, further affording the view of public opinion as the statistical aggregation of individual opinions. If we understand public opinion as such, in that private opinions are expressed to become public, then taking a representative cross-section of a population means that the people based on whom public opinion is measured have nothing in common with each other. Opinions are calculated based on formal similarity within the same virtual community. Krippendorff (2005) suggest that the priorisation of individual opinions might have ideological, cultural or economic explanation: *Ideological* in that it relates to the Western ideal of seeing humans as rational and autonomous; *cultural*, in that it relates to the democratic ideal of one citizen one vote and *economical* in that it is cheaper to interview individuals, rather than to study deliberations and conversations in the places in which public opinion is formed. Another view of public opinion that does justice to including conversations and deliberations as well as providing a solution to the question of how to measure those can be found in the deliberative view.

5.3.2   The Deliberative View

Whereas the standard view considers aggregated individual opinions to be a reflection of public opinion, a different practice of polling emerged as a critique of those very assumptions.

Deliberative polling, first described by Fishkin (1988), diverges from the standard view in that it does not consider individual responses to a poll to lead to real representations of public opinion, not even through methodological fixes. Instead, it suggests including deliberative aspects in that respondents are given time and energy to engage in and contemplate on the questions the poll asks. In this case, a sample is polled and then invited to participate in a gathering with other sampled people to discuss and contemplate on those issues. The same people are then polled again, so that the resulting changes can be seen as what the public would think if they were more informed and engaged in the subject matter. In doing so, one might say, deliberative polling acknowledges some of the central criticisms of public opinion polling (e.g., Blumer, 1948; Bourdieu, 1972; Ginsberg, 1989), as presented in the previous chapter. This view can also be seen as a response to early critique of whether average people are competent enough to express valuable views on matters of public concern (Lippmann, 1993). What Perrin and McFarland take from those considerations is that deliberative polling "constitutes an important ontological difference from traditional public opinion studies" in that they "may be successful in correcting for the skewed picture presented by traditional public opinion [polling]" (Perrin and McFarland, 2011, p. 98), centering on the idea that participants in deliberative polling become more interested and informed, leading to increased efficacy and a stronger sense of community. In this sense, deliberative polling addresses the challenge of measuring a collectivity and not only a collection, since it "locate[s] public opinion as a characteristic of collectivities" (Perrin and McFarland, 2011, p. 99).

5.3.3 The Performative View

Both the standard view and the deliberative view share the intuition that there is something out there that opinion polls can measure. Another way to think of the ontology of public opinion and its relation to polling is the performative approach, a view that aligns with the constructivist account of representation. As just elaborated on, this understanding of representation turns the

classic view of representation, as developed famously by Pitkin (1967) and with it the standard and the deliberative view on its head. It no longer posits that polls measure and represent a given public opinion; it argues that polls bring public opinion about, they play a role in constituting publics and do not just describe them. This view is not fundamentally at odds with the standard view, but it adds the layer of the social life of polls, the question of what polls do after they are taken and what is done to them. It does acknowledge polling data to be relevant, without claiming it to be overly authoritative or that it possesses only unidirectional representation. To explore this position more thoroughly and to expand the analysis to the landscape in which polling takes place, let's look at an example that I will explore in more depth later on in chapter 8. This case leads us to a widely reported incident involving former Austrian chancellor Sebastian Kurz and the Austrian People's Party (ÖVP). When looking at how the ÖVP did in the polls, it can be observed that the party has long lagged behind the Freedom Party of Austria (FPÖ) and the Social Democratic Party of Austria (SPÖ), until May 2017 when they overtook both parties in the polls.

Figure 1: Avopeas, CC0, via Wikimedia Commons
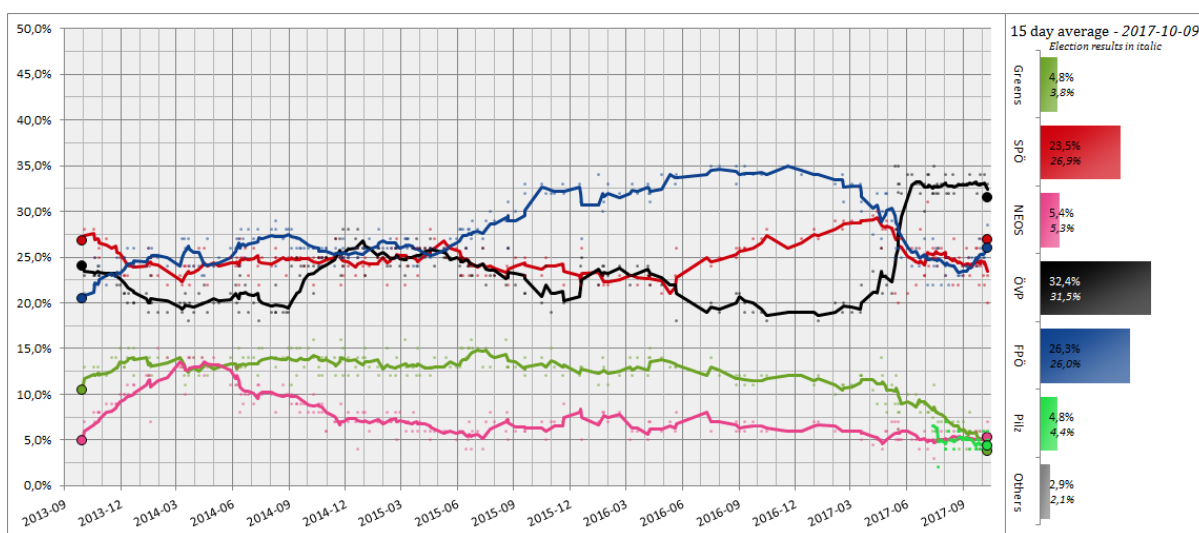
The black line in this figure of poll aggregates from 2013 until 2017 shows the jump that the ÖVP made in 2017, from a party that was mainly behind the other dominant parties, to the most successful one. Did the polls describe or enact the ÖVP as the leading party in the runup to the 2017 Austrian legislative election? Later on in this dissertation, I will explore the hinterland of

this shift in public support, pointing to a whole network of pollsters, embezzled public money, chats, newspapers, ads and politicians that brought this shift about, that, so to speak, performed it. In this sense, it is a case in which a public has been *forged* in the way that it became dominant and stable.

5.3.4   Notes on Different Logics of Representation

The previous three sections have outlined different onto-epistemological understandings of representations, the standard view, the deliberative view and the performative view. An additional notion that should be taken into account when studying the making of representations of public opinion through polls and surveys are the different logics inherent in the agendas of clients and distributors of those very reprsentations. In his article *Public Opinion Does Not Exist*, which I already drew upon in chapter 4.3, Bourdieu (1972) brings to the attention that the "problematics devised by the polling institutes are subordinated to a specific kind of demand" requiring one to "ask who can afford to pay for an opinion poll" (Bourdieu, 1972, p. 124). When people answer a series of survey questions, a public opinion emerges, no matter how important those issues were to the people answering, raising the question of whether there is "a correspondence between the common citizen's actual interests and the concerns of those in high places" (Herbst, 1992, p. 222). The existance of this correspondence is what Bourdieu doubts and should thus not be uncritically assumed. What Bourdieu is concerned about is that the problematics public opinion polls and surveys represent is usually the dominant problematic, the "problematic which essentially interests the people who hold power and who consider themselves to be well informed about the means of organizing their political action" (Bourdieu, 1972, p. 127). We can extend this line of argument to also consider the different logics and agendas behind the conduction and distribution of polls by the media. To highlight this, I want to revisit a response Urlich Rendtel gave me towards the publication of a certain poll, as discussed in chapter 4.3.2.4. When he told me that "those who publish such a survey in a very

public newspaper, have an interest in that this candidate is being discussed" (Rendtel, Interview, Translation), highlighting a differerent sense in which public opinion is being represented by different kinds of media. Highly commercialized media outlets, alongside social media platforms, exhibit a pronounced preference for content that is eye-catching, viral, provocative, and disruptive, diverging from the tendencies of traditional liberal mass media. This inclination suggests a heightened awareness and strategic engagement with their own performativity. Conversely, academic pollsters, as exemplified by numerous of my interviewees, who generally operate without the influence of commercial imperatives, may subscribe to a distinct ontological framework, one that prioritizes a different form of representation.

The next section draws on the accounts of those interviewees to explore the shift in polling and survey research towards online methods, a shift that has led to an enduring, highly fraught controversy – the stakes of which are apparent in my interviews, where strong emotions around the shift became clearer to me. I thus focus next on how my interviewees have reacted to changing conventions in the production, representation and dissemination of polls and surveys, raising new insights into the question of which groups are authorised to speak in the name of the public and their opinions. This section is focused on how experts feel about changes in survey research methods, specifically the transition from telephone interviews to online surveys, and the subsequent questioning of the dominant position of random sampling in recent times. Progressing, but still remaining within the general context provided in the previous historical chapters, the following chapter, alongside the remainder of this chapter, moves to the present time, commencing with an overview of historical and technological developments in the field of survey research and polling, with a particular emphasis on the period since the 1990s. This discussion will be followed by an exploration of the persistence of the random sampling paradigm and its gradual erosion, which became evident in the early 2000s – demonstrating a pattern of waning epistemological authority that is reminiscent of the pre-1934 era.

5.4   The Death and Resurrection of the Purposive Research Programme

5.4.1   The End of the Longue Durée of Probability Sampling?

Today, as battles between Civey and Forsa make clear, there is no doubt that public opinion research is currently in the midst of a significant transition. While some have argued that the current situation is comparable to the transition from face-to-face interviews to telephone interviews in the 1970s, others have argued that the prime days of survey research and polling are coming to an end. Savage and Burrows have pointed out in a 2007 article that while "in the years between about 1950 and 1990 sociologists could claim a series of distinctive methodological tools that allowed them to claim clear points of access to social relations [,] […] in the early 21st century social data is now so routinely gathered and disseminated, and in such myriad ways, that the role of sociologists in generating data is now unclear" (Savage and Burrows, 2007, p. 886). This refers to the period explored in chapter 4 in which a particular technique, the random sample, emerged as a reliable and trusted means to produce representations of the social world and public opinion. The monopoly pollsters possessed in reliance on the random sample, as one might put it, has become fragile. In a 2014 paper, the same authors state that the crisis in which the discipline finds itself at the moment is one "which does not unite experts in a quest to explore the potential of new modes of Big Data, but instead is likely to polarize and divide" (Burrows and Savage, 2014, p. 5). But what are the changes that occurred in this time and how can those polarisations be understood? These are the questions that guide the remainder of this and the subsequent chapter.

If one wants to understand the current situation in which survey researchers and pollsters find themselves in, it is important to understand the period between 1895 and 1948, in which the development and widespread adoption of sampling as a means to access populations gained ground and probability sampling became the preferred and superior approach in the discipline. This period, as we have seen, laid the ground for the widespread adoption of probability sampling

as the preferred way to conduct surveys and polls. A second crucial period is the time in which telephone surveys based on probability samples became the central hallmark of survey research and polling. This is often dated to the year 1974, when survey researcher Seymour Sudman (1928 – 2000) declared it possible to use telephones for health surveys. In less than a decade, as Dillman (2002) states, the "telephone had moved from being an unrecognized possibility to an accepted method of data collection" (Dillman, 2002, p. 475). Importantly, in adopting the telephone mode, researchers and companies were able to switch the *survey mode*, whilst maintaining the paradigm of probability sampling, the *survey design*. These two developments, the development of random sampling and the establishment of telephone interviews, were of crucial importance for the emergence of what has then become the industry standard, which is landline telephone interviews based on random digit dialling (RDD). Random digit dialling is a method of selecting survey participants by generating phone numbers at random and then calling those numbers in order to conduct an interview. The adoption of RDD was possible because of the wide distribution of telephones in the population and crucially because it substantially reduced costs compared to area sampling, which involves itinerant interviewers. In being able to balance profits and practices of knowledge creation, this has become the dominant mode of data collection for polls and surveys up until, as we will see in this chapter, it has been widely challenged by researchers and entrepreneurs, who aimed to tackle the problems and anomalies of this technique that became more and more apparent in the last decades. The solution that many suggested is in tapping into public opinion through the internet, using non-probability methods.

Despite the longstanding monopoly and authoritative status of surveys and polls relying on telephone interviews based on RDD, this mode began encountering increasing challenges in terms of rising non-response rates. While according to a report by Pew Research Center, their response rates were still 36% in 1997, they continued to decline and reached around 6% in 2018,

a trend that does not seem to be stopping (Kennedy and Hartig, 2019). This situation is widely acknowledged by members of the survey community.

> Yes, […] on the one hand it is of course true […] that the expenditures I have to make to reach someone are increasing. That has something to do with the mobility of the population. And if I used to be able to reach all the people via a landline telephone and a landline sample, that is no longer the case today. […] So, I think we have to take this discussion seriously and we also have to share the unpleasant truth that the effort is increasing and the costs are rising in order to process a reasonable probability sample (Interview, Thierhoff, Translation).

Response rates in the single digits are the new normal in the field, which resulted in elevated costs and raised concerns regarding potential sample biases. Consequently, the once unchallenged monopoly of random sampling in accurately measuring and representing public opinion faced significant limitations and drawbacks in the face of these mounting issues. In a recent poll conducted by The New York Times, Nate Cohn (2022), who is its chief political analyst, emphasized that completing a single interview required a staggering two hours of continuous dialling. Among various reasons for the decline in participation is that in contrast to previous times, individuals now perceive being interviewed via telephone as an intrusion rather than an honour. These changing values and perceptions have nicely been illustrated by Savage (2010), who presents fieldwork diaries from a 1962 survey conducted in Manchester. Savage describes this as a time in which people felt it was a "badge of honour to be specifically chosen for interview" (Savage, 2010, p. 1). What played a role in this development is that certain telephone service providers actively block calls originating from polling companies, further exacerbating the challenges faced by survey researchers. Furthermore, the growing mobility of many citizens presents significant obstacles to conducting sample surveys, as reaching these individuals via landline or postal methods has become increasingly difficult. These factors collectively contribute to the increasing difficulties encountered in obtaining representative samples, leading many practitioners and researchers to look for new ways to conduct polls.

5.4.2 Externalising the Decline

These considerations assume, one might argue, that surveying and polling practices contain an "implicit geography" (Legg, 2005) in the sense that they presuppose a particular ordering of the world through which surveys and interviews can be conducted. In the case of face-to-face appearances, questionnaires in newspapers or mails, as well as landline phones, interviews were bound to the home as the primary locus. All surveys and polls, be they face-to-face, telephone or web-based, assume a particular ordering of the world in and through which epistemic practices of representation take place. The current context is one in which it has become particularly difficult for researchers and pollsters to study public opinion by calling people at home, since many people no longer use landline phones or do not answer calls from unknown numbers.

Hence, the reasons for the challenges that probability sampling faces are located in changing conditions, in a changing geography, socio-technical infrastructures and habits of communication, external to the field of survey research and polling. External circumstances are discursively made responsible, not the theory of probability sampling itself, as many of my interviewees indicated.

> But now one thing remains, namely the people as such, they do not change. And the laws of statistics don't change either. So, on the one hand, it is true that this sector is of course in a state of upheaval because communication habits are changing massively and because people's mobility is changing. But the basic rules of empiricism and what Lazarsfeld, for example, defined as the basic rules of interviewing are [not]. And of course, it makes a difference whether I choose an approach that selects the interviewees randomly or whether I work with self-recruited people who have volunteered (Thierhoff, Interview, Translation).

Schnell also points in a similar direction arguing that as long as the conditions, underlying probability sampling, are met, there is a guarantee that the results are correct.

> You can prove the central limit theorem [laughs]. […] You can prove it mathematically. You can prove it empirically. It works. The central limit theorem works. That is, if the conditions are met, the result is certain. [...] That is, given the premise, it works, for sure. There is no mathematician who doubts the central limit theorem. There is also no statistician who doubts

the central limit theorem. And the question is, have the conditions of application changed? That is something else. A method only works as long as the preconditions are given under which one can arrive at a certain result (Schnell, Interview, Translation).

The central limit theorem posits that when random selections are repeatedly made, the resulting summary statistics will resemble a normal distribution. As the sample size increases, the variance of estimates decreases. This principle aligns with the law of large numbers, which asserts that with each successive random selection from a population, the observed characteristics of sampled units will progressively converge towards the characteristics of the entire population. Through the disappearance of landline phones throughout the last decades and an increasing aversion to participating in telephone interviews, the dominant mode of data collection became fragile. It does not mean, however, that the principles are no longer in place. While Schnell stated that his own way of dealing with those changing circumstances is to focus on databases, rather than surveys, others see the rise of the internet to offer a solution to the emerging anomalies of what is seen as the normal science. As a result, non-probabilistic sampling methods were increasingly brought back into the discussion and suggested as a serious alternative. It was, however, not that random sampling or the theory of probability lost its power as a tool or a principle, rather, the practical circumstances changed, leading to a new competition between two research programmes. Whereas, for instance, in the case of Lakatos' (2015) description of Euler's theorem, the world of mathematics is considered to be stable, but the theorem is being confronted by counterexamples; in the case of probability sampling, the world in which the theory is applied to changes. Because the world changes, research programmes lose strength. Thus, while pollsters continue in their efforts to know public opinion, they do not wish to know it at any cost. It has thus become necessary to increase efforts to reaching mobile phones, to use statistical tools to adjust for biases and to tap into public opinion through the internet. While the paradigm of probability sampling remains theoretically justifiable and valid, changing circumstances seem to push it towards the periphery of contemporary practice, leading to the

(re-)emergence of other practices. Many pollsters and survey researcher, however, aim to make their samples look like what an ideal probability process would have led to, showing how its primacy and theoretical superiority is still in place.

5.4.3 Technical Shifts

Since the mid-1990s, the rise of the internet appeared to offer a viable solution to many researchers in the field. It presented novel opportunities to efficiently acquire sizable and cost-effective samples of individuals and to ask them questions. This development fostered optimism among researchers, as they envisioned leveraging the internet's potential to overcome the challenges associated with traditional survey methods. The development of the internet opened new possibilities and brought about a new mode of data collection. As stated by Bethlehem (2018, pp. 27–29), computer-assisted web interviewing (CAWI) particularly gained prominence from 1995 onwards, when HTML 2.0 became available. One of the new functions was the possibility of data entry forms on a computer screen, differing from HTML 1.0 and e-mails in the sense that it allowed for the transmission of data from the computer of a potential interview partner to the server of the researchers. Whilst e-mail polls were already experimented with in 1983 (Sproull and Kiesler, 1986), the year 1995 and the introduction of HTML 2.0. marks the beginning of web polls and online sampling methods.

There are many reasons as to why using the internet became so prevalent in this time (Comley in: Hamersveld, Bont and Hamersveld, 2007, p. 402; Callegaro *et al.*, 2014, p. 5pp). Three reasons are of particular importance. First, in the United States, research buyers had already been utilizing panels that used paper questionnaires that were completed and mailed in. The transition to online thus did not mark any great challenges and was smoothly accepted. Second, the advent of the dot-com boom, which was a period of rapid growth in the technology industry, specifically regarding the Internet, from the late 1990s to the early 2000s, witnessed the establishment of prominent online panels in the United States, coinciding with a period of significant investments

in online ventures. Third, research buyers were particularly interested in reducing the cost of data collection. Thus, the combination of pre-existing familiarity with paper-based panels, the supportive climate of the dot-com boom, and the cost advantages of online data collection, alongside decreasing response rates and increasing costs and turnaround time, created a favorable context for the successful introduction of online research methods.

New technological developments have thus made it very easy and cheap to collect samples via surveys conducted online. Online surveys are also often remunerated and are generally more attractive to participants. According to a report by ESOMAR (2017), the turnover from conducting research using online methodologies rose from 26% in 2011 to 44% in 2016, an upward trend that is ongoing.

What distinguishes traditional methods from those using the internet is that in the case of addressing people through mail or landline, there exists a list of the people one wants to further know, from which researchers and pollsters can draw a sample. When aiming to reach people through the internet, no such list exists. There is no list of all registered email addresses or internet users, which is, however, a prerequisite when aiming to draw on random processes. Survey researchers are thus faced with two options: Embracing the unknown terrain and try to find ways beyond the paradigm of probability sampling or finding ways to apply probability sampling methods online.

5.4.4 Playing by the Rules

Due to the lack of a sampling frame and difficulties of maintaining a probability sample, early adopters of the internet and new technological means as a way to recruit participants aimed to remain within and maintain the paradigm of probability sampling whilst moving to online methods. Furthermore, in order to better understand the conflicts and controversies around online samples, which I will present throughout this and the subsequent chapters, it is important to

acknowledge that online methodology does not necessarily equal non-probability methods, as Cornesse emphasised in my interview with her:

> But of course, one must not make the mistake of equating non-probability [methods] with the online survey mode. For example, with the German Internet Panel we also have the online survey mode, but a probability sample, and in the same way, for example, laboratory experiments with students are offline, but still non-probabilistic samples. So, I think you also have to differentiate, which is often not done in the debate, between the survey mode and the sample design, and I think that is quite essential, because then it is often said that the non-probability samples will be more valuable in the future because more and more people are using the Internet, and that is just not right. That is true for the survey mode, but not for the sample design (Cornesse, Interview, Translation).

And indeed, when online methodologies were first developed, the move towards non-probability methods was not the first choice. As I will show in this section, many early adopters aimed to follow the rules underlying normal science, that is, probability sampling.[13] Many of those early attempts to use the Internet for surveys were done through what has been termed a "pre-recruited panel of Internet users", or a "pre-recruited panel of full population". The difference lies in the fact that in the former, the sampling frame consists of all internet users and in the latter of the whole population (see Couper, 2000). Due to the low coverage of the internet at this time, a pre-recruited panel of full population was the better methodological approach. In this sense, randomly selected respondents are provided with the necessary technology to participate in the surveys and become members of an access panel. Access panels are a group of people who have agreed to participate repeatedly in surveys.

One of the first to conduct such an approach was the Dutch Telepanel developed by Willem Saris. Willem Saris, a Dutch sociologist and Emeritus Professor of Statistics and Methodology,

---

[13]   In using the Kuhnian concept of "normal science" to describe probability sampling, it's important to clarify that this characterization primarily reflects the perspective of actors within the field. Numerous practitioners regard probability sampling as embodying the principles of normal science in survey research and polling. Therefore, the attribution of this term should not be misconstrued as emanating from my individual analysis but rather as an interpretation that emerges from my understanding of collective viewpoint of the field.

has made notable contributions to his field, focusing particularly on measurement errors and survey quality. The idea of the Telepanel was to maintain random processes when using those new technologies (Saris and De Pijper, 1986). During the years 1990 to 1995, he served as the director of the Telepanel, a market research company in Amsterdam that was established using university funds. The Telepanel got its name "because it used the telephone for the communications and the television as the computer 'monitor'" (Saris, 1998a, p. 410) and set out when the first modems entered the market in 1985. Thus, even during the mid- and late-1980s when the internet was still relatively unknown outside of specialist circles, Saris employed innovative techniques by leveraging the capabilities of connecting TV receivers and modems to create a networked system linked to a server. This pioneering approach, coupled with competitive pricing, resulted in Telepanel's considerable success.[14] (Schmitt, 2004, p. 284) The Dutch Telepanel started with 1,000 households, which were randomly selected and provided with the required technology, as well as modems and, if necessary, a telephone connection. Through this system, a new questionnaire was downloaded to the PC every weekend and after the survey was conducted, the central computer was dialled, and the data was uploaded. The system was later bought by the Dutch Gallup organization to become the first nationwide computer-based panel in 1986 (Hays, Liu and Kapteyn, 2015, p. 685).

Through this procedure, Saris and others were able to maintain the sample design of random sampling whilst changing the sampling mode through which people were interviewed. This "Interviewing without Interviewers" (Saris, 1998a) was a successful attempt to follow the rules of normal science, whilst shifting the technological apparatus through which surveys were conducted. As I demonstrated earlier, the problems that challenged random sampling whilst

---

[14] Due to its low prices and its new method, the Telepanel was very successful. Telepanel's activity were terminated by a court due to unfair competition, as they utilized university infrastructure to offer lower prices than their competitor. (Schmitt, 2004, p. 284)

being external to the theory itself, were not completely out of the scope of pollsters and survey researcher, as they were able to extend the paradigm of probability sampling through technological means. Probability sampling appears as a "progressive research programme" (Lakatos, 1978), in that it succeeds in increasing its predictive power in new circumstances.

Similar shifts already happened earlier when companies and practitioners changed, for instance, from face-to-face to telephone interviews. Whilst the adaption of the telephone raised controversy, the telephone and with it random-digit-dialling made it possible to extend the paradigm of random sampling through different technological means. Changing the geographical context, it can be observed that other players in the field adopted similar strategies. Manfred Güllner, founder and CEO of the German Market and Opinion Research Company Forsa told me.

> Initially, we saw very well and […] very early on that there are definitely advantages to using the internet as a survey instrument. We set up Germany's first online panel back in the 1990s. One of the advantages of the online survey was that we could show pictures, which you can't do with the telephone. […] And we then set up the first panel with online Internet users of 6,000, relatively costly, but the promised projects did not come. That is to say, we discontinued it. But then we started again by building up a panel. At first, successively from about 100 to 1,000 to 10,000, we worked with an auxiliary construction at a time when all the transmission paths were not yet there as they are today, namely by using the television screen to display questions on and show pictures. We selected households that got a set-top box that was connected between the telephone and the television, through which we could send the questions. [...] The basic idea, however, was still that we had selected households that were then, as was so classical, a reflection of the general population. [...] That was really a sample representative of the population. [...] The problem was that the acceptance was not yet great enough, it was a similar situation as in the mid-80s with the telephone (Güllner, Interview, Translation).

The development of Forsa's panel forsa.omninet thus also plays a crucial role in these developments and shows how pioneers were able to use new technologies in the development and execution of surveys. The name omninet stems from it being an omnibus survey, a method through which data is gathered on behalf of different organisations, who can add their questions. In some ways similar to the Telepanel, questionnaires were presented on the panel participants'

television screens through a transmission box. This device independently connects the television to the telephone network and was able to access the internet using a toll-free number and to establish a connection with the server. Once the questionnaire is created, programmed, and the survey period and sampling are determined, it is made available online. The chosen panel participants then receive a message on their TV screens, prompting them to complete the current questionnaire using an infrared keyboard. Bonus points are awarded to the panel participants for each fully filled questionnaire, which are paid out twice a year (Schmitt, 2004, p. 284).

As those examples show, survey research has undergone various shifts since the 1950s in terms of survey mode, but the survey design was maintained throughout those shifts. Pollsters and researchers thus adhered to the rules of what was considered normal science. There are more examples of early adopters, including the US company InterSurvey, which will be subject to a deeper analysis in the next chapter, set up with the aim to use the internet with the prerequisite of following random processes. The great split within the community into followers of the probability and those following non-probability approaches to sampling happened after the time when people like Saris developed their techniques and infrastructures to maintain probability-based approaches; the question thus is, how and when the paradigm of probability sampling started to disintegrate?

I asked Willem Saris about what has changed in the survey world since he launched the Telepanel. Saris told me that not much of the general procedure has changed, except for one thing, which he rather passingly mentioned:

> Well, what they do is, um, I think there is not so much difference. I think the whole thing is more or less the same, but the cost for us was much higher than for them because nowadays you have it on your computer automatically and most people have Internet, so at least in the western world, I think it's now very easy. […] But in principle, I think, there is not much change. Actually, there's one difference with the general procedure nowadays, that is that we tried to get a random sample of the population. We wanted it to be representative and that is nowadays not the case. So, they are in general, with few exceptions, they are just collecting data from people who are willing to participate (Saris, Interview).

The three differences Saris mentions are decreasing costs, decreasing efforts and a decreasing appreciation for random samples. While the first two aspects seem to be rather gradual changes, they can all be considered to designate paradigmatic shifts in the way polls and surveys operate, turned to in more depth in my next chapter.

5.5 Conclusion

This and the previous chapter have laid a foundation for understanding controversial changes in public opinion research, exploring how survey methodologists themselves have responded to those changes. Throughout this chapter, I explored different notions of representation and how they relate to the practice of polling. The majority of polling is done under the assumption of polling as a mirror, established through a sample that resembles the population in important ways. It now moves away from direct resemblance, emphasizing instead the process of crafting similarity through adjustments. This next chapter expands upon that focus through a closer engagement with the lived experience of practitioners, as well as an analysis of the theoretical implications of those experiences.

CHAPTER 6

6 The Making of a New Paradigm

6.1 A Revived Disagreement over Methods

At least since the aftermath of the 1948 US presidential election, the principle of randomisation has been the gold standard in various stages of survey administration, including face-to-face, mail, and telephone interviews. It established probability sampling as an unobjectionable convention in the field of survey research and polling. And yet, today, thanks to a shift that began in the 1980s and 90s, non-probability sampling is back. This is due to the capitalist logic within which they function: "With the rise of the internet in the late 20th century, however, nonprobability sampling rose to popularity again as a fast and cheap method for recruiting online panels" (Cornesse *et al.*, 2020, p. 3). Furthermore, push-to-web surveys, understood as a form of "data collection that uses mail contact to request responses over the Internet, while withholding alternative answering modes until later in the implementation process" (Dillman, 2017, p. 3) is more and more becoming a standard since most people have access to the internet (Lynn, 2020).

This chapter continues with a presentation of empirical material to illustrate how practitioners in the field have either embraced or resisted the evolving epistemic assumptions associated with the shift to online surveys. A crucial element of this will be an analysis of two narratives of conversion, wherein practitioners engage in critical self-reflection and begin to question their underlying epistemic foundations. These accounts highlight the transformative experiences of individuals within the field of survey research and polling, as they grapple with the evolving landscape and re-evaluate their previously held assumptions. The chapter provides a narrative of the increasing popularity of non-probability sampling, drawing on interviews with prominent

figures in the field of survey sampling and polling. These experts shared their perspectives on the development of this method, either by recalling their own contributions or presenting their views on how it came to be.

6.2 A former 'closure' opens again

This newfound re-emergence of non-probability samples and online panels has been periodised by Callegaro *et al.* (2014), who differentiate between two periods. They consider the time between the mid-1990s and 2005 as a period of "explosive growth of online panels" (Callegaro *et al.*, 2014, p. 5), especially in the US and in Europe and in the context of the dot-com boom. During this time, a large number of start-ups were established, many of which were focused on leveraging the newly emerging Internet technology. The second period[15] has been termed a "period of consolidation" (Callegaro *et al.*, 2014, p. 5), which was driven by two mutually reinforcing forces: The requirement to construct larger panels to cater the increasing demand for online survey participants more efficiently and due to an ongoing globalisation of the field of market research.

This time not only witnessed the emergence of online polling but also a significant paradigm shift in the way respondents were selected and the methodology for identifying representative samples. In other words, it deals with the fundamental question of how the epistemological grid of sampling should be constructed. This period, like previous sampling episodes, was marked by waves of controversy and consensus. Specifically, we can frame it as an initial period when the dominant paradigm of random sampling was rigorously adhered to, followed by a period of challenge, and subsequently, a period of defence and controversy. The subsequent sections, along with the next chapter, focus on the period of survey research spanning from 1990 to the present

---

[15] Callegaro et al. (2014) do not provide a concrete timeframe for the second period.

day, a period in which, as Groves puts it, "survey research methods adapted to changes in society and exploited new technologies when they proved valuable to the field" (Groves, 2011, p. 861).

Couper (2013) puts those considerations into context and states that there are three developments underway that significantly challenge the field of survey research and polling. (1) The emergence of Big Data, which has brought about the ability to collect information on a vast scale, encompassing entire populations rather than just a small sample, (2) the rise of opt-in panels, offering a fast and cost-effective way to gather data and (3) the rise of do-it-yourself survey tools (e.g., SurveyMonkey), giving everyone the means to conduct surveys and polls by themselves. Opt-in panels and do-it-yourself survey tools challenge the field of polling and survey research in fundamental ways. Opt-in panels designate a mode of surveying populations, in which participants are not chosen to participate, but where interviewees recruit themselves.

Those developments changed the way publics and their opinions are represented, challenging especially the authority of those who were traditionally endowed to produce legitimate representations, resulting in polarizations and divisions that occurred and that are still very much prevalent within the community. This "bifurcation in the survey industry" (Couper, 2000, p. 466) seems to have separated the field into, on the one hand, high-quality surveys characterized by probability samples and, on the other hand, data collection methods prioritizing low-cost and quick completion over representativeness and accuracy.

Important to note, however, probability sampling never dissipated, particularly in high-quality surveys and situations where budgetary constraints are less critical, such as in governmental surveys. In today's polling world, there is rather a co-existence of two different research programmes. This is reminiscent of the rhetorical closure by Neyman, after which both approaches became competing norms. Some practitioners are far more comfortable with that co-existence than others, as my interviews make clear.

6.3 Stories of Conversion

To better understand the history of this change, I want to draw on biographical aspects of Douglas Rivers' and Jon Krosnick's scientific and entrepreneurial journeys, which will give us interesting insights into the nature of the shift that occurred in survey sampling towards the end of the 20th and the beginning of the 21st century. Douglas Rivers is the CEO of YouGov and professor at Stanford University. Jon Krosnick is also a professor at Stanford and is often described as an antagonist to Rivers. In both of their narratives, we can observe stories of conversion told against the backdrop of these changes, which, as we shall see now, had different starting and endpoints.

I took inspiration for the analysis of interview material as stories of conversion from Ulmer (1988), who developed a framework for analysing such narratives. Whilst Ulmer is concerned with autobiographic narrations of religious conversions, the narratives I present in the following structurally correspond to those in important ways. The religious connotations, however, are implicit in the interview material. Rivers, for instance, called Krosnick and Langer's defence of the probability paradigm a "crusade" (Rivers, Interview) and others, such as Schnell attributed notions such as "believing" and "faith" (Schnell, Interview, Translation) to proponents of the non-probability paradigm. In addition, the structure of the following narratives resembles Ulmer's framework in that narratives of conversion follow a tripartite structure, which first presents a pre-conversion biography, then the actual conversion event and finally a post-conversion phase of life. In all conversion narratives, he shows how there is a biographical crisis presented as the occasion for conversion, a notion that is also visible in the narratives within the field of polling and survey research. Furthermore, while the pre- and post-conversion biography takes place in the realm of external reality, the conversion itself takes place in the inner world of the convert. In both narratives that I will present in the following, we encounter moments in which an inner realization took place that their way of thinking was incorrect, leading the path to a changed perception of the external reality.

A second source of inspiration comes from Kuhn, according to whom the existence of those narratives is a normal part of the nature of scientific research: "The transfer of allegiance from paradigm to paradigm is a conversion experience that cannot be forced. Lifelong resistance, particularly from those whose productive careers have committed them to an older tradition of normal science, is not a violation of scientific standards but an index to the nature of scientific research itself" (Kuhn, 1996, p. 151). In the following, I will present two such narrations, which involve a period prior to the moment of conversion, a moment of conversion in which both protagonists refer to a trusted figure as those who put the conversion in motion and a time after the conversion. Both narratives started at a similar time and are developed in reference to each other. Both narratives are told in dialogue with statements made by other participants in the field.

6.3.1 Challenging the Dominant Programme

6.3.1.1 An Unconventional, Conventional Start

One company that followed the path set out by the developments around the Telepanel was InterSurvey, which would later be renamed to Knowledge Networks, founded by Douglas Rivers and the SPSS co-founder Norman Nie (1943 – 2015) in 1998. Just as Willem Saris, Rivers and Nie considered it to be of great importance to maintain the paradigm of random sampling. Rivers told me that they "thought about how we could do it, using conventional sampling methods" (Rivers, Interview), which means that they wanted to work within the framework of normal science. Rivers told me that in the year InterSurvey was founded, they became aware of the Telepanel subsequently inviting William Saris to Palo Alto to consult with Rivers and Nie on how to further develop the company. As Rivers told me, they also invited Mick Couper to consult with them.

The development of InterSurvey, following some of the general tenets of the Telepanel was taken up positively. To Jon Krosnick, this "was a great idea" (Krosnick, Interview), as they were able to maintain their general approach to sampling, based on the random sampling paradigm.

The main difference between InterSurvey and the Telepanel was that the Telepanel used dedicated hardware and dial-up access – it was not really an Internet survey. The approach of the Telepanel to send interviewers to respondent's homes and to install the necessary hardware was, as Rivers told me, economically not feasible in the U.S. Norman Nie had the idea to use commodity hardware (WebTV) which was cheaper and more user-friendly. In a conference paper to the American Statistical Association, Rivers (2007) further explicates their initial thoughts:

> Norman Nie and I founded Knowledge Networks (originally InterSurvey) in 1998 with the idea of bringing probability sampling to the Web. The approach used by Knowledge Networks addresses the fundamental problems of sampling for Web surveys. The panel was recruited using RDD, so the sampling methodology was not particularly controversial. Coverage of households without Internet access was accomplished by providing respondents with an inexpensive device to connect to the Internet. For many items, the panel produces estimates that appear quite similar to conventional RDD surveys conducted using telephone interviewing (Rivers, 2007, p. 2).

Panel participants were recruited through offline random sampling from the US population. These participants were furnished with the required hardware at no cost and were also reimbursed for their network connection expenses. In exchange, they were expected to partake in a survey approximately once per week. Those chosen as participants would receive an email requesting them to complete the questionnaire within a specified timeframe.[16]

Nie and Rivers were thus highly successful with their approach, winning the 2001 Innovators Award from the American Association for Public Opinion Research "for developing a probability sampling method for internet-based surveys in the United States" (AAPOR, 2006). Their approach thus seemed to be most fruitful and successful, but doubt also became a companion

---

[16] As Schmitt (2004: 284) adds, the primary drawback of this approach lies in the fact that participants receive an internet connection, leading them to become experienced internet users over time. Consequently, there is a possibility that this could influence their media consumption patterns and leisure activities. Therefore, it is important to note that the findings might only be generalized for internet users and not the general population.

quite early on. At the beginning of InterSurvey, Rivers told me that they "got what looked like an initial response rate of around 50 percent" and that "people liked the idea" (Rivers, Interview). After inspecting the data, however, he would find out that after taking attrition[17] into account, the response rate actually was in the single digits, "not anywhere close to a respectable response rate" (Rivers, Interview). They thus realized that they were confronted with similar problems associated with telephone interviewing, which is that the sample runs the risk of being biased due to low response rates.

6.3.1.2   Beyond Convention: Challenges to the Status Quo

Throughout the same time, other competitors came up with different suggestions that no longer took place within the framework of normal science, that is, probability sampling. Those new approaches broke with the paradigm of choosing participants randomly. Couper (2013) describes the rise of those early opt-in and access panels as "meteoric", such that those promoting them "were claiming that they make other methods of survey data collection obsolete" (Couper, 2013, p. 149). Early adaptors, such as Gordon Black, then chairman of Harris Black already stated in 1999 that "[a]ll research is going to migrate to the Internet" and that if they can do what they "want to do in 2000, it will all but eliminate telephone polling going forward". Harris Black International Ltd was among the first to conduct opinion polls online, which "ignited a fiery debate among public-opinion researchers" (Simons, 1999). In a press release from August 1, 1999, Harris Interactive positioned themselves in the history of polling and survey research: "Market research began with door-to-door household surveys which gave way to telephone polling in the mid-1960s and is now making a quantum leap forward with new Internet research techniques" (Harris Interactive in: Couper, 2000, p. 482p). This quantum leap forward was,

---

[17]   Panel attrition describes the phenomenon where individuals who had previously agreed to participate in a longitudinal study or panel survey then dropped out, resulting in a decline of the study participants.

however, not welcomed by many in the industry and academia as it challenged the gold standard in the field.

Rivers remembered this time, saying that he felt "offended" (Rivers, Interview) by other groups diverging from the probability-based paradigm.

> So, we started InterSurvey in 1998 when Harris Interactive was starting as well. […] And they'd done a deal with […] Excite, where they'd gotten people off the Excite website. […] Anyhow, so they had a, you know, just a sample of the users, allegedly millions of users from the site. And they were doing quota sampling and that offended us […], you know, this has a bad reputation in the sampling world and so forth. (Rivers, Interview)

At this time, with the help of Excite Inc., Harris Interactive had a database of three million Internet users. Participants were motivated in that they could win prizes and cash awards (Simons, 1999). What Harris Interactive did was to use a technique called propensity scoring, a statistical technique developed by Rosenbaum and Rubin (1983). This technique aims to achieve comparability between the populations by accounting for all variables that may affect the comparison. Therefore, the surveys conducted by Harris Interactive included questions to capture the general attitudes that are believed to differ between the online and general population (Schonlau, Fricker and Elliott, 2002, p. 66). Harris Interactive published its approach in the form of a white paper, which ends with the following statement:

> Some of our colleagues argue that we have forgotten the past, notably, the mistakes made by polling organizations that depended on non-probability samples in the 1948 pre-election polls. We instead contend that our use of propensity score adjustment shows that we are indeed keenly aware of the mistakes of the past and have successfully moved beyond them. (Terhanian *et al.*, 2000, p. 9)

Insisting on the sophistication of their new methods is a way to endow their practices with legitimacy. In doing so, they legitimise a methodology that goes against the dominant paradigm, that is, that probability sampling is the only way to reach a true representation of the social world. To many people at the time, this was a huge problem. Economically, it made much sense, but it was, and still is to many, a strong violation of industry standards, which also resonates with the

accounts of other participants in the field at this time. Furthermore, how exactly they applied those new methods remained largely proprietary, violating the Mertonian scientific ethos of communism, which "is incompatible with the definition of technology as 'private property' in a capitalist economy" (Merton, 1974, p. 275).

> The problem came, unfortunately, because that mode of doing things scientifically, random sampling, computer self-administration, giving people […] the Web TV device to hook up to their television [was no longer applied]. That was expensive and at the time, other companies, […] create[d] a much less expensive business model […], which was no random sampling, just advertise, let people volunteer (Krosnick, Interview).

The bad reputation of their approach has to do with the history of sampling, as we have seen throughout the previous chapters. To Krosnick, a fierce critic of non-probability sampling and a highly respected scientist when it comes to the measurement of the accuracy of probability and non-probability samples, this violates the principles that have become accepted and trusted, especially ever since the theoretical proof for the superiority of random sampling was firmly established in the 1930s. Krosnick thus evaluates this as a pushback to "what we were doing in [the] 1920s" (Krosnick, Interview).

In essence, the size of the sample was seen as trumping the design, which is not right from Krosnick's and Rivers' perspective. Furthermore, this bad reputation is not just connected to the history of survey sampling itself, but also with its application for polling and election forecasting. The contemporary discourse thus also relates to the miscalls that happened in 1936 and 1948 that still shape the ideas of many people in this time. The end of the 20th century can be seen as the beginning of a new chapter in survey research, one in which traditional methods began to be put into question and one that aims to overcome the historical trauma[18] of 1936 and 1948 that tied

---

[18]   The 1936 US presidential elections was, of course, not a trauma for Gallup and the "scientific" pollsters, since they were able to use this event to promote their own approaches and to become the central players in the field.

people towards the anchor of randomisation. We can say that in this time, many problems that came with traditional random sampling methods were known, however, the belief that its method was superior was still widely shared, since its mathematical principles were never challenged.

6.3.1.3 We Have Never Been Random

Rivers, who became aware that his real response rates were not as high and promising as he initially thought, went on to tell me that Kirk Wolter, who is now Statistics Professor at the University of Chicago and Principal Statistical Advisor and Distinguished Senior Fellow with the National Opinion Research Center, but was Vice President of Statistical Design Worldwide for A.C. Nielsen Co. at this time, thus a respected and trusted figure in the field, told him that they gave up on trying to do probability sampling, which "surprised" (Rivers, Interview) Rivers, because he was convinced that "there is no inference you can do from these things", that "it is impossible" (Rivers, Interview). But he also realised that most samples that were called probability samples really did not fit the assumptions that are usually associated with it. He stated:

> And, you know, so I realized what I was saying actually wasn't correct. It doesn't mean that, you know, you shouldn't do probability sampling. I think where you can and can control it is great. But there are these applications, where people are calling stuff probability sampling, but it, you know, I mean something with a five percent response rate can by no measure be called a probability sample. Calling it probability-based is marketing bs. You know, it's a self-selected sample (Rivers, Interview).

What he ultimately states is that also in every probability-based survey, one needs to make assumptions and that both, a probability and a non-probability sample are necessarily self-selected samples. This view is shared among other participants in the field as well. One interviewee, for instance, told me that they "wish for the entire world to follow a normal distribution and that everything would unfold in a purely random manner", adding, that they, however, "know that it would be ridiculous" (Anonym, Interview, Translation) to think so. To Raimund Wildner, who holds, besides his other roles, the role of vice-president of NIM,

chairman of the Council of German Market and Social Research and who is a member of the Professional Standards Committee of ESOMAR, the probability-based sample appears as a unicorn: "The random sample described in the textbooks is something like the unicorn of market research: it is beautiful, everyone also knows what it looks like, but – unfortunately! – no one has seen it yet" (Wildner, 2020 Translation). Those views challenge the common belief that probability-based sampling is inherently superior to non-probability sampling due to its random and unbiased nature, comparing it to a mythical creature. To paraphrase Bruno Latour (1993), Rivers and Wildner proclaim, in essence, that We Have Never Been Random and that the belief in the existence of purely probability-based sampling is constructed through textbooks, academic literature, and conventions in the field, but does not really exist. In this sense, Wildner echoes Rivers in arguing that we should try to maintain random processes, but that we should not lie about the reality of it:

> So, we should […] try to maintain random processes as long as possible. But at some point, we must be honest enough […] to admit that it's no longer possible. And then you can't really call it a random sample anymore. Then we would perhaps have to say approximate random sampling or some other term. But to call it a random sample is simply lying! And if you want to make statistics and create facts, then you must not lie. Otherwise, you destroy trust (Interview, Wildner, Translation).

Starting with random selection from a well-defined sampling frame, but with a particularly small response rate becomes, so Rivers and Wildner, a self-selected sample. By labelling the practice of neglecting to acknowledge this aspect and promoting it as a probability sample as "marketing bullshit" (Rivers, Interview) and a "lie" (Wilnder, Interview, Translation), they not only highlight their emotional involvement in these advancements but also establish a moral distinction between themselves and those who persist in presenting their methods as pure probability sampling. In doing so, they position others as being outside the realm of the scientific and professional ethos. In a way, those characterisations can be seen as a legitimation of their

own practice qua delegitimising their opponent, a form of boundary work that can also be found on the other side of the controversy, as we shall see in more detail in this and the next chapter.

6.3.1.4   The Event of Conversion

Rivers continued retracing his own narrative of conversion against the backdrop of the insights and frustrations that he developed. By the end of 2001, Rivers left Knowledge Networks and went back to teaching, before, in 2003, he started to think about alternatives, which resulted in the foundation of a new company, called Polimetrix. His "initial plan for Polimetrix was still to use probabilistic selection" (Rivers, Interview), but he also told me how, by this time, the cost of recruiting one person who actually participated was at over 100 dollars, which means that creating a sample involved high costs. So what he did was to contact various pollsters and pay them for that at the end of their interviews, they would ask respondents if they wished to take part in additional surveys, which would allow him to recruit panel members at a much lower cost. However, he mentioned that since the panel was very hard to manage, someone suggested that he recruit new members through Google AdWords, an online advertising platform that allowed businesses to create and manage advertisements that would show up on Google's search engine results pages (SERPs) and other Google-affiliated websites. Highly sceptical at first, but eventually, he would see its benefits and started to shift the whole model to an opt-in panel.

> [A]nd someone suggested, try using Google AdWords to recruit people. And I said, well, yeah, the stuff I'm getting is not, you know, random in any level. Let's see how this works. And it worked unbelievable well, we could get to recruit people for pennies and a fair number of them stayed on. And we then shifted the model of Polimetrix to a large opt-in panel. And the goal there was to have everyone matched to a voter file that we could use as the sampling frame. And at that time, I came up with sample matching as a strategy for implementing the selection as opposed to quota sampling (Rivers, Interview).

Sample matching means that they utilized data from the U.S. Census Bureau's American Community Survey in order to create a random sample based on 15 variables of representativeness. The result of this strategy is that you get a huge amount of people to take the

poll, but since you only need a part, you choose the part that matches the population you want to further know.

Giving more context to this observation, Wildner puts this in context of the longer history of sampling, acknowledging the historical dispute between quota sampling and random sampling. In doing so, he criticizes the use of dishonest tactics and challenges the misconceptions surrounding random sampling.

> So, something like that has actually always existed. In the past, there was a rather bitter dispute between quota sampling and random sampling for many years. That was such a dispute. Then and now there are new companies […], which of course also grab a part of the market and are then fought by others. One thing has to be said, however, that some of this fighting has not been done by honest means. And that's what I wanted to draw attention to […]. So, there was [someone, who] said that you have to make random samples and a random sample is characterized by the fact that I know exactly beforehand the probability with which each participant will be included in the sample. Sorry, there is no such thing as random sampling. And they should say that honestly and not just claim things like that, following some chimera (Wildner, Interview, Translation).

This quote highlights further forms of delegitimation, arguing that others are not only morally in the wrong but also cognitively. Following a chimera means to follow an illusion, which is something scientists certainly should not focus on if they want to be taken seriously.

Douglas River's entrepreneurial and scientific journey can be described as one of conversion from being a proponent of probability sampling to one of non-probability sampling. As we can trace on the basis of these interview excerpts, Rivers' conviction of the paradigm of random sampling to be the only legitimate and sound way of doing polls and surveys came into a crisis, by the end of which he became a defender of non-probability methods. The story is one in which he realized that what he believed in for a long time, actually was an illusion. The authority of a trusted figure, in this case Kirk Wolter, led him to reconsider his own convictions, a constellation which we will also observe in the story of Jon Krosnick, who, as mentioned earlier, is critical of Rivers.

6.3.2 Defending the Challenged Programme

6.3.2.1 An Unexpected Opportunity

Krosnick also went through a conversion – but it was in the opposite direction. His conversion, as we shall see, might as well be conceived of as an entrenchment of a prior position, whereas it also displays a strong turning point in his methodological orientation. Krosnick, author of many well-cited studies on the accuracy of probability and non-probability sampling methods, explained to me how he first got interested in the endeavour of measuring accuracy and how he became a fierce proponent of the probability-based paradigm. His narrative starts in 2003, when he "came to Stanford as a new faculty member" (Krosnick, Interview) and where he met Morris Fiorina, a political scientist at Stanford University. At that time, YouGov, which would in 2007 acquire Rivers' company Polimetrix for an undisclosed amount,[19] had already achieved success in the UK and had set its sights on conducting a survey for the 2004 US presidential elections between George W. Bush and John Kerry. Conducting a poll on the US presidential elections already heralds YouGov's later expansion onto the US market, which by now even constitutes a bigger market for the firm than the UK, where it was founded and is still publicly listed. The plan involved applying YouGov's methodology from the UK to the US and sharing the resulting data with the weekly newspaper *The Economist*, who would then publish the polling results. In the course of them planning this project, YouGov approached Fiorina and inquired about his interest in collaborating on their survey for the upcoming elections.

Fiorina then approached the newly hired Krosnick to ask if he would be interested in collaborating on the project, leading to several meetings between YouGov and The Economist

---

[19] What is known is that YouGov reported a revenue of £1.1 million and a pre-tax loss of £0.7 million at the end of 2006. At that time, Polimetrix had gross assets totalling £4,212,000, and YouGov, which already owned a 32% stake in the company, acquired the remaining portion for £11.7 million (mrweb, 2007). Wikipedia states that it was acquired for approximately $17 million, but does not provide a source. (Wikipedia contributors, 2023)

editor, during which the plan and possibility of conducting a survey such as this one was discussed. YouGov designed the survey, which was planned to be conducted online from July to November 2004. They utilized online advertisements on platforms like Google and Overture, along with emails sent to commercial opt-in mailing lists, as a means of recruitment. The objective was to gather around 20,000 American participants for the survey. In order to motivate participation, incentives like monetary payments and opportunities to win prizes at the end of the survey were offered.

The problem was that, as Krosnick recalls, they "didn't really have any control over the methodology", which means that it was actually "YouGov doing what they wanted to do" (Krosnick, Interview). Krosnick and Fiorina had to trust the methodological approach invoked by YouGov. They agreed to participate and "kicked it off by writing a little paper, which I'm going to call a white paper" (Krosnick, Interview). In this paper, they reviewed various approaches and past polls to make a cautious case for the use of non-probability sampling methods in this particular case.

6.3.2.2 Challenging Dogma: Embracing Unconventional Scientific Approaches

The white paper is deemed particularly interesting, since it not only provides the background to "the Economist/YouGov Internet Presidential Poll," as the title of it reads, but it provides a contextualisation of their position in the light of the history of sampling and the context illuminated in Rivers' journey. The paper starts with a reference to the 1936 US presidential election and how probability sampling "became the gold standard of American public opinion research" (Fiorina and Krosnick, 2004, p. 1), which would become more and more difficult, due to the reasons outlined earlier in this dissertation. After describing those developments, they state that while "the traditional survey method is very valuable and often spot on, […] it is not perfect" (Fiorina and Krosnick, 2004, p. 3). They thus acknowledge the continuous value and importance

of probability sampling, whilst pointing out that it might no longer be the perfect technique to gauge public opinion.

As an alternative, they emphasise two possible approaches using the internet. The first is to use traditional probability methods to draw a sample and to then connect everyone who does not yet have the relevant hardware. They thereby refer to Douglas Rivers and Norman Nie's InterSurvey, as presented in the previous subchapter. As an alternative to such a 'pre-recruited panel of full population', they also mention the possibility of a 'pre-recruited panel of Internet users', an approach that "entails a re-examination of first principles; namely, the principle that a probability sample is the only means to achieve a representative sample" (Fiorina and Krosnick, 2004, p. 4), a statement resembling the white paper put forth by Harris Interactive four years earlier. While not being fully convinced by this new and at the time still highly unconventional approach, they still deemed it worth testing: "We're just going to give it a try, let's see what happens" (Krosnick, Interview). Their white paper goes on stating that although "the practice would seem to be a regression, a return to the straw polls of the pre-1930s, such procedures cannot be dismissed so simply [as a] great deal more information about the population is now available – its demographic characteristics, interests, and activities – than was available two generations ago" (Fiorina and Krosnick, 2004, p. 4). After weighting existing research applying this methodology, they conclude that "[o]nly as more such research is done can we understand how volunteer Internet polling unfolds and what determines its accuracy in particular contexts" (Fiorina and Krosnick, 2004, p. 5). While not fully endorsing the shift towards non-probability sampling, they argued that under changed circumstances, it might now be a feasible approach.

6.3.2.3 Doubt and Recognition of a Danger

This paper was not taken up too positively by some in the community, as Krosnick continues to explain.

And so, I got an email after that and I would say, maybe a somewhat angry email from a friend of mine whose name is Kathy Frankovic […] And she was the director of polling for CBS in New York. […]. And she's retired now, but she's a major, major, very important person in survey research. And she wrote me, you know, basically saying, Jon, how could you, this is terrible, non-random sampling is crazy. You seem to be endorsing it. It's a really bad idea. And so, wow, that was pretty striking to me because I respect her, and I trusted her, and she thought this was a dangerous pathway to go down and that we would be giving a spotlight to a methodology that would cause problems (Krosnick, Interview).

At this point, Krosnick started to get worried because he realized that he is on the brink of endorsing a methodology not backed by sound theoretical principles. Similar to the case of Douglas Rivers, where the change in mind was initiated through a trusted person, this also was the case for Jon Krosnick. He continued, stating that Frankovic was right, but also inspired him to pursue a path that gets him out of it.

Turns out she was right. She knew this long before I did. But what she did was basically, I mean, it's kind of like too late for me to get out. So, what she inspired me to do was to evaluate the data, to evaluate the accuracy of all the data by various different methods. So, in many ways I credit her as the person who inspired me to write all those papers […] and once I did one of them […] I kind of was stuck, I have to add to keep doing it (Krosnick, Interview).

Krosnick's conversion thus started at the time when non-probability sampling has already found its way into the practice of polling, as we have seen in the discussion of Rivers' journey. While Krosnick admits that he initially did not fully recognize the potential issues associated with non-probability sampling he then realized the gravity of the situation and the potential dangers of promoting a methodology that could compromise the accuracy of survey data. After all, it made him reconsider his course and motivated him to engage in the study of which of both approaches fares better. The Economist / YouGov poll, at least, showed John Kerry winning with 50% of the popular vote and 47% for George W. Bush.[20] The actual results were the other way around, with current president Bush gaining 50.73% and Kerry 48.27%.

---

[20]   The final report, showing the predicted results, is stored and available through The Economist's website: https://www.economist.com/media/pdf/FinalPrediction.pdf (last accessed: 12.09.2023) The authors nevertheless acknowledged that a Bush win is still possible considering the poll's margin of error.

6.3.2.4 Looking for the Truth

This led Krosnick to engage in research to evaluate the accuracy of surveys and polls. What he realised by this time was that "there was not a long tradition of published papers in the literature evaluating the accuracy of surveys and so I kind of had to figure out how to do it" (Krosnick, Interview). He sought a method to achieve this and devised a plan to compare specific benchmarks: For instance, he aimed to determine the accuracy of surveys by comparing respondents' statements regarding possession of a passport or a driver's license with corresponding information from government statistics. Additionally, he considered utilizing other information, such as inquiries about medical conditions like diabetes or asthma, which could be evaluated using high-quality face-to-face surveys conducted by the federal government. "That then gave me what I came to call benchmarks, [which] […] were kind of measurements of truth. And the idea was [to] […] find out how similar the survey measurements are to the truth as a way of assessing accuracy" (Krosnick, Interview). Among those studies was one involving Harris Interactive and Knowledge Networks. The paper was titled "The accuracy of self-reports: Comparisons of an RDD telephone survey with Internet surveys by Harris Interactive and Knowledge Networks" (Chang and Krosnick, 2001).

> And Harris Interactive was one of those non-probability volunteer Internet firms and so that was the start, but then we did bigger and bigger versions of it with more and more companies, and we got the same result every time that the random digital telephone calls doing probability samples were startlingly accurate, even though their response rates were nowhere near one hundred percent. [A]nd that the probability sample Internet survey created by Douglas Rivers and Norman Nie was also really, really accurate. But the non-probability samples never were, they were always less accurate, and one of the findings that we have seen is that they are sometimes shockingly inaccurate (Krosnick, Interview).

In the course of them conducting accuracy studies, many of them were widely read and discussed. An argument he often came across, as he said, is that people say he is "claiming that random samples are good", but people would say "they're not as good as he says because the response rates are horrible. They're not actually random samples" (Krosnick, Interview),

resonating with his own position in the context of the YouGov/Economist poll. While Rivers and other's argument is to a great extent based on those conclusions, Krosnick would first agree: "And that's true that these days that if you draw a random sample of telephone numbers, like by random digit dialling, and then you call those telephone numbers, there are lots of phone numbers where probably there's a person at the other end who you will not interview. And that was true from the very beginning of our accuracy papers, that response rates were not 100 per cent" (Krosnick, Interview). Importantly the principle of random sampling is based on the idea of a 100% response rate, that if you draw a purely random sample and interview them all, the results should be extremely accurate. The critics, so Krosnick, "are right", but the "amazing thing" is that "if you start with a random sample and you work hard to interview as many of those people as you can, even with response rates in the single digits, […] those data continue to be extremely accurate" (Krosnick, Interview). Low response rates are, so Krosnick, not guaranteed inaccuracy, it merely means that there is a risk of inaccuracy. Even though, if we follow the argument of Rivers and Wildner, that probability samples are unicorns, Krosnick would argue that chasing the unicorn still yields the highest quality of results.

When asking other participants in the field about those accuracy studies and why they seem not to be taken up by many pollsters and research buyers, Rainer Schnell, supporting Krosnick, told me the following: "There are all empirical studies on the selectivity of web surveys show the same results. They all show the same thing. And it's a mystery to me how anyone can ignore that. I stand there and marvel" (Schnell, Interview, Translation). On the same question, Thorsten Thierhoff responded with a rhetorical counterquestion: "Yes, that is a good question. Why are there charlatans and why are they not always discovered" (Thierhoff, Interview, Translation).

There appears to be a prevailing belief in the field that certain individuals and organizations deviate from the principles of trustworthy and honest scientific inquiry. The stakes for defenders of probability sampling could not be any higher. To Krosnick, the developments that took place

in the survey and polling world constitute the end of trustworthy science, a science that adheres to the truth. And he thereby thinks of truth in two senses, first "to describe the claims that researchers make about the accuracy of their data", but also in referring to "the accuracy of measurement" (Krosnick, Interview). While we have so far talked about truth of the second type, he also referred to truth of the first type. In referring to the period explored in River's journey, he stated that most of what happened then was based on falsehoods.

> And unfortunately, the beginning of the demise of all of this happened then. It happened when the companies that were starting this said - they disseminated information to the world that was not true - what they said was our surveys are quicker than telephone surveys. That's true. Our surveys are cheaper than telephone surveys. That's true. But they also said our surveys are more accurate than telephone surveys. That was not true (Krosnick, Interview).

To Krosnick, the rise of non-probability sampling is ultimately related to lower accuracy, but also to faster processes and cheaper prices, hinting at the political economy in which the survey and polling landscape takes place. Many media companies welcome the rise of cheap and fast data, which serve them in times of economic scarcity, they, in Krosnick's words, "no longer care that whether they're good or not, because […] [they're] having huge financial problems" (Krosnick, Interview). This, he fears, can herald the end of trustworthy survey research.

> If the non-probability sample world continues to lie to the public and its customers and continues to take away more and more data collection opportunities from the probability or scientific surveys, eventually what's going to happen is that all the companies that are collecting and selling probability sample data will go out of business. And so, when you have a facility that has hundreds of employees who make phone calls from random digit dial telephone calls and you close that facility, fire all those people and sell all the computers and you don't have the building anymore. It's gone. I mean, that's it (Krosnick, Interview).

He continued saying that many no longer care about the quality of polls, as long as they create clicks. In this context, according to Krosnick, nobody controls whether the data provided by surveys and polls be anchored in something tangible.

> Nobody is protecting the public anymore, nobody is telling the public, these are scientific methods, you should believe these, these are non-scientific methods, you should not believe

these or at least view them with a lot of scepticism. So that's essentially where we are today. We are in a world where it's far too inexpensive to produce crappy surveys that the public can't tell the difference. Lots of people don't want to know the difference. They say if it's cheap and I can use it and I can publish a paper with it, I want to do that. I don't care about accuracy (Krosnick, Interview).

To him, this has serious consequences, not only for the field of survey research and polling itself but also for democratic life and governance. His argument ultimately comes down to the point that surveys and polls ought to represent realities accurately in order to foster adequate policies and decisions. What he and other proponents of probability sampling fear is a bifurcation of the survey industry into expensive, but accurate and cheap, but inaccurate data. Representations need to be based on strong and stable recordings of social reality; they need to mirror reality correctly. Ultimately, Krosnick, however, sees the landscape shifting: "And I think, you know, we are losing, I mean, I would say the good guys are losing the war because the amount of money being given to the nonscientific nonrandom sample companies is much bigger than being put into the scientific work and I don't think their customers understand the truth about what they're collecting for them" (Krosnick, Interview).

6.3.3   The Outcome is Open

The stakes involved in those controversies over the right methodological approach also became visible outside personal viewpoints and developments. The last decade has seen an upsurge in headline-making mispredictions by pollsters when it comes to high-profile elections and referendums. These are reminiscent of the miscalls in 1936 and 1948 and may, a century from now, be treated in the same paradigm-shifting way that Gallup's famous triumphs and failures

back then are now treated. Most prominent among these are the US elections of 2015 and 2020 as well as the British House of Commons election of 2015 and the Brexit referendum of 2016.[21]

Following these mispredictions, many investigations were carried out to explore why such sometimes severe miscalculations occurred. In the case of the 2015 British House of Commons election, for example, Sturgis *et al.* (2018) concluded their "assessment of the causes of the errors in the 2015 UK general election opinion polls", with the main reason for the misprediction being a strong bias in the samples. Important to note, all election forecasts were based on non-probabilistic methods: "Our conclusion from these analyses is that the polling miss in 2015 occurred primarily because the procedures that were used by the pollsters to recruit respondents produced samples which were unrepresentative of the target population's voting intentions. These biases were not mitigated by the statistical adjustments that pollsters applied to the raw data. Other factors made, at most, a very modest contribution" (Sturgis *et al.*, 2018, p. 760). Sturgis *et al.* state that "[t]his collective failure led politicians and commentators to question the validity and utility of political polling and raised concerns regarding a broader public loss of confidence in survey research" (Sturgis *et al.*, 2018, p. 757), highlighting the critical role of robust methodologies to mitigate biases and enhance the reliability of pollsters' predictions, or in other words, to increase their predictive stability and to thus regain public trust.

In the case of the 2020 US election, however, something different may be observed. Under the title "Revisiting the 'goldstandard' of polling: new methods outperformed traditional ones in 2020", Enns and Rothschild (2020) were able to show that non-probabilistic methods performed

---

[21]   A very recent example are the polls in the run-up to the 2023 Turkish presidential and parliamentary elections, where the majority of polls saw the opposition leader Kemal Kılıçdaroğlu (CHP) above the current president Recep Tayyip Erdoğan (AKP). On election day, on May 14th 2023, Erdogan was, however, able to secure almost an absolute majority, leading the opposition parties and much of the population in disbelief about the value and accuracy of polls. In the run-off elections two weeks later, Erdogan won against Kılıçdaroğlu. Some mentioned the inherent political bias of pollsters towards the opposition in a nation characterized by highly polarized and firmly entrenched perspectives. But, with the reasons behind this polling disaster not yet thoroughly explored, I decided not to analyze it in more depth.

better than probabilistic ones: "The overwhelming majority of election surveys in 2020 included non-probability sampling methods, raising the question: Did the shift towards less expensive, opt-in samples hurt polling accuracy in 2020? Our research shows the answer is no. Non-probability surveys and surveys combining probability and non-probability methods outperformed probability-based surveys in the 2020 election!" (Enns and Rothschild, 2020). One possible explanation that Enns and Rothschild suggest is that the declining response rates for telephone surveys and Donald Trump's efforts to undermine trust in mainstream surveys, might have "hurt traditional methods, whereas innovative sampling appraoches were better able to reach the types of voters low in social trust that some have suggested were missing from the polls" (Enns and Rothschild, 2020).

This observation suggests a recurrence of the controversy that emerged in the early 20th century, as discussed in the earlier chapters of this dissertation. Probability and non-probability sampling approaches, or as referred to in those earlier years as random and purposive sampling, continue to coexist, albeit not always harmoniously. In the broader discourse, and resonating with the narratives presented in this chapter, three distinct positions can be identified regarding the debate over whether probability or non-probability sampling represents the superior method for surveying or polling a population. Carina Cornesse succinctly summarized these positions during a GESIS summer school on probability and non-probability sampling, in which I had the opportunity to participate. The positions can be outlined as follows: 1. The adoption of a probability sample is always necessary, particularly when aiming to make inferences to a broader population; 2. A probability sample is never essential, as statistical techniques can address and resolve any associated issues; 3. The need for a probability sample varies depending on the specific research objective, with instances where it is required and instances where it is unnecessary.

The clash happens between positions 1 and 2, constituting a strong divide in the field. Paradigmatically, those two positions from the scientific literature should illustrate this divide. As a proponent of position 1, Yeager *et al.* (2011) highlight that "probability sample surveys were consistently more accurate than the non-probability sample surveys, even after post-stratification with demographics" (Yeager *et al.*, 2011, p. 709). As a proponent of position 2, Wang *et al.* (2015) stated that "[w]hile representative polling has historically proven to be quite effective, it comes at considerable costs of time and money" and that "as response rates have declined over the past several decades, the statistical benefits of representative sampling have diminished" (Wang *et al.*, 2015, p. 980). Both evaluations here stem from a different epistemic context. Whereas the first seems to suggest that a sample ought to be free of bias as much as possible, the second acknowledges this bias and bases its inference in systematically taking it into account.

The field seems to be in a position in which there exists no generally accepted convention as to how to best "know" a population. Before moving on to the next chapter, which explores the forms of boundary work that takes place in the field, I briefly present two perspectives that encapsulate this situation of uncertainty. Drew Linzer, director and chief scientist at the American polling company Civiqs, for instance, told me in an interview that despite there being no doubt that the theory of probability sampling still holds true, no researcher absolutely trusts that probability sampling currently serves as a stable guide for the epistemic practice of polling. Rather than sticking to the old paradigm, which no longer seems to be able to be adequate, it is necessary to look out for better approaches:

> What I would say is that no, and I literally mean no survey researcher using any method currently available has absolute fidelity to that theoretical ideal. And online or non-probability sampling confronts that problem head-on. But I would argue that people who claim to be doing probability sampling should be confronting it head-on with just as much focus as we do (Linzer, Interview).

In a similar vein argues Raimund Wildner, who similarly sees the field in a crisis, emphasizing the need to think about new ways and means to continue.

> At the moment we don't have a reasonable science of sampling. [...] I think it's better to have the debate, rather than to have the debate brought to us from the outside. If someone comes from the outside and says that market research claims random sampling, that it's all lies and deception, and then they prove it, then we're in a very stupid defensive position. And because we can then say, yes, that's a problem, we know it, here, it says so. But we'll think about it. As long as we don't have anything better, we do the best we can (Wildner, Interview, Translation).

Wildner's statement already emphasises methodological self-reflexivity, which can be seen as a response to existing practices of contestation, which I will explore in greater detail in the subsequent chapter. The solution to the declining strength of random sampling is, to many, to draw on online non-probability samples. Against the background that historically, the superiority of probabilistic sampling has been demonstrated and established, as explored in the first two chapters, it is now interesting to see this old debate resurfacing. What we can observe here is that a, for commercial pollsters presumably dead method, now seems to offer a solution to several problems that probability-based surveys increasingly face. If scientific paradigms are held in place by networks of commitment, those strong networks that maintain and uphold the probability sampling paradigm are losing their strength, whereas the weak networks that previously held up and maintained non-probability sampling are gaining in strength.

6.4 Conclusion

The aforementioned discussions illustrate that both probability and non-probability adherents hold the potential to assert their legitimacy in certain respects. However, the future trajectory, in terms of which, or if any, approach will ultimately prevail, remains uncertain. From a sociological perspective, this state of uncertainty also highlights the idea that determining the representation of social reality on specific issues is not only influenced by the chosen methodology, but also by the authority attributed to them and by their ability to assemble

predictive stability. In this sense, the subsequent chapter undertakes an exploration of the struggles pertaining to authority, delving into how each camp aims to establish boundaries in order to present their own approach as superior.

CHAPTER 7

7   Contesting Legitimacy, Creating Boundaries

7.1   Introduction

In revisiting those stories of conversion, I identified typical positions from which actors in the field speak. As we shall see in this chapter, those actors not only engage in practices of justifying their own position, but also in practices of delegitimising other positions with the goal of creating boundaries between what is considered scientific and what is not. The controversy under study appears as one that is marked by retributions, it is a controversy in which colleagues, competitors and peers deploy heavy artillery to de-legitimise each other's claims, methods and practices. As we shall see and already have seen in the previous chapter, accusations such as magic, scam, fraud or charlatanry have been and are raised by members of the survey and polling community with the aim to delegitimate certain methodologies or practices, tendencies that remind of the work by Roger Burrows and Mike Savage (Savage and Burrows, 2007; Burrows and Savage, 2014), as introduced at the beginning of the previous chapter. What Burrows and Savage ultimately state is that whereas the social sciences were able to claim certain methodologies and techniques as their own and thus inhabiting a position in which they had an authority in producing knowledge of the social world, their role, authority and techniques are more and more pushed towards a marginal position in the light of the rise of new ways of data gathering and analysis. This leads to a situation that will be characterised by polarization, divide and contestation.

In times in which the monopoly over the representation of publics and their opinion started disintegrating, practices of boundary work evolved again, leading to increasing polarization in intradisciplinary struggles over how to best "know" and represent the social world. While the

previous chapter ended with the factors and developments that led to the emergence of this controversy, the current chapter will examine its dynamics. Specifically, we will explore how proponents of each research program seek to legitimize their own approach while simultaneously delegitimizing the opposing viewpoint.

7.2 Setting the Stage: The Resurfacing of an Old Conflict

7.2.1 Disrupting or Destroying the Survey Landscape?

In September 2019, Manfred Güllner, head and founder of Forsa sent out an e-mail to politicians and journalists, an E-Mail which can be seen as emblematic for the dynamic of the controversy around probability and non-probability sampling. Forsa is the Institute for Social Research and Statistical Analysis, established in 1984 and now among the leading market research and opinion polling companies in Germany. As journalists Robert Pausch and Fritz Zimmermann (2020) state, the subject of this E-Mail read, translated into English, "Nonsense, Fake-News and Manipulation". In the email, Güllner accuses Civey of being "unscientific" and of threatening the "hard-won acceptance of empirical social research".

Civey is a Berlin-based startup, with a name composed of the words Citizen and Survey, a new market and opinion research company founded in 2015 under the name OMNI TT GmbH. In a similar way as actors like YouGov, which we got to know in the previous chapter, Civey set out to revolutionise the field of opinion and market research by drawing on non-traditional methods and new sophisticated techniques. Civey now employs 110 people and has a turnover in the mid-single-digit millions, which has grown by 85 per cent in the past four years (Scheppe, 2023). Even though Forsa was able to obtain in court that Civey is no longer allowed to call itself the "market leader", it is hard to get around them, when exploring the field in Germany.

Why is Civey's approach so controversial? Why do so many traditional polling experts in Germany heap disdain on Civey? The discussion around the photo shoot is an example around

the question of, as Lusinchi, put it in a similar context, "what constitutes legitimate knowledge when the source of that knowledge is a sample survey" (Lusinchi, 2017a). Different to traditional surveys, in the case of Civey, participants chose to take part in surveys, rather than being chosen. This violates the principle underlying the practice of surveys, as already introduced by Bowley more than 100 years ago, which is that "[w]e can obtain as good results as we please by sampling, and very often quite small samples are enough; the only difficulty is to ensure that every person or thing has the same chance of inclusion in the investigation" (Bowley, 1906, p. 553). For a long time, this principle marked the boundary of what counts as legitimate statistical knowledge, a boundary that companies such as Civey – but also many other companies – put into question and people like Güllner aim to fortify.

As an observer of this controversy, I asked myself whether these internal attacks decrease the credibility of the whole industry and enterprise of survey research. Why did the controversy gain such a momentum and such harsh language? Why, one might ask, should one foul one's own nest? Market research and opinion polling has, in the US alone, an estimated industry revenue of 17.8 billion USD in 2020 (Statista, 2020), so reducing this controversy to one about market shares might seem to be an obvious point – and surely plays a role – but nevertheless misses a central aspect that can be linked to the biographical and educational background of the participants and the way how practitioners are attached to certain research programmes. The stakes involved can be viewed in the statements by participants. Shedding light on his disagreement with Civey, Güllner's view on the practices and activities of Civey stands in interesting ways for the larger controversy. In contextualizing his dispute with Civey, Güllner told me the following:

> What is important for me to emphasise […] [is that] [t]his is not a dispute about competition, it is not […] that we fear for our sinecures here. After seeing how difficult it was to bring empirical research back to Germany after the collapse of National Socialism, to gain acceptance for it, I am really concerned about how laborious it was, just as laborious as educating Germans to become democrats, from subordinates to citizens [*vom Untertanen*

zum Staatsbürger]. And that's why it annoys me that black sheep like Civey come along and endanger the whole enterprise, destroy the whole acceptance through a way of working that I once called a scam and I stick to it. That is what annoys me about it. It is not, so to speak, something that concerns Forsa. And I have also dealt with this scientifically, I know the history of the individual institutes […], and I see this laborious process that has still not been completed. There are still politicians here who say that they don't think much of polls, and they don't understand that we report what people tell us and that's what Civey is ruining (Interview, Manfred Güllner, Translation from German).

This extract reveals several highly interesting points, which guide the subsequent analysis. The interviewee draws on the difficult history of survey research and the laborious process of it becoming an accepted and trusted scientific endeavour. In doing so, it also points to the historical development of a certain convention of how surveys and polls should be conducted to be granted the label of being scientific. The reference to the end of World War II reminds of the US context after the 1948 US presidential elections, as discussed in chapter 3, that put the social sciences in a crisis of legitimation. The interviewee then mentions the appearance of Civey as a danger to the hard-won legitimacy of the discipline he represents, a history that is not yet complete, since, as he states there are still people who do not yet understand the task of survey research and polling to provide legitimate and sound empirical representation of the social world. What is noteworthy in this context is that the polling and survey research field relies on external validation to ensure a portion of its legitimacy. For instance, the media plays a significant role by commissioning polls, thereby granting polling institutes their legitimacy. Unlike other knowledge-producing disciplines like physics, where the validation of theories lies solely within the purview of physicists, polling and survey research, being an applied science, is subject to the need for external recognition and approval.

Other statements go in a similar direction. Jon Krosnick, for example described to me how he sees the whole industry of survey research and polling under threat, a threat caused by the adoption of new and non-traditional methods and practices. The following statement, even though he does not address Civey directly, but does, at different stages of the interview, show how this line of argumentation can be further generalised.

So, if the infrastructure of survey research disappears, if all of the phone facilities and all the face to face and companies go out of business. How can we recover from that? And that, to me, is the problem. That's what I'm worried about. And if we if we insist on mediocrity in order to make profits, we're going to pay the price that, you know, and it may be that that the entire research industry collapses (Interview, Jon Krosnick).

Both extracts point to the downfall of a laboriously and well-established framework, which can be referred to as a paradigm in the Kuhnian sense or a convention in the sense of the Economics of Convention approach. Probability sampling has proven to be a reliable framework to create representations of the social world and of public opinion. It has become a successful model of what Kuhn (1996) has termed "puzzle solving", the normal activity of scientists in dealing with problems that they know have a solution, as long as their methods are guided by the paradigm of normal science. Tensions arise from the fact that, on the one hand, as we have seen in the previous chapter, the paradigm of normal science, i.e., interviewing with probability sampling reached its limits due to declining response rates, while, on the other hand, the general superiority of the principle of probability sampling remained stable and uncontested. It is especially due to the conditions in times of economic struggles, e.g., newspapers that rely on affordable data and the rise of the internet, that the framework of normal science came under scrutiny.

This climate of changing epistemic and economic practices of knowledge creation led to many controversies, running through different themes, displaying different social dimensions. The following is an attempt to organize the disputes and controversies around non-probability methods based on their underlying structures. Similar to the previous chapter, the following also mainly draws from interviews conducted with pollsters and survey researchers, but also extends its analysis to secondary material, especially in cases where no interviews could be conducted. Throughout this chapter, I present a thematic analysis of the main arguments mobilised in this controversy, alongside an interpretation, drawing on major concepts in sociology.

7.2.2 Two Survey Worlds, Two Research Programmes

In the introduction of this dissertation, I introduced the controversy around the photoshoot between the German national players Özil and Gündoğan with the Turkish president Erdoğan. Just to briefly recall this discussion, due to strongly diverging pictures of how the German population thinks about this photoshoot and the role of those players in the German national football team, three major polling and survey institutes filed a complaint against Focus Online, the news magazine that published the Civey survey, with the German Press Council. The complaint brought to the German Press Council stated that the survey uses a design, "which, according to the generally accepted scientific criteria of empirical social research, is fundamentally not suitable for delivering 'representative' results" (marktforschung.de, 2018, Translation). In framing their complaint in this way, they aim to stabilise the current paradigm of survey research and polling. The press council's decision that it is not the task of journalists to check the scientific basis of the surveys they publish has been taken up by some in the industry as a green light to publish whichever representation of the social world, as long as its declared to be representative.[22] As Güllner warned, "if an institute says that something is representative, then the press is entitled to accept that and does not have to verify it further" (Güllner, Interview, Translation).

A different spin on the relation between polling methodology and media standards to publish them is the case of The New York Times. In their own polling standards, the newspaper is quite clear:

> In order to represent the population statistically, a survey should be based on a probability sample. This means everyone in the population must have an equal chance of being selected

---

[22] The Press Council consists of members of media outlets and editorial offices. The complaints committee is elected by a panel and consists of half representatives of publishers and half journalists, many of whom have cooperated with Civey. The decision thus did not come with great surprise to many, as representants of one of the complaining companies told me.

to participate in the survey or at least a known chance of being selected.[23] […] Non-probability samples are commonly used in Internet polls, call-in polls, blast e-mail polls and a variety of others. The entire population does not have an equal chance of being contacted. Most Internet polls are based on panels of self-selected respondents, and Internet access is not yet evenly distributed across socio-economic and demographic groups. The Times does not publish most Internet polls (The New York Times, 2008).

Whilst The New York Times still has those policies in place, their own poll reporting diverges quite strongly from those policies. Their poll reporting is, as an informant told me, a direct contradiction to its own policy, which they secretly overcame. This raises concerns about the quality control measures employed by major media companies, as it appears they are either becoming less stringent in implementing them or not taking them as seriously as before.

Polling appears to be an epistemic practice that enacts different representations of reality, not only depending on the measurement conventions in place, but also depending on the dissemination and publication of those very results. In the light of this, it also makes sense why they filed the complaint with the Press Council and not the Council of German Market and Social Research (Rat der Deutschen Markt- und Sozialforschung), whose task would be to ensure compliance with the professional principles and rules of professional conduct.

7.2.3 Dynamics of Inclusion and Exclusion

Coming back to the case of Civey and Forsa, it is interesting to see that whereas in the statements by Manfred Güllner, we see a rhetoric of excluding Civey as a legitimate member from the survey community, Civey itself engages in a rhetoric of inclusion, in the sense that they identify themselves as a valuable part of the survey community, one that is about to solve, and already has done so, some of their central current problems.

---

[23] One should add that everyone in the population must have a known chance > 0.

In a 2018 article, published under the title 'Opinion Research: Make Polls Great Again' in the magazine "Junge Wirtschaft", the magazine of the Junior Chamber of Germany, Gerrit Richter, Civey's Co-Founder and CEO gives some more context as to how the company aims to tackle the problems the field currently suffers form. After mentioning some famous mispredictions from recent years, he raised the question of what can be done to overcome those challenges, which, in his understanding, are mainly due to declining response rates:

> But what can be done to re-engage all segments of the population in opinion research and reduce mispredictions? When we founded the opinion research start-up Civey, we wanted to solve the problem with a crazy idea: We don't annoy people on the phone, we don't pay them money either, we just show every participant the representative result in real time. No result is concealed – no matter whoever commissioned the survey. In addition comes transparency about the raw data,[24] statistical error, and all participant numbers (Richter, 2018, Translation).

Civey sees itself as a new actor in the survey and polling world that is able to solve some of the most pressing issues the field is facing at the moment. Janina Mütze, founder and CEO of Civey, stated in a Startup podcast called Gründerszene, that they have overcome crucial problems that the survey industry suffered from in recent years.[25]

> So, I think what I can claim for us is that we have solved the survey problem of this sector to a certain extent, because of course the response rate, the willingness to participate in telephone interviews […] has simply declined. That's what makes these surveys so expensive and time-consuming. The whole thing is easier online, you click quickly […]. But we also have to show a valid result. So just because the AfD [Alternative für Deutschland, a far-right party in Germany], for example, mobilises its voters here, we can't show that somehow 95 percent are against taking in more refugees or something. That wouldn't work. So that's kind of the quality standard we have, that we show a valid image of society from this skewed set of data and don't let the loudest win. And why do the users participate? Yes, I think it's the low threshold […] [and] the interest in the result. So, in quotation marks, I get it for free, but in exchange for my personal opinion. I get the representative result, which was otherwise

---

[24] It is important to note that, since 2021, Civey no longer publishes the raw data of their polls and surveys. Previously, as Richter describes in the 2018 article, it was possible to easily switch between displaying the raw and the adjusted data, labelled as representative.

[25] Unfortunately, I was not able to obtain an Interview with Civey myself, which is why I draw on secondary interview material in form of e.g. podcasts or reports published on their website. Civey's response rate to my many requests was, so to speak, in the zero digits.

part of the business model of the traditional institutes (Mütze, Gründerszene Podcast, Translation).

The promise of Civey is thus to offer a way through which people can express their opinions online, without pre-selecting potential participants through a sampling frame. They acknowledge that this would at first lead to a skewed picture of society, which they promise to overcome in re-weighting responses to reach a representative picture of the population. In this sense, they tackle the problem of low response rates and the increasing economic struggles. Furthermore, they appeal in very innovative ways to the economic struggles that the journalistic field suffers from and they position themselves in an inclusive manner as an actor that helps the field recovering.

The theme of inclusion and exclusion has been prevalent in sociological thought ever since but has famously been developed by Georg Simmel. Simmel describes the stranger as somebody who is neither in nor out, they can be in and out at the same time. But overall, the stranger is someone who comes today and stays tomorrow and non-probability sampling has, as others in the field said, "come to stay" (Cornesse, Interview, Translation). Simmel describes the stranger as somebody, who is not the "owner of soil" (Simmel, 1950, p. 403), be it in a physical sense or in terms of social figurations. This is insofar as Civey does not follow the standard way of conducting polls and surveys and thus appears as a stranger, they do not own the soil and have become subject of various rejections, not only from the side of Forsa, but quite prominently so. On the other hand, Civey and with it similar companies have become more and more integrated in the overall landscape of polling. In March 2021, for instance, Civey has been included in the ADM, the Working Group of German Market and Social Research Institutes (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V), which gives them legitimacy. Approaches based on non-probability sampling are thus not only widespread, they are also being integrated in traditional and established organisations and learned societies.

This line of thought brings us back to the case of the New York Times and its poll reporting. Not only does the New York Times, in contradiction to its own standards, publish polls with various kinds of methodological backing. In July 2014, the New York Times and CBS News revealed that the market research and data analytics firm YouGov, which we became acquainted with in the previous chapter, would conduct their polls for the upcoming mid-term elections. This shook up the polling world, as YouGov relies on an Internet panel, which stands in contradiction to the New York Times' own polling policies.

YouGov, which has been expanding beyond the UK since 2006, has made strategic acquisitions to strengthen its global presence. Notable acquisitions include the Dubai-based company Siraj, the Scandinavian firm Zapera, the German firm Psychonomics, and the mentioned company Polimetrix. Between 2009 and 2011, YouGov continued its growth in the United States with the acquisition of several research firms, including Clear Horizons in Princeton, New Jersey, Harrison Group in Connecticut, and Definitive Insights in Portland, Oregon. In the same year, YouGov established an office in Paris. In 2014, YouGov extended its reach to the Asia Pacific region through the acquisition of Decision Fuel. Most recently, in July 2023, YouGov made a significant move by agreeing to acquire the consumer panel division of the German market research company GfK for €315 million. YouGov is currently listed in the UK, but the recent acquisition of the consumer panel division of GfK, which, as YouGov co-founder and former CEO Stephan Shakespeare said "increased our size by 50 per cent overnight" (Shakespeare in: Thomas, 2023) made them consider moving its listing from the UK to the US, as the company would "be introduced to a bigger market" (Shakespeare in: Thomas, 2023). YouGov grew from a company with one office, 20 employees and 1,000 panel members in 2000 to a company with 39 offices, 1,650 employees and 24,000,000 panel members in 2023.

This decision from The New York Times to hire YouGov to conduct their polls led Michael Link, then president of the American Association for Public Opinion Research (AAPOR) to

criticise this decision in arguing that selecting a sample via the internet has "little grounding in theory" (Link in: Lusinchi, 2016) and lacks transparency. This response from AAPOR led, in turn, to other responses, among them one from Andrew Gelman, professor of statistics and political science at Columbia University and David Rothschild, economist at Microsoft Research. Gelman and Rothschild (2014) countered AAPOR's statement in saying that their "rigid faith in technology and theories or 'standards' determined in the 1930s [is] holding back our understanding of public opinion [and is] putting the industry and research at risk of being unprepared for the end of landline phones and other changes to existing 'standards'". The accusation became even stronger in a second statement, an outburst, when Gelman (2014) ridiculed Michael Link in a blogpost with the title of "President of American Association of Buggy-Whip Manufacturers takes a strong stand against internal combustion engine, argues that the so-called 'automobile' has 'little grounding in theory' and that 'results can vary widely based on the particular fuel that is used'". He wrote that what Michael Link wrote "really upset" him "in that it reminded" him "of various anti-innovation attitudes in statistics" he "encountered over the past few decades". He goes on saying that the "kind of aggressive methodological conservatism just makes" him "want to barf". The theme of accusing one's opponent of being a luddite is a common and recurrent theme in the controversy and points in interesting ways to the paradigmatic shifts that the field is undergoing. The Luddites in the nineteenth century and their famous machine-smashing were not a movement about technology denial or hatred of progress, but about fears of shifting power balances and of being replaced (see e.g., Stilgoe, 2020, p. 27). The force of this criticism is significant, not only due to the harsh language used. While, as we have seen in chapter 3, the field of survey research has always been one of controversy, the kind of boundary work engaged in contemporary debates is notably distinct and emphasizes the high stakes at play in this discussion, often resulting in exclusionary and confrontational discourse.

In contrast to the form of boundary work engaged, for instance, by proponents of established companies against Civey, this form of boundary work, targets and attacks the established core in the community. One might see this form of boundary work as opposed to the one that stems from the inside, from the side of the probability camp, it is however, interwoven with it. It is a quasi-mirror-image boundary-making, constituting a rhetoric of exclusion from the outside. While Michael Link, speaking for AAPOR in this case, aims to further manifest the historically grown boundaries, similar to people like Krosnick or Güllner, others like Gelman, Rivers or Wildner aim to shift and extend those very boundaries. It has become clear that discussions about those boundaries often take on rhetorical forms of boundary work, through which practitioners discredit each other, denouncing their role as legitimate actors in the community.

7.2.4   A Transparent Boundary Object

We have just seen there is a rhetoric of exclusion and one of inclusion, but this is not all there is – throughout the interviews, one theme that came up many times was that of transparency. Transparency, it seems, became an obligatory passage point for many learned societies and organisations, in order to include non-traditional methods into their repository of accepted and legitimate approaches. The Roper Center's transparency and acquisitions policy is a case in point here. Established in 1947 by Elmo Roper, the Roper Center's core mission revolves around the collection, preservation, and dissemination of public opinion data. Its purpose is to function as a valuable resource aimed at enhancing the methodologies of survey research and fostering a more comprehensive understanding of public sentiment through the utilization of survey data. In 2019, the Center's archives, the world's largest database of Public Opinion Research, began accepting non-probability research to be housed in a separate area of the archive, on the condition that full methodological disclosure is provided. In addition to its "Longstanding Methods Collection", there is now also a "Recently Developed Methods Collection". Gary Langer, vice chair of the

Board of Directors and chair of the transparency committee told me that they "opened the collection to non-traditional methods, happily so":

> There are now two collections in the Roper archives […]. And one is for traditional methods or existing methods, long-standing methods as we call it and the other for newer experimental methods. The only requirement we have to accept data into the new methods collection is that we have full transparency. Because then we can serve the vital role of opening these data to the research community for assessment of the research claims, by evaluation of the means by which these data were collected (Langer, Interview).

Langer makes a distinction here between 'long standing methods' and 'newer experimental methods', a separation no longer between in and out; he offers a conditional invitation. Transparency here becomes a "boundary object" in that it allows different communities to collaborate. Boundary objects are "artefacts, documents, terms, concepts, and other forms of reification around which communities of practice can organize their interconnections" (Wenger, 1998, p. 105). Under certain conditions, the stranger is being allowed to enter, however, only into a separate space. While boundary objects may allow to overcome the boundaries between different, otherwise separate communities, the practice of boundary work creates and stabilizes those very boundaries.

From both sides of the continuum between probability and non-probability sampling, the issue of transparency has been raised many times and in different forms. Whilst some interviewees used notions such as "magic" or "fraud" to describe a general lack of transparency on the side of their opponents, Langer did so more straightforwardly, implicitly endorsing Merton's scientific ethos of communism:

> This is the point at which I personally am willing to go and do battle, transparency, transparency is essential. And to claim some proprietary method of producing data that I can't tell you about, you have to trust me and believe, is antithetical to the cause of science and good research. […] And any claims of proprietary methods that cannot be disclosed, to me, is fundamentally disqualifying, for taking any data seriously. It should be thrown out the window, if we don't have a full and fair description of every piece of the methods involved (Langer, Interview).

This points to an interesting relation between polling and public opinion, a relation that chapter 4 has already foreshadowed. The practice of public opinion polling as an attempt to measure public attitudes and wants is an endeavour that oftentimes takes place behind closed doors. It is the demand of publicness towards those that produce accounts of the public. Following Porter (1995), whereas descriptive statistics has traditionally been used strategically as a transparent and public means of counting, inferential methods might increase public suspicion. In this context, transparency acts as a boundary object for inferential statistics, enabling methods and approaches to become subject to public and scientific scrutiny.

7.3 Dynamics of the Controversy

7.3.1 A Shifting Paradigm?

While some regret this changing landscape underlying the production of representations of the social world, some view it as a natural process in the way how the field adapts itself to changing circumstances. Throughout the interviews conducted for this research, the notion of paradigm shifts, as suggested by Thomas Kuhn, has been mentioned by two participants, drawing, however, different conclusions from it. Rainer Schnell, one of the key actors in Germany and known as a fierce enemy of non-probability sampling methods and a loud critic of Civey puts it this way:

> Well, Kuhn's concept of paradigms, as vague as it is in the original, has as its essential criterion that there must be a successful model solution [*puzzle-solution*]. That is, someone must have demonstrated that it actually works. And the term paradigm is always used so vaguely, in the sense of a worldview. But that's not what Kuhn means. To Kuhn it means that there is actually a successful application […] and with non-probably samples, I don't see one. I don't see a successful demonstration of solving a problem, at least beyond magic tricks. Being right once it's not a problem; the problem is being right often [laughs] (Schnell, Interview, Translation).

The interviewee here provides an interpretation of current developments in the field from the perspective of the sociology of science. He does so in referring to Kuhn's notion of what

constitutes a paradigm, which in his understanding means to have been demonstrated successfully. The interviewee rejects this status for non-probability methods and adds that the only success was achieved through magic tricks. This notion here serves as a fundamentally delegitimizing description of the scientific practice of other, deviant approaches. What those from the other camp do, testifies to cognitive inferiority and adheres to long-established forms of non-scientific knowledge creation: magic. In claiming there to be magical knowledge, to Schnell, there is no scientific discourse when talking about non-probability sampling, which in turn thus also cannot constitute a paradigm. Magic cannot irritate knowledge and practice in a scientific discourse, but only in one that is conducted beyond the boundaries of meaningful scientific conduct.

The notion of magic has been invoked in discussions around statistics for long a time, see for instance the notion of 'statistical magic'. The notion also found its way in an article by Humphrey Taylor and George Terhanian, representatives of Harris Interactive, who titled an article defending their non-probability approach as "No Witchcraft Here". They write: "The object of our efforts is to develop and refine a weighting routine, that works on those occasions when we are unable to mount parallel telephone surveys. This is not witchcraft: it is painstaking, theory-driven work." (Taylor and Terhanian, 1999, p. 42,43). In doing so, they mockingly pre-empt the accusation of practicing magic rather than legitimate scientific research. In describing efforts to mitigate selection biases in non-probability samples, Münnich, statistics professor and Chairman of the German Statistical Society, states that difficulties arise in regard to the question of whether one can compensate or not, as well as which methods are used to do so. And that, he adds, "is, of course, a bit, let's say, between trade secrets and magic" (Münnich, Interview, Translation).

In the field of sampling, there are, however, suggestions, as to what might count as a successful puzzle solution for the case of non-probability sampling. Wang *et al.* (2015), for instance, were able to produce forecasts in line with those from leading polling analysis, based

on daily voter intention polls for the 2012 US presidential election conducted on the Xbox gaming platform. The sample one gets from such a platform is highly biased in terms of age and gender, structurally, a highly skewed sample that reminds of early straw polls. Wang *et al.* (2015) used poststratification to mimic a representative sample of likely voters. The idea behind stratification is to subdivide a heterogenous population into separate and homogeneous units, called strata. While pre-stratification means that strata are created before the sampling begins, post-stratification describes the creation of strata during or after the sampling has taken place. Their conclusion is as follows:

> [A]s response rates have declined over the past several decades, the statistical benefits of representative sampling have diminished. […] We conclude by arguing that non-representative polling shows promise not only for election forecasting, but also for measuring public opinion on a broad range of social, economic and cultural issues (Wang *et al.*, 2015, p. 980).

The argument that this serves as an example of a successful puzzle-solution for non-probability samples is, however, not shared among all members of the survey community. Carina Cornesse evaluates this as such:

> The majority of empirical findings that we currently have clearly say that non-probability samples cannot achieve a good representation of the population. […] There are studies that show that you can accurately calculate election predictions from X-Box data […] if you use the appropriate statistical weighting procedure. But people sit there for months and think about the perfect weighting procedure for exactly this one variable. And so, in individual cases, it may be possible to make a correct representation of the population from non-probability samples. But to say that we have a non-probability sample, and we also measure everything possible by means of this sample, best of all in real-time, perhaps weighting or quoting for age, gender and education and thus we have covered all the relevant [laughs] population characteristics, that is actually absurd and lacks any scientific knowledge, is not based on any comprehensible statistical theory (Cornesse, Interview, Translation).

Being scientific is here equated with statistical theory, similarly, rejecting non-probability methods as a paradigm in its own right. To others, the development towards non-probability sampling is part of the normal changes a scientific field undergoes. Agreeing with points raised by other participants over the general superiority of random sampling, Rendtel states that "it's

like the book and the modern media, they also say it's the death of the book […], but the book still survived". In the same way, the random sample will survive "for highly qualified samples" (Rendtel, Interview, Translation). But he puts the current rise in non-probability sampling in a position comparable to earlier episodes in the history of sampling, which we have explored earlier, where different modes coexist with each other. But with regard to paradigm shifts, he is quite clear in his conviction that we experience a gradual shift, which is comparable to previous shifts in the field. Regarding the future of non-probability sampling, he states as follows:

> And I am deeply convinced that this form of surveying will continue to exist. It is not possible to stop it by the verdict of a very convinced defender of design-based methods, no! […] But at some point […], there is a saying that impressed me that mostly paradigms don't change because people have been convinced, but because the followers of a certain paradigm simply retire, die and others grow up [laughs]. But, yes, as I said, I already am an emeritus and that's the way things are and that is perhaps also quite good, that new people who simply have a different attitude to these things grow up (Rendtel, Interview, Translation).

Rendtel also implicitly endorses Kuhn, however other aspects of his work. While Schnell drew on Kuhn's work regarding the question of what constitutes a paradigm, Rendtel draws on Kuhn's notion of scientific revolutions. One could assume that Rendtel stumbled upon Kuhn's quotation of Max Planck's *Scientific Autobiography*, which resembles his statement in central ways: "[A] new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it" (Planck, 1949, pp. 33–34; Planck in: Kuhn, 1996, p. 151). The notion of paradigm shifts has also been politicized by Harris Black, who is quoted in the Wall Street Journal in 1999 with the statement that "[i]t's a funny thing about scientific revolutions[.] People who are defenders of the old paradigm generally don't change. They are just replaced by people who embrace new ideas" (Black in: Simons, 1999).

Whether or not this holds true for the case of probability vs. non-probability sampling, those considerations lead us to the question of how boundaries are further negotiated in the field of

survey research, and how stability is being produced by those who continue to defend either of both approaches. Before the subsequent subchapters will explore those mechanisms more closely, I would like to highlight another aspect as to why this controversy has, for quite some time, been stuck in a deadlock, as to why closure seems to be so difficult to achieve.

7.3.2 The Glasshouse Metaphor and the Struggle over Non-Response

The controversy displays many aspects where the same scientific arguments are mobilized from different sides of the controversy, an observation that emphasizes questions around interpretative sovereignty, rather than being right or wrong. Despite it being a controversy that plays out between different fields of survey research, e.g., academic vs. commercial research, we often have to deal with disagreement between epistemic peers, who are people with an equally good epistemic position with respect to a certain topic or issue. Whereas the philosophical literature (see e.g., Christensen and Lackey, 2016) on peer disagreement focuses on questions as to how one should react to cases of such epistemic disagreements, when one should defer to other opinions, or how to identify who the real experts are in areas of disagreement, in this section, I aim to explore the situation of epistemic disagreement from a different perspective, leaving valuations around who the better knower is aside. Instead, I focus on what needs to be in place to gain and maintain a position of counting as the 'better' knower.

A central aspect that we have already come across in chapter 6 is cases in which members of the polling and survey community have diverging views on fundamental issues or accuse each other, using the same arguments. We can call this theme the glasshouse metaphor. The metaphor of the glasshouse was raised in several of the interviews, mainly pointing to mutual accusations of a lack of representativity and in doing so, discursively claiming 'interpretive sovereignty' over, for example, non-response rates. Carina Cornesse highlighted this aspect, emphasising that most companies in the field do not fulfil the promises they make.

I am also not convinced by the sincerity of many of the accusations made. I believe that many of the accusations are true, but they are often voiced by people who give me the impression that if you live in a glass house, you shouldn't throw stones. So, I think there is a lot of insincerity in this dispute. Because many people accuse Civey, for example, of propagating representativeness, even though they don't achieve it themselves, and I think that's a valid argument, I think that's true, but it's just as true for almost all the other competitors. It's true for YouGov, for example. So, if you go to the website of YouGov, they also say you have representative data and that is just not justifiable from my point of view (Cornesse, Interview, Translation).

A similar statement has been made by Gerd Bosbach, agreeing with Civey that the established institutes are "clearly reaching their limits" and thus welcoming "doing things different", adding, however, "to draw representative conclusions, and that that's what Civey does, I think is simply critical" (Bosbach, Interview, Translation). Ulrich Rendtel, who argues in a similar, yet opposite vein, also mobilizes the Glasshouse metaphor: "But ultimately, they [the traditional polling companies] are sitting in a glass house themselves. With a response rate of 5%, that's a laugh" (Rendtel, Interview, Translation).

On the one hand, we have to deal with the accusation of not being representative due to a flawed methodology, but, on the other hand, it is due to the high non-response that renders a sound methodology impractical. The issue of non-response and missing data has been taken up by interviewees from both sides, with quite starkly different connotations. Rainer Schnell, echoing the argument made by Jon Krosnick, as presented earlier, distinguishes between Missing at Random and Not Missing at Random.[26]

---

[26]   According to van Belle (2008), statistical terminology defines three kinds of missing data, which each refer to one type of unknowns : 'Missing Completely at Random' (MCAR): MCAR describes situations in which missing data have no systematic relationships with other variables. This means that non-observation has nothing to do with the person or object under study. For example, a survey gets lost by sending it through the post or the probability that data on income is missing is the same for all individuals. To van Belle (2008, p. 185), this is the "most benign situation […] [and] means that the available data can be analyzed without worrying about bias". 'Missing at Random' (MAR): MAR describes situations in which missing data occurs only conditionally at random. This means that non-observation has something to do with the person or object under study, which, however, can be controlled for. Even though data are missing, their missingness can be explained based on observed variables. For example, females might be less likely to fill a survey on

I would say the empirical evidence that there is systematic non-response is very thin. If we say that technically, non-responses is, for social science topics only, according to all empirical evidence […] missing at random and that means we can correct for that. If you have something that is not missing at random, then you can no longer correct (Schnell, Interview, Translation).

Schnell here refers to cases in which participants are not missing at random, also known as NMAR. This is the case when one aims to measure something, which is at the same time the reason for non-response. If age is what I am interested in and people are missing because of their age, I cannot correct for the non-response. Such cases are different from Missing at Random, as the reason why certain people are missing has nothing to do with the characteristics of the people themselves. The missingness is not due to, to stay with the example, age, but due to other observable factors, for which it is possible to adjust. To Rendtel, the problem lies somewhere different:

But now comes the practice. If you do a survey like this and you will find that of the 2,000 people who should answer, perhaps only half answer, or in the meantime only […] 5%. Then you don't know the selection probability anymore. And then you have to make assumptions about which characteristics are important for someone to have participated in such a voluntary survey. And there you end up again with a statistical assumption, which in this case is not so easy […], because you do not know the characteristics of the people who do not want to participate […]. That is a tricky situation (Rendtel, Interview, Translation).

Rendtel diverges from Schnell in the sense that Schnell assumes the evidence for systematic non-response to be very thin, a point that has also been made by Krosnick. Rendtel's point is, how could be it be different, also echoed by Douglas Rivers who commented on it as follows: "But let's be clear, you don't have a probability sample at that point and your inferences are

---

stress, but the fact that they don't do so has nothing to do with their level of stress. Or if the probability that data on income is missing differs according to age or gender of the respondent but not according to the income of respondents of the same age or gender. 'Non-Ignorable Missingness' or 'Not Missing at Random' (NMAR): NMAR describes situations in which the value of a variable is missing related to the reason why it is missing. For example, individuals filling out a survey on (illegal) drug usage might leave fields empty out of fear of prosecution. Or if the probability that data on income is missing varies according to income for respondents of the same age or gender. The fields are thus not left empty randomly but on purpose. In such cases modelling and adjustment becomes necessary.

going to be based on modeling assumptions, not on the procedures that you follow" (Rivers, Interview). What Rivers and Rendtel state is that ultimately, in our time, every supposedly probability sample in fact is a non-probability, or a model-based sample, because the participating respondents no longer represent the population. To make up for this gap, probability samples rely on models on how their sample differs from the general population.

In contextualizing this controversy, and in referring to Ulrich Rendtel and Rainer Schnell, Ralf Münnich, stated that both positions we find here are, despite being both true, extreme positions.

> In this respect, I also understand what Ulrich Rendtel says, who is also a statistician – and I'm prepared to pay admission when Ulrich Rendtel and Rainer Schnell discuss this […] [I]t's certainly the case that both represent extreme positions, Ulrich is one of those who say that it's not so simple with the many non-responses. Rainer Schnell argues that we first need a probabilistic sample as a basis. And both are correct. Of course, the question is in what order of magnitude are we talking? (Münnich, Interview, Translation).

Using the metaphor of the glass house points to the current shortcomings of both sampling frameworks, and, as we will see in the following sections, it is not necessarily scientific rigour, but to a large extent also due to other social practices through which either of those approaches are being justified. The question thus becomes what proponents of these approaches mobilise in order to assemble legitimacy. By this I mean the rhetorical or cognitive figures that they apply in order to create boundaries and establish themselves as representing a scientific and the other as an unscientific approach. Taking a more positive interpretation of the metaphor of the glasshouse, we can, again, point to the issue of transparency as a prerequisite for both approaches to meet.

7.3.3 Shifting Boundaries and Purity Claims

Throughout this section, we remain within the framework of boundary work. To briefly recapitulate, boundary work is a term coined by Gieryn to analyse how science was able to gain

authority, as compared to other epistemic practices, such as religion. Other accounts, such as suggested by Jasanoff (1995), also describe disputes over authority and recognition within a scientific community, such as in this case, polling and survey sampling. In the following, we can observe notions of purity and impurity as attributes being assigned to research programmes, which serve as a marker between what counts as legitimate. One way to further conceptualize this has been suggested by Swedlow (2007), drawing on Bloor (1991), Gieryn (1999) and Douglas (2005), including pollution and purity claims as a way to "create and maintain boundaries *within* science, allowing distinctions to be made *among* scientists" (Swedlow, 2007, p. 636 italics in original). In this way, boundary work is used to contrast "sacred, pure versions of scientific activity with profane, polluted ones" (Swedlow, 2007, p. 636). This distinction between purity and impurity can also be found in the discussion around probability and non-probability sampling.

7.3.3.1 Stabilising Existing Boundaries

What we have seen throughout this and the previous chapter, different actors in the field engage in practices to maintain and further stabilise the boundaries that have governed survey research and polling since the 1940s. Others, on the other hand, engage in practices to extend or shift those very boundaries. Paradigmatically, Rainer Schnell, found strong words, comparing non-probability sampling and Civey's methodology with pseudo-science and homoeopathy, ultimately stating that they do not follow the rules of science, thus pushing them outside the boundary of the scientific terrain. Schnell's evaluation reminds of Gieryn's notion of boundary work as the making of a distinction between scientific and "'non-scientific' intellectual or professional activities" (Gieryn, 1983, p. 791).

> There is no empirical evidence that it works. The procedure is not documented. It is not comprehensible. And the fascinating thing is that they then find lunatics who believe it. It's just like homoeopathy. Yes, you have someone there who is fooling you. And of course, there are people who believe it. It's cheap. It's fast. It's sexy. But the problem is, they don't explain

it to you so that it's comprehensible. There's no reproductive proof. There's no prognostic power. And the interesting thing, the really sociologically interesting thing is that there are people who believe them, so it's really believing. There is no mathematics behind it, nor is there empirically convincing evidence that it works. And you see this persistence despite the absence of evidence with homoeopaths for, I don't know, 120 years, and why should it go faster in statistics? (Schnell, Interview, Translation).

Depicting those who consider the numbers produced by Civey as sufficiently plausible for the respective purposes as lunatics seems to be the strongest antonym to science and scientific method. In Schnell's view, their legitimacy needs to be challenged, as what they do does not take place in the scientific arena. They are, so to speak, not on equal footing with the legitimate and scientific practices of knowledge creation. His comment ultimately refers to practices such as homeopathy to illustrate the absence of evidence he sees in the work of Civey. In this statement, he also highlights how there are different logics operating in which either of those approaches are more sensible. From a logic of production, it makes sense, as it is cheap, fast (and sexy), but in a scientific logic the interviewee would expect more, which is prognostic power and reproductive proof. His argument condenses into a commitment to a demarcation between science and non-science: "And from that point of view, what Civey is producing is incredibly fascinating, as a phenomenon in the sociology of knowledge, it's incredibly fascinating. But that is not science. It's the opposite of science. It is counter-enlightenment." Furthermore, he distinguishes between faith and knowledge, stating that "Civey believes, classical inferential statistics does not believe" (Schnell, Interview, Translation). Those claims can, following Swedlow's (2007) conceptualizing, be understood as pollution claims, understood as practices of "persuading relevant others that rivals' factual claims are illusions or lies, the result of ulterior motives or faith, for example, not scientific investigation" (Swedlow, 2007, p. 635).

This antagonism of science and non-science and thus pollution claims run through many interviews conducted with survey researchers and pollsters. It is mainly raised by those who defend the paradigm of probability-based sampling, serving as a rationale to delegitimize Civey's work in attributing a misorientation in their epistemic practices. By framing the discussion in

this manner, Schnell argues that if there exists only one legitimate form of science, any alternative endeavor deviating from these established principles cannot be categorized as science but rather as a commercial activity masquerading as science to garner a following. Once this semantic position is established, Schnell employs linguistic devices such as "magic tricks" (analogous to marketing tricks) to engage in a discourse that portrays the followers of Civey as deluded. He contends that a "sociology of knowledge" ought to uncover and expose these delusions.

To give another example, the following quote by Krosnick goes in a similar direction, opening a divide between scientific and non-scientific methods. It also states how the public plays a further role, first as one that is being deceived, but also as one that wilfully ignores the scientific basis behind polls and surveys, thus allowing the pollution to take place: He states that nobody "is telling the public, these are scientific methods, you should believe these, these are non-scientific methods, you should not believe these. […] Lots of people don't want to know the difference" (Krosnick, Interview). The distinction between science and non-science here serves as a way to tighten the borders surrounding the practice of survey research and polling. Only what counts as science, which is probability sampling, can legitimately be used as a reliable way to understand society, inform political decisions or serve marketing purposes. Similar practices of boundary work, however, are also employed from the other side. There, we identify a counter rhetoric, one that aims to extend the boundaries to include non-probability surveys.

7.3.3.2 Extending Existing Boundaries

While pollution claims have been raised by defenders of the probability sampling paradigm, others counter those claims by either positing an always-already existing impurity or demanding a re-evaluation of what counts as pure. Under 6.3.1.3, I already explored this notion of an always-already existing impurity, summarised in the dictum of 'we have never been random'. The argument was that pure probability sampling has always been an illusion and that we should not

fixate on it if we want to make progress in the field. We also already came across people like Douglas Rivers, members of Civey and others, who engage in what I earlier called a rhetoric of inclusion, aiming to shift the boundaries in which legitimate social research takes place in a way that it includes non-traditional, non-probability-based approaches. To Sabine Zinn, vice-director at the Socio-Economic Panel (SOEP) of Germany, non-probability sampling constitutes a new approach that needs to be given its due space, without being pre-judged. Countering impurity claims such as those just presented, she argues as follows:

> I don't know why we have to demonise it, it's not to be demonised. […] I don't know where that comes from. Maybe it's some kind of old professorial textbook attitude. I don't know. You don't always have to demonise things; you also have to see the advantages of them. And that's exactly it. And the people who collect such non-probabilistic samples should perhaps be told that they shouldn't claim that they can now explain the whole world. It is a kind of learning, a mutual learning, and that also takes place (Zinn, Interview, Translation).

Zinn suggests both approaches have their advantages, advocating for a balanced approach that acknowledges both advantages and limitations while encouraging a mutual learning process among researchers.

Others shift the perspective, stating that error one gets through sampling is rather marginal, while maintaining the argument of a general superiority of randomisation. To Saris, the founder of the Telepanel, this looks as follows:

> You can defend the random sampling procedure […] [,] [but] in principle, there is no defence for any other procedure, which is just arbitrary asking people to participate. There is no proof that you get the proper results. On the other hand, if the samples are very large, you can also re-weight the samples and you get an approximately, well, reasonable results and I think that's my reservation for this thing. In the beginning when we started, we were also in the mood that we saw that we have to take care that they are representative. But as I say, I mean, I am convinced that the errors from the questions are much larger than the errors by the statistical procedures (Saris, Interview).

What we encounter here is also an attempt to extend the boundary of survey sampling, to also include non-traditional methods. Zinn and Saris share a similar stance, yet their arguments operate on distinct levels. Zinn primarily offers a meta-observation, cautioning against the

demonization of certain practices and impurity claims, in order to safeguard the potential for innovative heterodox knowledge formation. On the other hand, Saris approaches the issue from a distinct perspective, particularly focusing on the statistical impurity that might be deemed forgivable. Saris suggests that such impurity is of lesser significance compared to the strategies employed in formulating survey questions.

7.3.4   The Role of Conventions

Questions around sampling take place within different realms, official statistics, academia and commercial polling. Those different realms, it can be said, follow different logics imposed by the field in which they operate. To Bourdieu, the social world consists of different 'fields', with their own sets of rules, and knowledge practices. The concept of field originated in the theory of magnetism, referring to the range of action of charge. The core of the field theory was the concept of the force that acts on bodies over a certain distance and thus determines the nature and limit of the field. Bourdieu's fields are structured areas of social life that afford socially patterned activities. We may also conceptualise the survey landscape as a field, defined by disciplinarily and professional criteria. In adapting Bourdieu's definition of the juridical field, we can call the survey field the "site of a competition for monopoly" (Bourdieu, 1987, p. 817) over the right to produce representations. A strong force that structures the field of survey sampling is that of randomisation. Randomisation is thus a central force many participants in the field drive towards. As we have seen in the previous sections, Schnell, for instance, described his commitment to the probability-based paradigm in demarcation to proponents of the non-probability-based paradigm (in this case Civey), as "believers" in the sense that they are followers of an ultimately unscientific form of knowledge creation, based on unverifiable certainties of faith, but not found in a reason-based, rational scientific ideal. This line of thinking seems to be based on an ultimately positivist belief in science, which suppresses the conventional character inherent in scientific practice, that statistical data and facts are always made against the background of

guiding assumptions, concepts and theories, that truth claims are only correct against the background of a paradigm supported by knowledge bearers and their institutions. As we have seen in chapter 3, the paradigm of probability sampling also grew historically and had to establish itself as an accepted way of doing science.

Following this tradition, the economics of convention offers a perspective that further helps to illuminate the contrasting views. What can be observed here, looking from different conceptual lenses, is the collision of different 'data worlds'. In the tradition of the 'economics of convention' (EC), Diaz-Bone and Horvath introduce the concept of 'data worlds' to explore "how data infrastructures (historically formed cognitive and organizational frameworks and resources that structure the production, distribution, and usage of statistical data) are anchored and embedded in social rationalities ('conventions')" (Diaz-Bone and Horvath, 2021, p. 220). Those "worlds of data production vary in their methodological cultures, their epistemic values, their quality criteria and their collective understanding of and relation to the 'common good'", implying a "plurality of co-existing ways of producing and using numeric representations of social phenomena" (Diaz-Bone and Horvath, 2021, p. 220). The co-existence of different paradigms in the field throughout different time periods can be explained by referring to the different conventions underlying the different epistemic practices. Thus, probability sampling, for instance, continues to remain a standard for high-quality research.

If we understand the different logics of data production, such as academic or market-oriented, as different data worlds, it can be argued that they follow different conventions when it comes to the making of statistical knowledge. Those views align with general statements on the nature of statistical knowledge made by my interviewees. The general point made by British statistician David Spiegelhalter resonates in interesting ways with the presented perspectives and the assumptions underlying the economies of convention approach.

So, statisticians have always known that numbers are constructed. The way in which they collected the definitions that are given is absolutely essential to their understanding. And […] the labels assigned to events is always a matter of judgement and choice […]. And so, every statistic is to some extent constructed. It doesn't mean it's made up. […] [B]ut it does mean that the categories in which things are placed is as a matter of judgement. And they've also known that statistical models are, to use the cliche, all wrong. All models are wrong, but some are useful, a phrase invented by a British statistician, George Box. Now, that is completely familiar within the statistical world. We know this. We know that our data is deeply limited, but it's very valuable. We know our models are all wrong, but they can be very valuable (Spiegelhalter, Interview).

In the context of what has been stated so far in this section, we can say that the survey landscape consists of different conventions that make up different survey worlds, which we might relate to different business models or areas of application. Ralf Münnich, for instance, told me that at the core of the conflict between probability and non-probability sample surveys, is a conflict between two business models:

You have to say that there are two business models that come together. There are the market and opinion research institutes that work with classical probability samples and there are, so to speak, these Internet start-ups that deal with non-probability samples. And of course, in the end, both groups want to sell their products. And the question is, obviously, which one is the best? And I would like to bring in one of these aspects right from the start, which I have already mentioned several times. For me, of course, it's about what a statistic is used for. And at the moment when, let's say, political actions are derived from them, I have the demand of high-quality statistics. And there is no free lunch, one simply has to say (Münnich, Interview, Translation).

What Münnich defines as different business models, is similar to what we can call survey worlds, containing different orders of justification (see also: Vogel, 2019). What is interesting in Münnich's statement is that he takes the commercial aspect of sampling for granted in saying that "of course, in the end, both groups want to sell their products". The fact that survey research had entwined into the logic of capitalism, however, marks a particular historical development and has, for instance, not been part of any discussions on sampling during the early years of the ISI. The quality of a representation thus does not solely lie in the statistic or representation itself but can only be evaluated in relation to a particular convention and usage. Statistics produced in a capitalist logic of production are thus not solely to be evaluated based on accuracy, but on

whether or not they serve a particular goal. In cases in which fast, rough estimates are enough, other quality criteria might be applied compared to those policy-guiding situations, in which high-quality surveys are necessary. Those different survey worlds can thus be justified in different contexts and applications:

> And these are two, I would say, quality features that are absolutely antipodal [...]. This means that if you want to have information very quickly, you can of course obtain it very well with the modern platforms [...]. But you then have the difficulty that you can no longer quantify the accuracy and are, so to speak, poking in the fog. [...] For many questions, this may work quite well, but there may also be questions where this does not work well. And the moment a renowned ministry jumps on it and wants to have data from them, for example, to change standard rates, when that comes, I'm the first to stand there and say no way, that's not possible at all (Münnich, Interview, Translation).

The central argument here is that different survey worlds do not only differ in terms of their methodologies but that those worlds are based upon different knowledge systems. The production of data within those data worlds is only "possible on the basis of conventions of measurement"; on "[w]hat counts as relevant, acceptable, and fair process of quantification" (Diaz-Bone and Horvath, 2021, p. 221). The case of the controversy between probability and non-probability sampling might thus be understood as a controversy over the acceptable convention of measurement. In this sense, statistical data is always built upon social, moral, and political orders.

Those social, moral, and political orders, however, mainly become visible in cases in which conventions and practices of representation break down, i.e., situations in which their predictive stability becomes fragile. Those cases can be observed in the contemporary controversies, but also in the context of the 1936 and 1948 US presidential elections, during which observers and participants become aware of how numbers always depend on the conventions established to bring them about. Furthermore, the different positions in this controversy ultimately serve as empirical substantiation for the perspectivity, historicity and relationality of scientific knowledge, including statistical knowledge, a position that has not always been taken for

granted. While it is almost commonsensical to view scientific explanation as "always located in, and shaped by, the specific historical, political, economic and social contexts within which they are proposed and developed" (Mahfoud, McLean and Rose, 2017, p. xii), Karl Mannheim, for instance, as we have seen in the literature review, excluded mathematical knowledge from such characterizations. To him, the natural sciences and mathematics were cognitive systems that, for reasons of principle, defy sociological scrutiny. This epistemological exceptionalism in regards to the natural sciences was not only held by Mannheim, but commonsensical among sociologists by this time, a situation that changed in the course of what can be termed the anti-positivist turn in the philosophy and sociology of science, associated with scholars such as Paul Feyerabend, Thomas Kuhn or Imre Lakatos (see Heintz, 1993).

7.4 Conclusion: From Legitimising to Forging

What we can see is that certain developments from the past decade, such as missed election forecasts and a general increase in distrust towards polling data led to a loss in trust and a loss in authority towards the producers of statistical knowledge. While, on the one side, established statistical authorities have lost their power in determining the will of the public, others have, to some extent, taken up those spaces, aiming to monopolise their access to the social world. With, at the moment, nobody fully incorporating this position, Will Davies' statement in the Guardian describes the situation quite accurately: "With the authority of statistics waning, and nothing stepping into the public sphere to replace it, people can live in whatever imagined community they feel most aligned to and willing to believe in" (Davies, 2017). Quantification holds a crucial role in the constitution of *imagined communities*, as "statistical categorizations both reflect and affect the structural divisions of societies" (Simon, 2012, p. 1368) and numbers allow to "anchor this imagination in something tangible" (Davies, 2020, p. 96). What we observe here is not only a competition of facts that must now "compete with other conflicting 'facts'" (Davies, 2020, p. 55), but it also highlights how difficult it is to maintain a shared basis for matters of public

concern. This becomes problematic as the decline in the authoritative nature of numbers and statistics can be understood as a decline of a shared notion of reality. Such a shared notion of reality is necessary to measure factual claims against reality so that they are not being reduced to mere interested-bound statements. This is also in line with what Münnich worries about:

> I think statistics suffer from the fact that everyone thinks they can do it, because everyone thinks they can deal with numbers and understand a table and a graph. And then, of course, many people have already realised that information is a means of persuasion and of getting their own opinion across, so to speak. And then we very quickly reach the point that this is of course misused at that moment, which is why I am a very strong advocate of open and reproducible research or data, and I would also like the data that the various political parties use to be comprehensible (Münnich, Interview, Translation).

The issue that different conventions of measurement bring about different representations of the social world highlights a crucial aspect, namely that the social realities measured by polls are not mere representations. The act of representing brings, to some extent, those realities into being, a notion that we will explore in greater depth throughout the next chapter. To capture this more thoroughly, the next chapter will work conceptually on what I term the power to forge stable statistical chains, adhering to a conventional approach to scientific practice, while upholding the dimension of power and the performative character of public opinion as the result of polls. Those forms of power struggle have been prevalent in early discussions at the ISI.

Similarly to now, as this chapter has shown, debates concerning the authority to establish industry standards and to construct authoritative representations of the social world persist. In this process, organizations, companies, and scholarly societies engage in practices aimed at demarcating the lines between legitimate and illegitimate methodologies, thereby discerning between credible and discredited scientific knowledge. In this sense, I will not only look at the conventions within the field of polling but into the broader landscape in which those practices are located. What the next chapter ultimately states is that looking at the conventions of how

statistical chains are made is insufficient to explain their power to represent, following both,

realist and constructivist understandings of the term.

CHAPTER 8

8   From Legitimising to Forging

8.1   Introduction

Throughout the previous chapters, I presented two central ideas and findings, which I would like to briefly recapitulate. I have shown how the history of sampling is strongly entangled with the history of polling, producing an understanding of public opinion as a statistical aggregate, established through the use and implementation of the sample survey. The result of this development was the establishment of a convention on how to best "know" a population and their opinions. The development of this idea of public opinion gave pollsters and polling companies a particular role and a monopoly over access to and the representation of public opinion. I then showed how in recent years, especially with the decline of landline phones, the relationship between public opinion and the random sample started to disentangle, giving rise to new methods and agents claiming their part in accessing public opinion for themselves. This situation led to strong controversies between the more traditional and established pollsters and new actors in the field, drawing on non-traditional methods. What becomes clear in this controversy is that, as should have become clear throughout the last three chapters, the rise of new methods for accessing and analysing data led to divide and polarisation, and not necessarily to mutual attempts in paving new directions. This controversy over method is also a struggle over the sovereignty to speak in the name of the public and their opinion, re-opening not only an old debate in the field of sampling but also the question of who holds the power over the delineation of the truth of public opinion.

But what does it take to establish a public by their opinions? While the previous chapter looked into the question of how practices of quantitation in the field of polling and survey

research are negotiated, this chapter wishes to look at what the results of those quantifications do, how they are distributed and stabilized to become acknowledged and commonsensical representations of the public and their collective sentiments. Therefore, this chapter looks at the ontological politics, i.e. the practices that bring certain realities into being, involved in survey research and polling. Through several steps, I will introduce the concept of the *power to forge stable statistical chains*, drawing on recent work that applies earlier convention theory to survey research, Actor-Network Theory, Agential Realism and their relevance to discussions of power. Positing that pollsters shaped the concept of public opinion to equate the statistical aggregation of private opinions, measured through particular data collection methods, public opinion is rendered an artefact, created and influenced by the questions asked, the inferences made, and the way the results are presented. By integrating these perspectives, I will explore how public opinion is shaped internally, by conventions around methodological and technical decisions, as well as externally, through the landscape in which the distribution, production and valuation of those representations take place. Drawing on Latour, one might say that the question guiding this chapter is how representations become "immutable mobiles". It is about how they become invested with authority to ensure its circulation, dissemination and treatment as an accurate and stable representation of public opinion.

In order to arrive at the conceptual framework of the power to forge stable statistical chains, I will proceed as follows: In the first step, I will begin with an elaboration on the notion of ontological politics and a first approximation to the concept of *forging*. In the second step, I will look into different cases and circumstances from the field of polling and survey research in which an immediate relation between methodological decisions and measured public opinion can be observed. Therefore, I will explore questions around matters such as fictitious issues, sampling, question-wording and the dissemination of polls. In a subsequent discussion, I will briefly explore three polling controversies, involving the scandal around manipulated opinion polls

associated with former Austrian chancellor Sebastian Kurz, the fabrication of a poll purporting that American musician Kid Rock was the top choice for a senate seat in Michigan, and the notorious relationship between former American president Donald J. Trump and the polls. The analysis of these cases underscores the limitations of the conventionalist approach to quantification when it comes to explaining the formation and sustenance of publics through surveys and polls.

8.2 On Ontological Politics

8.2.1 Making Dominant Representations

8.2.1.1 The Sample–Population Nexus

Public opinion is usually understood as what links the populace and the elected government and political leadership. The dependence of the coherence and maintenance of publics has been thought of in different ways. As Perrin and McFarland (2011) summarize, it has been positioned in the salons and coffee shops (Habermas, 1999), representative claims (Saward, 2006), print capitalism (Anderson, 2016), electoral rituals, or media representations (Warner, 2005). Those views have in common that publics and their opinions are not immediately observable but depend on and are upheld by social and technical practices and apparatuses. The sample survey therefore became the dominant instrument to create representations of people and their opinions, rendering public opinion the result of what public opinion polls measure. It has gained a hegemonic status (Page and Shapiro, 1992; Fishkin, 1997; Althaus, 2003; Sturgis and Smith, 2010; Goldthorpe, 2021) as a tool to access the will of the people. In this sense, the responses of a small number of people enter into a *standing for* (Pitkin, 1967) relationship with a population. But how is this *standing for* relationship to be perceived? Therefore, the following will explore what I term the sample-population nexus, whereas sample stands for the observed individuals and population for the calculated population that one seeks to know further. This view turns the questions from the previous chapters on its head in that it no longer asks how to move from a population to a sample,

but how to move from a sample to a synthetic representation of a population. If there is no such thing as a pure and unfiltered representation, the *standing for* relationship between a sample and a population must be thought of in considering how to move from a sample to a population, of which there are always more than one. Is the mere act of polling enough to establish a people by their opinions? What else is needed to forge this sample-population nexus?

8.2.1.2   The Partiality of Representations

From a methodological standpoint, representing public opinion and social reality adequately and without large distortion is a difficult undertaking. From a constructivist theoretical perspective, it is an impossible undertaking in that public opinion only exists as its statistical representation, a representation whose relationship to the represented is one that is mediated through conventions, power and politics. This relationship can be further illuminated by Ernesto Laclau. For Laclau, the conditions for a perfect representation, i.e. when "the act of representation is totally transparent" and the "opaqueness inherent in any substitution and embodiment" (Laclau, 2007, p. 97) is reduced to a minimum, never actually obtains.

The reason why such a perfect representation is impossible is due to the inherent logic in the process of representation: if somebody needs to be represented, their basic identity is taken from one place to another, through which the identity of the represented is always incomplete, as the relation of representation can only be a supplement. An interest that is being represented e.g., in parliament by a deputy, always undergoes transformation, as "the representative *inscribes* an interest in a complex reality different from that in which the interest was originally formulated and, in doing so, he or she constructs and transforms that interest" (Laclau, 2007: 98; italics in original). Due to the very logic of representation, there can thus be no pure representation. If the representation could be fully reduced to the represented, it would no longer involve any representation. There must, so to speak, be processes of transformation happening when calculating a population from a sample.

Those "production halls of social facts" (Raphael, 1996, p. 191, Translation) thus bring about populations, rather than just describing them. A compelling demonstration of this performative aspect can be seen in an experiment conducted by The New York Times during the 2016 US presidential elections. In this experiment, four pollsters and their own section, the Upshot, which features articles that blend data visualization with traditional journalistic analysis, were provided with the same raw data consisting of 867 poll responses and were asked to independently estimate the poll's result. Surprisingly, despite using exactly the same data, the five measures produced varying results, resulting in a net five-point difference between their estimates. Polling results, so Cohn, New York Times's chief political analyst, echoing the conventionalist approach, "rely as much on the judgments of pollsters as on the science of survey methodology" (Cohn, 2016).

Considering the importance of public sentiment in the realm of politics, understanding public sentiment is crucial for politicians. It is thus obvious that in the case of various polls on the same topic, one might choose the most favourable one if somebody wants to present themselves as successful and credible. Since public opinion, as measured by polls, holds significant value as an important currency in the political landscape and has developed to become a common topic in news coverage and political discussions, attempts to manipulate, invent, or discredit polls have accompanied polls since its beginning. This becomes especially pressing in times in which the legitimacy of those who speak in the name of public opinion is put into question.

8.2.1.3 Publics and their Multiple Spokespeople

Whilst the demand and importance of surveys have consolidated, the overall survey landscape has changed remarkably throughout the last years. Among many changes is the availability of survey instruments to almost everybody. I would like to emphasize those changes in the survey and polling landscape in reference to a presentation during a workshop on Survey Climate at the University of Kassel in October 2022, given by Don Dillman, Regents Professor at Washington

State University. In this presentation, Dillman presented various factors that led to a new context within the survey landscape. He identifies seven significant factors that I will briefly summarize here. He states that (1) the ability to conduct surveys has become accessible to almost everyone with digital skills, that (2) despite the use of administrative records, surveys have not lost their importance, and that (3) the variety and sources of survey requests in our daily lives have significantly increased compared to the past. Furthermore, (4) the time between recognizing the need for a survey, conducting data collection, and applying results has been greatly reduced, leading to a higher demand for more surveys, alongside (5) changes in the structure of survey requests and related communications. Importantly, (6) the scientific foundations of using sample surveys to estimate population characteristics are being overlooked or considered less relevant for many surveys and (7) the concept of "Fit-for-use" has gained increased significance as a criterion for evaluating and utilizing survey results.[27] Overall, more and more surveys and polls with various methodologies and approaches are conducted and distributed. A recent report by the Pew Research Center, for instance, has shown how, in 2000, around 29 national public pollsters were active in the US, of which nearly all used live telephone interviews; by 2022, the number of active public pollsters had more than doubled and the used methods had diversified (Kennedy, Popky and Keeter, 2023). As a consequence, the proliferation of entities and methods purporting to represent the public's views through polls raises a critical question: which of these representations become stable and durable? Which ones become, so to speak, the true representation of the public and their opinion?

---

[27]   Fit-for-use and fit-for-purpose refers to the suitability and appropriateness of a survey instrument or polling methodology for its intended purpose. Data thus does not need to be generally accurate, but nevertheless serve the purpose of a particular study (see e.g., Dever *et al.*, 2021).

8.2.1.4 Making Immutable

If there are always multiple potential representations, it is about the closure of possibilities for some to actualise. To better understand this, I suggest understanding this closure of possibilities as a form of *forging* and to be more precise, a *forging of statistical chains*. Adding the notion of forging to that of statistical chains is both a critique and an extension of conventionalist approaches to the practice of survey research and polling, in arguing that an analysis into the constitution of representations needs to broaden its gaze towards the landscape in which practices of quantification take place. The word *forging* thereby involves three meanings that all play a role in the subsequent analysis.

1. In the first sense, the notion *to forge* describes the practice of forging metals, such as chains. After forging, a blacksmith achieves strength and stability of the chain in putting it under water after the forging, a technique referred to as quenching.

2. In the second sense, the notion of forging describes the building of strong ties. To say that two friends have forged a close bond means that they developed a strong friendship or friendship network.

3. In the third sense, the notion to forge describes the malicious practice of forgery. A forgery can be the creation of a copy of a piece of art, one that comes with various degrees of accuracy. A forgery resembles the real thing whilst not being it.

All those aspects of the notion of *forging* involve practices of stabilisation and come with obstacles, power and efforts - they all involve practices of producing stability. A forgery must not be detected, a chain must not break, and a bond needs to be maintained to remain strong over time. In all those cases, stability needs to be inscribed; their reality must be upheld by various mechanisms. Following the making of polls through those lenses highlights the forging power of polls in the creation of publics. In the same way as one can forge a strong bond, polls can lead to the forging of stable identities and play a role in "making up people" (Hacking, 1996, 2013).

The concept of forging can further be understood in relation to Latour's notion of "immutable mobiles", which he defines as "objects which have the properties of being mobile but also immutable, presentable, readable and combinable with one another" (Latour, 1986, p. 7). The concept of immutable mobiles focuses on the question of how objects are transported from one place to another through processes of translation. The emphasis thereby shifts away from merely representing something through another medium or translating it into different forms. Instead, it centres on the essential material and technical requirements necessary for making a leap between registers or achieving spatiotemporal abstraction from the concrete object. Examples of immutable mobiles are maps, artefacts or diagrams in that they allow one to go out and collect things in the world, but to "come back with the 'things'[,] […] [which] have to be able to withstand the return trip without withering away." Those collected things are then made "presentable all at once to those you want to convince and who did not go there(Latour, 1986, p. 7). From this perspective, the goal of survey research and opinion polling is the creation of immutable mobiles. Pollsters and survey researcher go out, physically (in face-to-face interviews) or virtually (in telephone or online interviews), to collect opinions of individuals, in order to statistically reassemble them and to make them presentable, ensuring that those who did not collect the opinions themselves can still trust the validity of the reported results. Thus, individual opinions are assembled on a chart, without, obviously, actually moving the holders of those opinions onto the chart. Forging, thus understood, can be seen as a way to make public opinion mobile and immutable. As the following sections will show, however, the notion of forging adds a different spin to the practices described by Latour in that it accounts for the different agendas and interests of the agents involved in the process of creating representations of public opinion. To a great extent, this has to do with (in)visibility and the politics behind practices of making things (in)visible. If representations are, as I argued, never pure, it is about which representations become dominant.

8.2.2 Polls and Ontological Politics

The way polling and public opinion has been presented throughout this dissertation, the representation and the represented do not stand in a relationship with each other, as a 'correspondence theory of truth' might suggest. Polling results are not to be seen as a true and unfiltered representation of a pre-existing public opinion, out there to be discovered. Rather, the truth of public opinion is performed. In adhering to what has been termed the turn to ontology in the science and technology studies (Law and Lien, 2013; Woolgar and Lezaun, 2013), this chapter also emphasises a shift from epistemological to ontological and onto-epistemological questions. In this sense, we may add the notion of 'politics' to that of ontology (see Law, 2002), emphasising the "process of shaping, and the fact that its [reality] character is both open and contested" (Mol, 1999, p. 75). Exploring scientific practices in this way, shows how "reality is historically, culturally and materially located" (Mol, 1999, p. 75), a notion that links to the conceptual lenses offered by agential realism (Barad, 1999). In this sense, the measuring apparatus, the subsample of respondents and the inferred numbers become "ontological entanglements that are (re)configured via their intra-acting performativities" (Dixon-Román, 2016: 163). When we think about measurement in survey research through the lenses of intra-action, its performativity and its constitutive exclusions become visible. In the words of Dixon-Román, if we "think about the survey or measurement questions and their item responses as intra-acting performativities, then this would make the practice of measurement an ontological entanglement with that which it seeks to observe" (Dixon-Román, 2016: 164).

8.2.3 Do Practitioners Agree?

Such a view, even though it might seem odd to members of the survey community, is not detached from views to be found in the field of polling and survey research. The consequences involved in the interrelation between polling results and the conventions through which they are gathered are a central aspect of the practice of polling. Many interviewees mentioned, for

example, aspects in which the question wording plays a crucial role in changing the measured opinion.

> So, if you ask a question, you always get an answer, but you don't know on which it is based. It is just arbitrary what they think just at that moment. […] That's why I mean, I say, well, it depends on the form of the question, if you change the question, you get a different answer (Saris, Interview).

Others engage with this issue in mentioning the notion of truth. The response Drew Linzer gave is emblematic for the overall argument of this chapter. In many cases, there is no truth out there to be measured. The sampling, the questionnaire and its distribution become an onto-epistemological entanglement, through which we can explore the particular epistemic access and its world-making/ performative capacities through which social scientists tap into public opinion.

> And in many cases, what the discrepancy is due to is simply that there is no truth. We're asking people about attitudes where slight variations in question wording will impact what people respond. And it's not because the question wording is quote unquote wrong. It's because there is no truth! There is no fundamental truth in the world that an ideal question can gather (Linzer, Interview).

Those examples highlight a fundamental issue when it comes to polling. In many cases, such as questions around certain political attitudes, especially when not everybody has equal knowledge of the subject matter, there is no truth against which those polling results can be measured. This view might make more sense when referring to American pollster George Gallup again. In chapter 5.3.1, I presented a statement made by Gallup a few months prior to the 1948 US presidential elections and his famous miscalls. In this statement, Gallup refers to the progress the field made and that it is possible to accurately represent public opinion by the means of polls. This statement culminated in his view that poll results might be even more accurate than the outcome of elections. In the aftermath of the 1948 election, Gallup made a slightly less confident statement during a symposium on the question of whether public opinion polls should make election forecasts at all. He stated that polls are always subject to probabilities and also prone to

fail, but that pollsters will give their best and continue to improve their methods. After his statement during this symposium, Gallup was asked a question regarding the purpose of political polling on the background that it wastes a lot of time and energy only in order to see "who is going to get the election when you will know in a few weeks anyway" (Gallup in: Seymour *et al.*, 1949, p. 142). Agreeing with the questioner on the little social value to election forecasting per se, he justified it as a means to test polling methods:

> It has been my experience that we have had a greater urge, we have done more to perfect our methods, we have actually made greater progress because we knew we were going to have to face an acid test, than we probably would have made if we hadn't been making election forecasts. So I think the only justification of an election forecast is to test polling methods (Gallup in: Seymour *et al.*, 1949, p. 143).

Implicit in this statement is the creation of a justification and legitimation for the use of election forecasts in the aftermath of a great failure of the polling industry. It does, however, also highlight the indeterminacy of polling results. Polling results are subject to conventions, which might be strengthened by, for example, testing them based on polls to which there is a truth, such as elections. The existence of a controversy over the right approach is emblematic for this indeterminacy and the ontological politics of polling.

8.2.4   Forging Stable Statistical Chains

8.2.4.1   The Stability of Chains

Those considerations lead me to add one more element to the notion of forging: stability and statistical chains. According to recent work in the tradition of the French economies of convention, "statistical facts are seen as the result of concrete practices of knowledge production that take place in given institutional contexts and are structured by existing orders of (scientific, political, cultural …) knowledge which they build on, contribute to, and transform" (Diaz-Bone and Horvath, 2021, p. 220). This interplay of conventions, actors and technologies is thereby described by the concept of 'statistical chains', understood as "institutionalized social processes

which allow to generate data" (Diaz-Bone and Horvath, 2021, p. 220). Turning to early articulations of conventions, such as those by Desrosières, it is only when certain conventions are routinely engaged in that statistics turns into facts: "Reality appears as the product of a series of material recordings: the more general the recordings – in other words, the more firmly established the conventions of equivalence on which they are founded, as a result of broader investment – the greater the reality of the product" (Desrosières, 1998, p. 12). In other words, the reality and truth of representations comes through their stability.

We can make more sense of this by remembering the discussion under 4.2, in which I emphasised how after the 1948 US presidential election, when a stable network started to disintegrate, the polling industry was able to build much stronger forces in tying the production of their representations to the random sample. In characterizing their methodology as scientific, the polling industry made itself indispensable and became an obligatory passage point for everyone who wants to know public opinion. It is in this sense that we have to understand the stability of statistical chains. Various actors are thereby being enrolled in practices of translation, which is, as Callon nicely described, the "mechanism by which the social and natural worlds progressively take form", leading to a situation in which some entities control others. In what is usually understood as a power relationship in sociological literature, it comes down to the "way in which actors are defined, associated and simultaneously obliged to remain faithful to their alliances" (Callon, 1984, p. 224). Translation occurs when different actors, both human and non-human, come together to form a network and collaborate towards a common goal. In becoming an obligatory passage point for getting knowledge about public opinion, polling practices establish a regime of visibility. The power to forge stable statistical chains is the power to make certain representations (in)visible.

8.2.4.2 Forging as a Regime of Visibility

In Griessl (2023), I explore the truth-making capacity of polling through Rancière's notion of the political difference.[28] In the following, I will provide an overview of the part of my argument that is relevant to discussions of forging. In his oeuvre, Rancière refers to the *political difference*, in which he distinguishes between *politics* and *police*. The difference is in the mode of appearance. Whether a protest expresses politics depends on whether or not it is able to appear in the current regime of the perceptible: "nothing is political in itself. But anything may become political if it gives rise to a meeting of these two logics" (Rancière, 1999, p. 32). Whereas the logic of police refers to the regime of visibility, politics is an interruption into this very regime, it is "an intervention upon the visible and the sayable" (Rancière, Panagia and Bowlby, 2001, p. 21). My discussion of forging refers to Rancière's notion of the police. The forging of stable statistical chains establishes an apparently unchallengeable representation of social realities. It brings about "an uninterrupted count that presents the total of 'public opinion' as identical to the body of the people" (Rancière, 1999, p. 103). The forging of stable statistical chains forecloses the possibility to challenge representations, since the data no longer appears as a doubling or a construction of the facts of the world but an alleged presentation of those very facts. The forging of stable statistical chains makes representations more and more incontestable and immutable. Whereas democracy is always a struggle about who gets to be represented, about who and what is made visible, the dominance of polls and the forging of publics through polls creates a synthetic representation of reality that comes to feel real.

> What in actual fact is this identification of democratic opinion with the system of polls and simulations? It is the absolute removal of the sphere of appearance of the people. In it the

---

[28] My peer-reviwed article 'From Skopein to Scraping: Probability, Agency, and the Politics of Public Opinion Research' has recently been published by Parallax and is concerned with the notion of polling as a form of political participation. The article explores the implications of viewing polling and demos scraping through the lenses of Jacques Rancière.

community is continually presented to itself. In it the people are never again uneven, uncountable, or unpresentable. They are always both totally present and totally absent at once. They are entirely caught in a structure of the visible where everything is on show and where there is thus no longer any place for appearance (Rancière, 1999, p. 103).

In this sense, the regime of public opinion, "as gauged by the poll and of the unending exhibition of the real" is, for Rancière, the "normal form the police in western societies takes" (Rancière, 1999, p. 31). In doing so, the "effectiveness of the sovereign people is exercised as strictly identical to the calculations of a science of the population's opinions, which is the same as saying an immediate unity of science and opinion" (Rancière, 1999, p. 103). The appearance of the people, the political moment in which political agency can be exercised becomes overridden by calculations. Rancière describes democracy as the "kind of community that is defined by the existence of a specific sphere of appearance of the people" (Rancière, 1999, p. 99). The introduction of certain groups or interests into the political sphere is that through which those groups and interests are constituted. In contrast, post-democracy is the "government practice and conceptual legitimization of a democracy *after* the demos, a democracy that has eliminated the appearance, miscount, and dispute of the people" (Rancière, 1999, p. 102). The forging of statistical representations eliminates this very appearance of the people; it posits stability without adequately accounting for how this stability came about. It posits complete and true representation while only being partial.

Those considerations are of crucial relevance for the following two kinds of the ontological politics involved in polling. The examples inherent to polling emphasise the very making of visibility of particular representations through scientific practices. The examples external to the actual practice of polling emphasise attempts to override and manipulate representations, equally playing a role in the regime of the police. What is involved here, when talking about the notion of forging, is the closing of a gap between samples and potential representations. Forging statistical chains means making certain representations immutable and refers to a regime of visibility in which particular representations become stabilized. Furthermore, different agents

and institions that create representations of public opinion might seek to enact different representations, reflect their political-economic interests, also opening the space for potential discussions on the political-economy of polling practices, an interest that is, however, only peripheral to this dissertation. If representation is never pure and always to some extent opaque, the question of which representations become facts, or to put it differently, which representation gains interpretive authority, becomes the locus in which power operates.

8.3 Conducting Polls, Enacting People

Taking that stable representations are the result of a process of forging, I will now look into two kinds of examples of the way how public opinion is established through polls. I, therefore, suggest distinguishing between examples inherent in the practice of polling and cases that are external to its actual practice. What those phenomena unite is that they have to do with the ontological politics of polling in the sense that they enact one particular reality rather than another one, they close the possibility for other realities to actualize. The following examples are about the representation, establishment and potential manipulation of public opinion by means of polls, which lead us to the question of power to shape the public perception of groups and individuals. As we shall see, polls play a role in creating publics that possess different degrees in how they are anchored in reality. The difference is due to the stability of those very practices, the various degrees in the stability of their material recordings.

8.3.1 Inherent to Polling

8.3.1.1 Fictitious Issues

The first topic I would like to consider here is that of fictitious issues, which has to do with the "elicitation of opinions on non-existent or highly obscure pieces of legislation" (Sturgis and Smith, 2010, p. 67). Schumann and Presser (1980), for instance, have shown that 25-30% of respondents offered an opinion to a "highly obscure bill before congress" (Schuman and Presser,

1980, p. 1214), and Bishop *et al.* (1980) have shown that a similar amount of people responded to an invented 'public affairs act', therefore coining the phenomena in question "pseudo-opinions". Those studies resonate with Bourdieu's (1972) critique of polling, in stating that pollsters and survey researchers cannot expect that everyone has an equally well-informed opinion or an opinion at all. In the case of fictitious issues, people offer an opinion about invented issues, issues they must be ignorant about. When polls and surveys are the statistical aggregation of knowledge regarding a topic in question, fictitious polls seem to produce knowledge based on ignorance. They thus produce opinions in favour or opposed to something that does not exist. The later example about Kid Rock and manufactured public support during a senate race points to the same phenomena under opposite assumptions. In contrast to the case of 'pseudo-opinions', this constitutes the case of a 'pseudo-public', the creation of a public out of thin air. In most of those cases, however, the created representations do not display strong material recordings and are thus not firmly anchored. When exploring the making of publics by polls, we can see that in cases such as those, the creation of stability is of crucial importance and yet a highly difficult matter. Something different happens when considering the influence of question wording and sampling, as a way through which different publics might come into being.

8.3.1.2 Question Wording and Sampling

The next category is widely discussed in the survey literature and relates to what has been termed the 'total survey error framework' (e.g., Groves, 2009; Groves and Lyberg, 2010; Lyberg and Weisberg, 2016), linking "survey design, collection, and estimation into the error sources" (Groves and Lyberg, 2010, p. 856). Important to note are the notions of *measurement*, refering to the issue of what it is that a question measures and *representation*, pertaining to the question of how to arrive from a sample to a population. In a paper, titled *The Public Opinion about the EU Can Easily Be Swayed in Different Directions,* Saris (1998b), whose work we already became acquainted with earlier, shows that in situations in which the public does not have enough

information about European policies, "public support for any kind of policy in favour of or against European integration can be swayed rather easily" (Saris, 1998b, p. 406). To be more concrete, Saris could show that small changes in the questions led to quite different results and that it was easy to change public opinion from a nationalistic to a European position and back again. In a similar way we can see the already explored example over the controversy around German football players Mesut Özil and İlkay Gündoğan, and their photoshoot with the Turkish president Erdoğan. Part of this case was that not only different question wordings, but also different sampling techniques enacted completely contrary representations of reality. Those two examples are part of a wide range of examples in the literature, in which the ontological entanglements between questions and responses become intra-acting performativities.[29]

The examples of question wording and sampling aligns with discussions of power as the ability to set and reject industry standards, relating to Thomas Kuhn, to whom power comes in the ability to change the rules "governing the prior practice of normal science" (Kuhn, 1996, p. 7).

8.3.1.3 Selected Polls

The last class of examples is highly intuitive, the case of the selected publication and dissemination of polls. Who decides which polls are being disseminated and which ones are not? The more a poll is shared, talked about and referred to, the more truth is attached to the calculated population in question. This case refers to the power, for instance in the hands of media

---

[29]   In a recent book on polling, British pollster Mark Pack (2022) lists more such cases, such as the case of ComRes, a British polling company that asked in May 2009, if people agree that "Britain should remain a full member of the European Union", a question to which 55% agreed. But when they were asked if they agree or disagree that "Britain should leave the European Union but maintain close trading links", another 55% percent agreed. Despite there being more information provided in the second question, the distribution of either poll has the potential to create publics that are both in favour of the UK remaining in the EU and against it.

companies, to spread particular representations rather than others. The ubiquity of polls and surveys often allows to cherry-pick, thus, enacting certain publics over others.

A related theme concerns the interpretation of polls and how they enable different kinds of publics. Shahrokni (2012) explores this in showing how "opinion polling in Iran was utilized as a political strategy that contributed to the formation of a counter-public, which served as the backbone to the reform movement" (Shahrokni, 2012, p. 201). Shahrokni's argument is that during the reform era in Iran, public opinion polling "became a radical tool in the hands of the reformist to draw differences out of the undifferentiated" and to thus engage in what Bourdieu (2010) terms 'classification struggle'. Reformists aimed to claim an 'official' position in the political field for a counter-public, enabled and constituted by polls. Polling thus became a social movement, resonating with Bruno *et al.*'s (2014) notion of statistical activism: Rather than seeing polls as "conservative mechanisms that serves to control, discipline, persuade, manipulate or fool the people" (Shahrokni, 2012, p. 205), "polling contributed to the formation of a counter-public that would serve as a legitimizing force for reform" (Shahrokni, 2012, p. 216).

8.3.2   The Polling Hinterlands

While the preceding cases pertain to implicit issues within polling practices, the second set of issues broadens the analytical framework to encompass the context in which those practices occur. In this section, I will present three polling controversies from Austria and the US that relate to the production and maintenance of public opinion through polls, showcasing varying degrees of stability in their material recordings. The first case involves a potentially invented poll that claimed American musician Kid Rock was the leading contender for a senate seat in Michigan. The material evidence for this portrayal of public opinion is scarce, making it a case in which the represented public opinion shows little material recordings and therefore minimal stability. The second polling controversy centres on former Austrian Chancellor Sebastian Kurz, whose high approval ratings were ultimately attributed to an intense and coordinated effort to

manipulate polls and their dissemination in Austria. This case exhibits robust material recordings and a high degree of stability, which can be attributed to the extensive and institutionalized networks that were established. The third case pertains to how powerful individuals aim to delegitimize polls and their representation of public opinion. I will investigate this phenomenon by examining the contentious relationship between Donald J. Trump and the polls. These examples raise questions about the solidity of these constructs and the power that is wielded in their (de)construction.

8.3.2.1 Faking What We Value

The following cases are mainly associated with the idea of forgery when it comes to the forging of stable statistical chains. Practices of forgery in the realm of polls, can, following Civitella (2019), be understood as a subset of fake news: "Information which is manufactured to look legitimate, however, is of partisan or erroneous construction. Possibly constructed for purposes other than the elucidation of the true 'public opinion'" (Civitella, 2019). While the topic of fake and invented polls and surveys has existed for a long time, the proliferation of social media platforms and the blurred lines between trustworthy and untrustworthy polls has facilitated the spread of fake polls. In 2019, Raben and Francovic asked in ESOMAR's Research World the question whether we experience "the beginning of an epidemic of fake polls?" (Raben and Francovic, 2019). Unlike established media organizations that must adhere to reporting guidelines, individuals on social media may not have such restrictions, making it easier for fake polls to circulate.

This raises questions as to what the proliferation of fake polls tells us about polling and public opinion more generally. One interesting aspect is an understanding of fakes as a way to understand what societies value. In referring to Jones' (1990) preface to the catalogue of the British Museum's first exhibition on the topic of fakes, Coopmans (2021) highlights that faking and coveting stand in close relation to each other. "[T]he presence of fakes is indicative of

coveting. Fakes can teach us what we value" (Coopmans, 2021, p. 80). The existence of fake and manipulated polls thus elucidates the value contemporary societies put into polls and ratings as a sign of approval. When individuals or entities create fake polls and manipulate data to portray certain narratives, it reflects a desire to influence public perception and shape opinions in their favour. This coveting of public approval through fabricated polls underscores the extent to which favourable public opinion has become an influential factor in modern society.

8.3.2.2 Fake Polls: Kid Rock For Senate

The first example leads us to the US, where in July 2017, the rap rock musician and Trump supporter Robert Ritchie, probably better known as Kid Rock, announced on Twitter that he intended to run for the US Senate seat in Michigan. On a newly launched website www.kidrockforsenate.com, which is no longer active, he claimed that "[t]he democrats are 'shattin' their pantaloons' right now… and rightfully so!… it's game on mthrfkers" (see: Wright, 2020). And the confidence was not without reason if one considers his results in one particular poll. A poll conducted by a company called Delphi Analytica suggested that Kid Rock would gain 30% of the votes, whereas his democratic competitor and current senator, Debbie Stabenow, would gain merely 26%. If only considering those who expressed a preference, 54% of the respondents indicated that they would vote for Kid Rock, while the remaining 46% preferred Debbie Stabenow. If one takes those polling results at face value, it seems as if Kid Rock was more popular for the potential US Senate seat than current Senate Stabenow.

Delphi Analytica, however, is neither an established nor a known polling company, and the veracity of this poll was in doubt. According to their website, which came online on July 6, 2017 and was closed only two months and two days later, but is archived under web.archive.org, the company "was founded in 2017 by a group of individuals from diverse political backgrounds, united by their affinity for politics, who wanted to create a grassroots public polling organization" (Delphi Analytica, 2017). The company lacked any transparency and responded

with unhelpful responses to people aiming to inquire about the company. One of those was Harry Enten, senior political writer and analyst for FiveThirtyEight. The company did not reveal any names of the individuals involved, nor provided any further information about how they conducted the poll, except that they did so using Survey Monkey and a Google survey tool, which everybody with a Gmail account can set up. Enten (2017) published screenshots from a PredictIt Discord Chatroom, which he received from a source. PredictIt is an online prediction market that offers exchanges on political and financial events. In this chat, one user, @AutismoJones, who claimed to have started Delphi Analytica referred to Enten's inquiries, stating that they do not need Harry Enten, as they have "governors tweeting our polls" and that they "are already famous" (@AutismoJones in: Enten, 2017), revealing an interesting dimension in times in which questions around the legitimacy of polling methods and companies are widely discussed. Methodological scrutiny is seen as less important compared to the distribution and dissemination of those very polls. An e-mail exchange between the anonymous founders of Delphi Analytica and WXYZ-TV, an ABC-affiliated television station in Detroit, Michigan, speaks in interesting ways to the position they claim to take up in the polling world. Referring to the traditional pollsters, they wrote that "those dinasaurs (sic) will be stuck in the past and firms like delphi will continue doing great work".

Delphi Analytica published the raw data of their survey on their website. The data looks, as Enten states, similar to the data you get through Google Surveys. Google Surveys, for instance, infers the age, gender and region of respondents based on their online activities and their IP address. In the raw data to this poll on voting preferences in Michigan, age and gender are often labelled unknown, something that also often happens to Google Surveys, when the software is unable to infer this information. Even though the data seems like they actually conducted a survey, the data also displays several anomalies. For example, capitalisation is inconsistent in the responses labelled "I prefer not to answer", which indicates that the data might stem from

different sources. Potentially, they mixed responses from Google Surveys and Survey Monkey together, without adjusting the data. It is, however, also possible that they manipulated the data or did not conduct a poll at all and merely invented it. Enten (2017) states that it "remains unclear whether the person or persons behind Delphi Analytica conducted a poll" at all.

Even though it cannot be said for sure whether the poll has any grounding in tangible political subjectivities, the publication and dissemination of this poll had an influence. Kid Rock, for instance, retweeted the Delphi Analytica survey and his candidacy was endorsed by former New York governor George Pataki, who tweeted that "Kid Rock is exactly the kind of candidate the GOP needs right now. #KidRockForSenate @KidRock". Enten concludes that "[a] poll that may not even have been conducted could wind up being at least partially responsible for the election of a musician to the U.S. Senate" is "pretty amazing" (Enten, 2017). What Enten describes as 'pretty amazing', is the fact that Delphi Analytica was, even if only for a short time, able to create a public and a representation of Michigan's voters with overarching support for Kid Rock. In a similar way as the case of fictitious issues, that can establish a pseudo-opinion, in the case of Kid Rock, a pseudo-public was being enacted. The case of Delphi Analytica speaks, not only in the choice of the name, also to other debates, such as the scandal around Cambridge Analytica and the misuse of Facebook data to influence elections (e.g., Mahfoud *et al.*, 2018).

The conventions that forged this public were, however, not very firmly established, which is why this representation became highly fragile. In order to make statistical representations immutable and to create a regime of visibility, the networks in and through which publics are forged require a stronger degree of stability. This is different in the next example, where a representation of the Austrian people was established with great stability – at least until its fraudulent nature became apparent.

8.3.2.3   Faking Polls: Operation Ballhausplatz

While in the case of Kid Rock, we explored an example in which public opinion in support of the musician as senator has potentially been invented, this subsection explores a case in which the creation of a public in support of a political leadership has been orchestrated. Both cases strongly differ in their stability. While in the case of Kid Rock, the representations and thus the enacted population has been largely instable, the case of Sebastian Kurz is probably one of the most meticulous and organized examples in recent history, in which polls have been manipulated in order to create public support for a political leadership. In this case, a representation has been forged through a process that involves all aspects of the term forging.

As already introduced earlier in chapter 5 and depicted in Figure 1, the Austrian People's Party (ÖVP) has, since a long time, been behind the Freedom Party of Austria (FPÖ) and the Social Democratic Party of Austria (SPÖ). In May 2017, however, the ÖVP, under chairmanship of Sebastian Kurz, who, in July 2017 would be elected leader of the party and in December of the same year chancellor of Austria, overtook those two parties in the polls. Conservative Sebastian Kurz, the youngest-ever chancellor of Austria has been credited to give the party new credibility and to convince the Austrian people to cast their vote in favor of the party. As has later become clear, the changing public support towards Kurz can, if at all, only be partially attributed to him as a person and the hopes and trust he assembled. The changing circumstances that the polls represented were not prior, but posterior to the polls. The success story of Sebastian Kurz points to embezzled public money, chats, newspapers and more. In short, a successful network that forged public support in his favor. The issue is concerned with the making of a particular representation of public opinion, which was depicted in the daily newspaper *Österreich*, founded in 2006 and published by the media mogul Wolfgang Fellner.

At the heart of this story, which was, in reference to the address of the chancellery in Vienna, termed "Operation Ballhausplatz", lays a pact between politicians, pollsters and a newspaper.

This compact which became known as the "Beinschab-Österreich Tool", was significant for how Sebastian Kurz[30] and his allies meticulously planned the takeover of power in Austria. Purportedly through intermediaries within the Ministry of Finance, budgetary funds from the Ministry were illicitly diverted to support these political objectives. This was done by generating fictitious invoices to disguise the actual destination of these funds. The funds acquired through these means are alleged to have been utilized, among other purposes, for commissioning manipulated surveys from the pollster Sabine Beinschab. Beinschab was a former employee of Sophie Karmasin, a former pollster at the family owned *Karmasin Motivforschung* and its subcompany *Das Österreichische Gallup Institut Dr. Karmasin GmbH*[31], who was at this time crossbench family minister of Austria. Through Karmasin, they could recruit Beinschab, who held close links to the newspaper Österreich and was tasked to publish fictitious polls in favor of Sebastian Kurz. These falsified surveys were then strategically placed in the media, notably in collaboration with the Österreich media group, which received inducements in the form of advertising bookings. This orchestration aimed to exert influence over both internal ÖVP party dynamics and public sentiment.

Due to the public availability of leaked chat protocols from within the inner circle of the forgers, we get access to the 'backstage' of what happened there. This classical sociological concept goes back to Goffman's (1956) distinction between 'front stage' and 'back stage': The difference between both is about how we perform in public, following cultural norms and scripts

---

[30]   Up to now, Sebastian Kurz has not yet been convicted, and he maintains that he had no knowledge of the manipulations. But more and more evidence points towards the fact that Kurz not only knew about it, but that he also gave the order to do so. For the time being, the presumption of innocence must be maintained.

[31]   Das Österreichische Gallup Institut was established in 1949 as Austria's inaugural market and opinion research institute. In 1964, Dr Fritz Karmasin was personally granted the exclusive right by Dr George Gallup to run the Austrian Opinion and Market Research Institute under the appellation Österreichisches Gallup-Institut.

and what we do when nobody watches. The chat protocols inform how the Austrian People's Party (ÖVP) was able to shift public support in its favor through, what Wodak (2022) termed *message control.* By this, Wodak describes a "a new media logic based on favoritism, nepotism, and clientelism" that emerged "from the specific propaganda tool developed by the former Austrian Chancellor Sebastian Kurz and his followers and implies launching and thus controlling select information […], and to financially subsidize only those media that reported favorably about the activities of Kurz's government" (Wodak, 2022, p. 3). Wodak puts this notion in the context of Walton's concept of *orchestration*, the manipulation of "different media over time to produce a cumulative message" (Walton, 1997, p. 400), through, among others, the misuse of statistics and opinion polls.

The scheme to manipulate (published) public opinion became better and better and on January 8, 2017, the newspaper Österreich, alongside a poll, also published an interview with pollster Sabine Beinschab, with the headline that the ÖVP would "benefit from switching to Kurz" as party leader. Chat protocols give us insights into the backstage of what happened there, providing us a peek into the political economy of Kurz' success.

> Johannes Frischmann (a spokesperson of Kurz): Yesterday I told Beinschab what she should say in the interview.
>
> Thomas Schmid: I've never gone as far as we have. Brilliant investment and Fellner is a capitalist.
>
> Thomas Schmid: Who pays gives the orders. I love that.[32]

---

[32] Extract from the WhatsApp chat between spokesperson Johannes Frischmann and Thomas Schmid, former General Secretary of the Ministry of Finance from January 7, 2017. See https://kurier.at/politik/inland/oevp-ermittlungen-call-me-mr-umfrage/401761890 for a summary of those chats.

The "Beinschab-Österreich Tool" thus consisted of Kurz's allies determining the specific questions to be posed in the surveys and subsequently cherry-picking favorable outcomes. In some instances, these results were even manipulated to provide additional backing for Kurz' aspirations for leadership. Furthermore, Österreich was instructed on when and how the survey outcomes were to be presented. In exchange, Österreich received regular placements of classified advertisements. As Katrin Bennhold wrote in the New York Times, prosecutors stated "that many polls before that election were falsified and that Mr. Kurz and a small cabal of allies with cultlike devotion to him paid off one of Austria's biggest tabloids to ensure favorable news coverage" (Bennhold, 2021). Kurz, so the accusation, "institutionalized the system, using taxpayers' money to elevate the appearance of his own popularity and punish journalists and media outlets that criticized him" (Bennhold, 2021).

Those practices remained unknown to the public for quite a while, which shows that it was indeed successful to create support through polls, at least until the manipulation became public knowledge. On the 6th October 2021, Austrian anti-corruption authorities executed a search operation at the offices of Federal Chancellor Sebastian Kurz, the Austrian People's Party headquarters, and the Federal Ministry of Finance. The anti-corruption prosecutors alleged Kurz, along with nine prominent politicians and newspaper executives, of involvement in embezzlement and bribery. According to the prosecutors, public funds were unlawfully utilized to finance politically motivated and occasionally manipulated opinion polls. In the aftermath of the raid, Kurz faced intense criticism from his coalition partner, The Greens, as well as the opposition. The leader of The Greens, Werner Kogler, asserted that Kurz was no longer fit to hold office and demanded that the People's Party appoint a new chancellor with a clean record. Subsequently, on the 9th October 2021, Kurz announced his resignation, paving the way for Alexander Schallenberg to assume the role of chancellor. Comically, but inevitably, Kurz' ratings then turned very unfavorable towards him, with 65% in favor of him completely leaving politics

(Mohr, 2021). It can be said that manipulated polls were used to sway and forge public opinion and might have led Sebastian Kurz first to the lead of the ÖVP and then to the chancellery. Different to the case of Kid Rock, this case displays high, even though only temporal, stability.

8.3.2.4 Fake Polls(!) "If it's bad, I just say it's fake"

Quite distinct to the case of Sebastian Kurz is the relationship between former US president Donald Trump to the polls. This relationship is, to put it mildly, not quite straightforward. Whereas Trump loves to talk, especially, to tweet about polls, whenever they are in his favor, his relationship to polls made several shifts. In December 2015, he tweeted: "Wow, my poll numbers have just been announced and have gone through the roof!" After the polls seemed to have been in favour of Hilary Clinton, he stated in October 2016, that "[e]ven though we're doing pretty good in the polls, I don't believe the polls anymore". The most telling instalment of his love-hate relationship with the polls happened during the Conservative Political Action Conference (CPAC) in July 2021 in Dallas, Texas, where he said in front of the audience that he would evaluate the validity of polls based on whether or not they are favorable to him: "By the way, you [he refers to the 2024 GOP nominee straw poll conducted by the CPAC] have a poll coming out. Unfortunately, I want to know what it is. You know they do that straw poll, right? If it's bad, I just say it's fake. If it's good, I say, 'That's the most accurate poll, perhaps, ever'."[33]

Trump does not make a secret about his view of polls, nor does he actively aim to manipulate them, as we have seen in the case of Sebastian Kurz. What can be observed here is a politicization of polling to a degree that any statement runs the risk of being relativized as interest-bound such that statistical representations no longer grants epistemic legitimacy. To Wodak (2022), Trump

---

[33]   After the speech, Fox News delivered the results of the CPAC straw poll, which was actually overwhelmingly favorable for Trump and thus highly accurate and trustworthy if one follows Trump's own quality criteria. According to the poll, 98% of respondents expressed their approval of Trump's performance and when asked about their preference in the 2024 Republican presidential primary, a substantial 70% chose Trump as their top candidate, while 26% favored Gov. Ron DeSantis.

can be viewed in contrast to Kurz, in that Trump "delegitimized *all* investigative journalism *without* explicitly attempting to control it" (Wodak, 2022, p. 3 italics in original). Whereas the practices around the manipulation of polls have been conceptualized as a form of 'message control', Wodak views Trump rather as an 'agitator' (Lowenthal and Guterman, 2021), "as he supported and instrumentalized far-right and extreme-right media channels (such as *Breitbart* or *Fox News*) and extensively used Twitter to spread systematic disinformation" (Wodak, 2022, p. 3 italics in original). In this case, we can again observe a new phenomenon, since it is not about the power to shape polls in a certain way; rather, it is about the power to create boundaries between which polls are legitimate and which ones are not. This practice might, as I showed in 6.3.3, even have direct consequences on polling accuracy, as there might have been a tendency of Trump supporters not to respond to traditional polls and surveys in the runup to the 2020 US presidential elections. Polls might not be faked, but they can become so in the minds of people.

8.4 Forging Power

8.4.1 On the Relation between Power and Representation

Throughout this and the previous chapter, we came across various forms in which power is exercised over the determination of representations by means of polls and surveys. In particular, throughout this and the previous chapter we came across a variety of different types, including the power to set and reject industry standards, the power to shape and determine polling outcomes or the power to disseminate and (il-)legitimize certain outcomes over others. Generally, the complex relationship between power and knowledge has been conceptualised in different ways. Francis Bacon, for instance, emphasised the relation between knowledge and power as one of gaining control over nature (Bacon, 2000) or Shapin and Schaffer (2011) emphasise the power of testimonies in the experimental creation of facts. Crucial for what we have seen throughout this dissertation is an understanding of power following Kuhn (1996) as the ability to change the rules underlying the production of scientific knowledge. This final

subsection aims to allude to such a conceptualization of power, one that is intertwined with the authority to actualise potential representations or to make representations (in)visible.

8.4.2 Legitimate Conventions

When thinking about power, a central reference point is the work of Max Weber. In *Economy and Society*, Weber defines power "as every chance, within a social relationship, of enforcing one's own will even against resistance, whatever the basis for this chance might be" (Weber, 2019, p. 134). The central point here is that it is about a social relationship in which power is applied from one individual to another. The power of pollsters to create representations, is, however, based on legitimacy and validity, not so much on power over other people. Referring to his work on authority, authority is the legitimate use of power, legitimated through a certain belief towards it. Expanding upon this notion, it becomes evident that authority's effectiveness relies on the acceptance and recognition of its legitimacy within a societal framework, one might say a convention. The term convention is distinct to its understanding as customs and traditions, as, for instance, found in Weber's conceptualization of the term. Rather, conventions of measurement are construed as logics of coordination, as I have shown in drawing on second-generation thinkers within the economies of convention. Thinking with Weber about conventions of measurement is in any case fruitful, as it offers an insightful vantage point through which conventions metamorphose into sources of legitimacy.

As we have seen throughout this dissertation, legitimacy has been created and negotiated in different ways, culminating in stable conventions. The discussions at the ISI until the 1920s have, for instance, created a legitimation of sampling through, to some extent at least, legal-rational procedures, leading to a joint declaration from members of the ISI about the appropriateness of using sampling. Throughout this period, the rules underlying the production of knowledge, when the source of that knowledge is a survey, have significantly changed from a dominance of full enumeration to one of sampling.

8.4.3   The Power of Associations

Despite there being conventions about how to measure and represent public opinion, not every representation becomes durable and stable. As we have seen in the example of Sebastian Kurz, Kid Rock and Donald Trump, the stability of polls depends on the enrolment of various actors to become stable. Even though we might say that Kurz might have possessed legitimacy through a form of charismatic leadership, his power to shape public opinion lies in the ability to create a stable assemblage. To Latour "power is not something you may possess and hoard" (Latour, 1984, p. 263) but a "*composition* made by many people", which amount depends on the "number of people who enter into the composition" (Latour, 1984, p. 265) but which is usually attributed to only one person. In this way, power always depends on the stable assembling of heterogeneous actors.

In *Pandora's Hope,* Latour (1999) explores a case involving the sampling of soil in the Amazon Forest. This case serves as an illustration that challenges the notion of scientific knowledge production as a simple transition from the realm of physical objects to written statements about them. Knowledge production is not to be understood as a mere move from things to words, but as a complex and intricate chain of translation. Since the soil cannot directly be summoned as a source, it needs to be translated onto a scientific publication. By adopting such a viewpoint, a new dimension of understanding emerges regarding the dynamics of power within this process. It compels us to reevaluate how the various chains of association are meticulously crafted and solidified, contributing to the establishment of immutable facts and assertions. This perspective allows us to grasp the implications for comprehending the influence wielded by figures like Kurz or Trump, as well as the potent sway held by polling companies in shaping public opinion. The power that these entities hold lies in their ability to skillfully orchestrate an assembly of diverse elements, encompassing the media, internet users questionnaires and more. It is the very interplay of legitimized authority, and the ability to

assemble various elements that make up the sample-population nexus, and thus the apparatus in and through which publics are enacted.

Viewed through those lenses, the power to forge stable statistical chains involves legitimized authority and the ability to form stable assemblages, ultimately establishing a regime of (in)visibility of representations. In assembling, what I have earlier introduced as *predictive stability*, polling companies come to dominate the capacity to delineate what counts as legitimate knowledge about public opinion. If public opinion constitutes an important source in the legitimacy of policies and the course politicians should follow, polling, as a "science immediately accomplished as opinion" (Rancière, 1999, p. 105) has the potential to bring about immutable representations, a process I have come to term as forging. Furthermore, forging comes in three shapes, understood as the making of stability, the building of strong ties and the malicious practice of forgery. Just as forging involves the meticulous shaping and transformation of materials to produce something distinct, forging in the realm of representing social reality encompasses the act of rendering things perceptible and bringing them into the realm of shared understanding. Similar to the blacksmith's craft, statistical practice takes raw, unrefined data, and refines it into a structured form with enhanced properties and can evolve into stable representations. This alignment can occur through legitimized authority of those who create those representations, but also through acts of forgery, which entail deceitful manipulation of the sphere of visibility.

8.5 Conclusion

This final chapter has introduced the notions of forging stable statistical chains, designating the different ways in which publics are constructed in and through polls and surveys. This notion bears resemblance to Margolis' and Mauser's (1989) notion of "Public Opinion as a Dependent Variable", underlying the idea of a sample-population nexus, whereas one, out of multiple populations is enacted from a sample. Stable statistical chains are thus chains of translation and

association, that define which populations and opinions enter the visible. Furthermore, the forging of those chains acknowledges the broader landscape, including attempts to manipulate and shape not public opinion per se but rather its very measurement. Especially in times of uncertainty, when the monopoly of traditional pollsters disintegrates and in which not only spokespersons of public opinion but also potential public opinion itself has multiplied, the question of which ones become immutable gains particular relevancy. What this chapter ultimately distinguishes from the previous analysis of controversies is the argument that there is always more to the representation of public opinion than questions around methodological decisions - it is about the stability of networks of actors and techniques, in and through which their representations are forged.

CHAPTER 9

9   Conclusion

> It is important to remember that while the seismograph does not create earthquakes, this instrument may one day help to alleviate such catastrophes by charting the place of their occurrence, their strength, and so enabling those interested in controlling the effects of such disasters to obtain more knowledge of their causes. Similarly, the polls do not create the sources of irrationalism and potential chaos in our society. What they can do is to give the people and the legislators a picture of existing tendencies, knowledge of which may save democracy from rushing over the edge of the precipice. (Gallup and Rae, 1968, p. 270)

> Whereas in fact, I think that is my concern and that is also my resentment [...] that the use of the instruments of empirical research is worse today than it used to be many years ago when people started to use and see it as an instrument of policy advice in a positive sense. Today it has degenerated, also on the part of the political actors, as an instrument for superficial reflections of moods. (Güllner, Interview, Translation)

In this dissertation, I referred to Gallup's notion that polling protects the people "from the tyranny of the majority". According to Gallup, polls provide politicians with a "sampling referendum" on the "background of tested knowledge" (Gallup and Rae, 1968, p. 268). Rather than leading to a "mobocracy", in which the majority can overturn minority opinions, he argues that polls make the "maintenance of a free tribunal of public opinion in which rival protagonists can make their appeals" (Gallup and Rae, 1968, p. 270) possible, creating an atmosphere in which democratic methods produce intelligent results. In times in which the generally accepted criteria of how to enable such a free tribunal of public opinion have become fragile, the question of how to return to this ideal becomes pressing. The historical and contemporary trajectories of polling have revealed that persistent challenges related to methodology and transparency, coupled with the deliberate efforts to manipulate and undercut polling data, have to some extent eroded the

democratic significance of polling. The contemporary landscape in which survey and polling data is produced and disseminated falls short in cultivating greater public trust, an element that lies at the heart of realizing the democratic ideals envisaged by the early pioneers of polling. Rather than providing trusted and generally accepted representations of the social world and public opinion, people have at their disposal a variety of oftentimes conflicting representations to choose from. As we navigate the future of polling, addressing these issues remains crucial in restoring the role of polling data in shaping our democratic discourse and fulfilling the promises that early pioneers of polling once made when they marketed their craft.

Brief summary

The previous chapters have navigated through a complex terrain, beginning with a seemingly isolated event and culminating in an exploration of the intricate interplay between methodological debates, representation, and the construction of publics and public opinion. The starting point was a discussion around an eventful photoshoot between German national football players Mesut Özil and İlkay Gündoğan with the Turkish president Recep Tayyip Erdoğan. The photoshoot provoked a controversy, leading to several polls regarding their role in the German national football team. Those very polls, which, as I have shown, strongly diverged in their findings, themselves led to a controversy about polling methodologies, at the heart of which was the question, as the plaintiffs in a complaint against the publication of a Civey survey by Focus Online put it, of what constitutes the "generally accepted scientific criteria of empirical social research".

Throughout this dissertation, I unpacked the historical, sociological and conceptual dimensions surrounding the emergence and negotiation of those very criteria. Chapters 3 and 4 explored the history of survey sampling and polling from the end of the 19th century until the mid-20th century, bringing together the history of a particular social science technique, the random sample, and the emergence of public opinion as the statistical aggregate of private

opinions. As a result of those entanglements between public opinion and the sample survey, pollsters and polling institutes monopolised access to and representation of public opinion. They developed into an obligatory passage point for anyone who wants to arrive at knowledge about public opinion. What this implies is that the methodological choices of these custodians of public sentiment wield influence over the emergence of competing realities. Chapter 5 explored this relationship between polling and representation, arguing for an ontological understanding of public opinion that integrates realist views, which are usually held by practitioners in the field, and constructivist views, raised by contemporary theorists of representation. If public opinion, as this chapter argued, is to some extent brought into being by the techniques that purport to describe it, methodological discussions about how to best "know" a population become central for the making of truth about this very population. The question of whether the German population was in favour or opposed to Özil and Gündoğan playing for the German national football team hinges on the ability of either of those representations to, what Aradau and Huysmans (2019) called, assemble credibility. This became the starting point for the empirical analysis, starting at the end of chapter 5, but mainly elaborated on in chapters 6 and 7, which set out to explore the recent history of the controversy around probability and non-probability sampling. In this context, I drew on biographical narratives to explore conversion stories told by practitioners who changed their methodological standpoints. Based on those narratives and historical background, I explored the different ways in which practitioners discuss those methodological questions as a form of boundary work. From a historical perspective, this controversy is reminiscent of the early history of survey research when probability and non-probability sampling co-existed relatively peacefully. From a contemporary perspective, those practices of boundary work emphasize a struggle over the interpretive sovereignty over the public and their opinions, highlighting the relation between the making of publics and the methodological and statistical chains that lead to that. The analysis of this issue was the topic of

chapter 8, in which I introduced the notion of forging stable statistical chains. Whilst approaches from second-generation proponents of conventionalist understandings of scientific practice are able to explain the processes that lead to the statistical representations of publics and their opinions, they fall short in explaining what is needed to establish those very publics. Therefore, I looked into polling controversies that highlight different forms of power when it comes to the making of stable publics through polls. This concept allows for the analysis of polling to be enlarged to include the broader landscape in which polling takes place.

At the core of this dissertation was the finding that contemporary controversies over sampling led to a strong divide and polarization and not, as one might expect, to mutual attempts to solve the problems the field is concerned with. Considering the early 20th century, we have seen how a very similar controversy already took place, which, was, however, discussed and solved rather peacefully, raising the question of whether and how the current debate will resolve. One central difference was that now, those discussions are strongly intertwined with economic interests. While economic and commercial aspects of sampling were not part of the discussions at the ISI, commercial aspects have now become part and parcel of contemporary debates around sampling, potentially necessitating different solutions than in the past. Will full transparency of methods reduce the antagonism between proponents of probability and non-probability methods? Will we see new and generally accepted criteria for survey research and polling emerge as an outcome of this controversy? Or will the future be characterized by a multiplicity of competing methods? In the remainder of this concluding chapter, I will formulate a cautious approximation to those kinds of questions.

The Call for Transparency

In the last years, several attempts have been made to increase transparency and public disclosure of methods. I already mentioned, for instance, the Roper Center's archives that added a "Recently Developed Methods Collection", to store polling and survey data stemming from new and

experimental methods under the condition of full methodological disclosure. Another example is AAPOR's Transparency Initiative (TI), formally launched in 2014 to promote methodological disclosure by assisting survey organizations in routinely sharing their research methods for publicly released studies through an educational approach. While the TI refrains from assessing method quality, it is guided by the belief that complete disclosure enhances consumers' capacity to evaluate methods. In addition, public disclosure opens up a space for scrutiny and assessment of used methods.

These initiatives are undeniably significant and play an important role in bolstering trust. However, it appears that they may not be a sufficient corrective when considering the scarce resources and political economy of polling and survey data, the media and the public, as well as the prevalence of motivated reasoning in evaluations of published polls. Recent research (Kuru, Pasek and Traugott, 2020) on the trust towards and perceptions of surveys shows that people disproportionally find those polls more reliable and credible when they support their own political convictions or candidates. In the contemporary survey climate, people have a large amount of potential survey and polling findings to choose from, making it possible to cherry-pick and believe the polls they deem favourable. Those findings speak in important ways to partisan attacks against "rigged polls", as we have seen in the context of Donald J. Trump's statements in the context of the 2016 and 2020 US presidential elections, as explored in chapter 8. In times in which such a huge amount of, oftentimes conflicting, polling data is published, addressing this issue goes beyond mere transparency in polling methodology. Some suggest the deliberate and transparent reporting of competing polls as a potential remedy.

Deliberately reporting about competing polls, or highlighting that there are polls that say the opposite might lead to a greater debate about polls and their role in democracies, they might, however, also raise the issue known as false balancing, the practice that journalists present opposing viewpoints and evidence to balance out the variety of potential scientific findings. In

referring to other scientific fields, where some demand the media to consider the balanced publication of conflicting studies towards issues such as climate change or COVID-19, one might demand a similar treatment for polling data. While this might give the impression of considering and representing different methodologies and thus giving the public greater agency to form a decision about which one to believe, the goal should not be to produce and publish stories with the appearance of balance, but to seek the truth, or at least, the best available evidence.

Statistical Literacy

A theme that was often raised in my interviews was a perceived lack of statistical and data literacy on the side of the public, politics and the media. A recent initiative in Germany is the Data-Literacy-Charta, initiated in 2021. The Charta "formulates a common understanding of data literacy and its importance for educational processes", which is seen as "the key to systematically transforming data into knowledge and action" (Schüller, Koch and Rampelt, 2021, pp. 1–2). Statistical and Data Literacy deems important for everyone who deals with numbers and data in one way or another, especially in times in which statistics and numbers seem to have become fragile and contested. This is especially central for journalism, as Gary Langer made clear in my interview with him, saying that "journalists for far too long, engaged themselves in the lazy luxury of being both data hungry and math phobic. […] [R]eporters see a number and a percentage sign and it's just too compelling. They want it, they got to have it, they grab it, and they run with it. And running with data is like running with scissors, you can easily get hurt." (Langer, Interview)

Part of this consciousness must be an increased understanding of the practices that brought numbers about. Langer continued, saying that "it's not just numbers and percentage signs, we have to stop and understand where and how these numbers and percentage signs came from […]. But again, it's a fight that we spend many years fighting" (Langer, Interview). In other words, before considering certain representations and numbers as reliable and meaningful, it is essential

to comprehend the source of the numbers, the methods and the type of sampling employed, and the nature of the questions asked. Only after gaining a clear understanding of these factors can one genuinely evaluate the analysis and treat it with seriousness in reporting. Rather than reducing survey and polling data to "matters of fact", they must be seen as what they are, "highly complex, historically situated, richly diverse matters of concern" (Latour, 2004, p. 237). Educatory programmes on the side of journalists and media representatives can help to learn the craft of reassembling the steps, methods and histories that led to the polling findings, an important skill for judging their stability and robustness.

(When) will the Controversy Come to Closure?

Looking again at the concrete controversy between probability and non-probability sampling, and the development of a new accepted practice of how to conduct surveys and polls, the burden of proof is often put on either of both sides. Talking in ideal-typical ways, both sides, however, have strong arguments in their favour or against them. So why would the burden of proof be only on the side of proponents of non-probability methods, as proponents of probability sampling often claim and vice versa. Is it not somehow understandable to switch to new and experimental methods, considering strongly increasing costs and efforts, despite there not being a coherent theory? Didn't it also take over thirty years to provide a theoretical justification for sampling, after Kiær first proposed it in 1895? On the other hand, it is also understandable to argue in favour of traditional methods, since they have proven to provide a trustworthy picture of society, as long as the relevant conditions are met. But what seems equally important in the light of the historical development of survey and polling methods, is to put the burden of proof (also) on those who commission, communicate and consume polling data. Polling controversies did not necessarily come to closure through 'definitive' proof or 'knockdown' arguments. Closure emerged as the result of what I termed predictive stability, mutually established by pollsters, the media, consumers and the public.

When visiting conferences in the field of survey research, I observed strong and honest efforts in the community to acknowledge the problems the field is facing at the moment. Researchers aim, for instance, to integrate both approaches or to approximate probability samples in cases where no sampling frame is available. While those conversations take place in the scientific community, commercial polling and survey companies seem to continue their fights on different fronts. The dispute between Civey and Forsa illuminates this issue. The battle between both has gone on since 2017 and is still nowhere close to being settled. While Civey is, for instance, taking legal action to prevent Güllner from calling them a dangerous bunch of crooks (Gefährlicher Gaunerhaufen), Forsa was able to obtain legal action to stop Civey from calling itself the market leader. According to Scheppe (2023), Civey has spent six-figure sums on legal disputes with Forsa, an amount that could arguably be spent for better ends.

Towards Epistemic Humility

Similar discussions in polling and survey research can also be observed in different areas of statistical representation beyond the sphere of polling. The time period during which I worked on this dissertation was also the time of the COVID-19 pandemic, a time when numbers and representations of reality have been highly controversial and contested (Billig and Marinho, 2022). One question I asked almost all of my interviewees was how they perceived the role of statistics during the pandemic. The responses I obtained unanimously highlighted the lack of evidential, trustworthy and representative data. It became clear that we had to cope with a "virus that is shrouded in uncertainties and ignorance" (Gross, 2021) and dealing with ignorance became a reality for everybody. During the pandemic, viewed in the context of public discussions and the loss of trust towards scientists and governmental bodies among many parts of society, the question as to which data and evidence to trust became crucial. It forced us, however, to also acknowledge the role of ignorance. David Spiegelhalter commented on the role of ignorance during the pandemic, saying that "it's quite good to talk about things we don't know, to have […]

free epistemic humility, […] just admit you don't know things. It doesn't lose the trust of audiences. […] We should be admitting what we don't know" (Spiegelhalter, Interview).

Moving beyond the complicated and polarised situation in polling and other fields is crucial to protect the integrity of public discourse and ensure that polling renews its role as a safeguard against the tyranny of the majority and to enable a free tribunal of public opinion. Rather than insisting on the superiority of one approach over another, the better way to assemble credibility and make statistical representations fit for the future might lie in acknowledging the problems and challenges and fighting battles on the right fronts.

10 Bibliography

AAPOR (2006) 'AAPOR Awards', *The Public Opinion Quarterly*, 70(3), pp. 423–425.

Adorno, T.W. (2005) 'Opinion Research and Publicness: (Meinungsforschung und Öffentlichkeit)', *Sociological Theory*. Translated by A.J. Perrin and L. Jarkko, 23(1), pp. 116–123. Available at: https://doi.org/10.1111/j.0735-2751.2005.00245.x.

Almond, G.A. and Verba, S. (1972) *The civic culture: political attitudes and democracy in five nations*. 4. print. Princeton, NJ: Princeton Univ. Pr.

Alterman, J.B. (2001) *Statisticians of the Centuries*, *Egypt and American Foreign Assistance 1952–1956*. Edited by C.C. Heyde et al. New York, NY: Springer New York. Available at: https://doi.org/10.1007/978-1-4613-0179-0.

Althaus, S.L. (2003) *Collective preferences in democratic politics: opinion surveys and the will of the people*. Cambridge, UK : New York: Cambridge University Press.

Anderson, B.R.O. (2016) *Imagined communities: reflections on the origin and spread of nationalism*. Revised edition. London New York: Verso.

Aradau, C. and Huysmans, J. (2019) 'Assembling credibility: Knowledge, method and critique in times of "post-truth"', *Security Dialogue*, 50(1), pp. 40–58. Available at: https://doi.org/10.1177/0967010618788996.

Baber, Z. (2000) 'Ambiguous Legacy: The Social Construction of the Kuhnian Revolution and Its Consequences for the Sociology of Science', *Bulletin of Science, Technology & Society*, 20(2), pp. 139–155. Available at: https://doi.org/10.1177/027046760002000208.

Bacon, F. (2000) *The new organon*. Edited by L. Jardine and M. Silverthorne. Cambridge [U.K.] ; New York: Cambridge University Press (Cambridge texts in the history of philosophy).

Barad, K. (1996) 'Meeting the Universe Halfway: Realism and Social Constructivism without Contradiction', in L.H. Nelson and J. Nelson (eds) *Feminism, Science, and the Philosophy of Science*. Dordrecht: Springer Netherlands, pp. 161–194. Available at: https://doi.org/10.1007/978-94-009-1742-2_9.

Barad, K. (1999) 'Agential realism: Feminist interventions in understanding scientific practices', in M. Biagioli (ed.) *The science studies reader*. New York London: Routledge, pp. 1–11.

Barad, K.M. (2007) *Meeting the universe halfway: quantum physics and the entanglement of matter and meaning*. Durham: Duke University Press.

Barker, E. (1967) *Reflections on Government*. London: Oxford University Press.

Becker, H.S. and Horowitz, I.L. (1972) 'Radical Politics and Sociological Research: Observations on Methodology and Ideology', *American Journal of Sociology*, 78(1), pp. 48–66.

Bellhouse, D.R. (1988) 'A brief history of random sampling methods', in P.R. Krishnaiah and C.R. Rao (eds) *Sampling*. Amsterdam ; New York: North-Holland : Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co (Handbook of statistics, v. 6), pp. 1–14.

Bemmann, M. (2023) 'Introduction: exploring the International Statistical Institute, 1885–1938', *European Review of History: Revue européenne d'histoire*, 30(1), pp. 1–12. Available at: https://doi.org/10.1080/13507486.2023.2165436.

Bennhold, K. (2021) 'Fake Polls and Tabloid Coverage on Demand: The Dark Side of Sebastian Kurz', *The New York Times*, 17 October. Available at: https://www.nytimes.com/2021/10/17/world/europe/austria-sebastian-kurz-scandal-chancellor.html.

Berger, P.L. and Luckmann, T. (1991) *The social construction of reality: a treatise in the sociology of knowledge*. Repr. in Penguin Books. London: Penguin Books (Penguin social sciences).

Bethlehem, J.G. (2009) 'The rise of survey sampling. (CBS Discussion Paper; No. 09015)', *Statistics Netherlands*, pp. 1–28.

Bethlehem, J.G. (2018) *Understanding public opinion polls*. Boca Raton London New York: CRC Press, Taylor & Francis Group.

Billig, M. and Marinho, C. (2022) 'Preventing the political manipulation of Covid-19 statistics: The importance of going beyond diplomatic language', *Language in Society*, pp. 1–23. Available at: https://doi.org/10.1017/S0047404522000367.

Bird, G.L. and Merwin, F.E. (1971) *The press and society: a book of readings*. Westport, Conn: Greenwood Press.

Bishop, G.F. *et al.* (1980) 'Pseudo-Opinions on Public Affairs', *The Public Opinion Quarterly*, 44(2), pp. 198–209.

Blondiaux, L. (1998) *La fabrique de l'opinion: une histoire sociale des sondages*. Paris: Seuil (Science politique).

Bloor, D. (1991) *Knowledge and social imagery*. 2nd ed. Chicago: University of Chicago Press.

Blumer, H. (1948) 'Public Opinion and Public Opinion Polling', *American Sociological Review*, 13(5), pp. 542–549. Available at: https://doi.org/10.2307/2087146.

Blumer, H. (1969) *Symbolic interactionism; perspective and method*. Englewood Cliffs, N.J.: Prentice-Hall.

Bourdieu, P. (1972) 'Public Opinion Does Not Exist', *Communication and Mass Struggle*, pp. 124–130.

Bourdieu, P. (1987) 'The Force of Law: Toward a Sociology of the Juridical Field', *The Hastings Law Journal*, 38, pp. 814–853.

Bourdieu, P. (1998) *On television*. New York: New Press.

Bourdieu, P. (2010) *Distinction: a social critique of the judgement of taste*. London: Routledge (Routledge classics).

Bowley, A.L. (1906) 'Address to the Economic Science and Statistics Section of the British Association for the Advancement of Science, York, 1906', *Journal of the Royal Statistical Society*, 69(3), p. 540. Available at: https://doi.org/10.2307/2339344.

Bruno, I., Jany-Catrice, F. and Touchelay, B. (2016) 'Introduction. The Social Sciences of Quantification in France: An Overview', in I. Bruno, F. Jany-Catrice, and B. Touchelay (eds) *The Social Sciences of Quantification*. Cham: Springer International Publishing (Logic, Argumentation & Reasoning), pp. 1–14. Available at: https://doi.org/10.1007/978-3-319-44000-2_1.

Burrows, R. and Savage, M. (2014) 'After the crisis? Big Data and the methodological challenges of empirical sociology', *Big Data and Society*, 1(1), pp. 1–6. Available at: https://doi.org/10.1177/2053951714540280.

Callegaro, M. *et al.* (2014) 'Online panel research: History, concepts, applications and a look at the future', in M. Callegaro et al. (eds) *Online Panel Research: A Data Quality Perspective*. Chichester, UK: John Wiley & Sons, Ltd (Wiley series in survey methodology), pp. 1–22. Available at: https://doi.org/10.1002/9781118763520.

Callon, M. (1984) 'Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay', *The Sociological Review*, 32(1_suppl), pp. 196–233. Available at: https://doi.org/10.1111/j.1467-954X.1984.tb00113.x.

Chang, L. and Krosnick, J.A. (2001) 'The accuracy of self-reports: Comparisons of an RDD telephone survey with Internet surveys by Harris Interactive and Knowledge Networks', in. *American Association for Public Opinion Research Annual Meeting, Montreal, Canada*.

Chase, S. (1948) 'Are the Polls Finished?', *Nation*, pp. 626–629.

Christensen, D. and Lackey, J. (eds) (2016) *The epistemology of disagreement: new essays*. Oxford: Oxford University Press.

Cicourel, A.V. (1964) *Method and measurement in sociology*. New York: Free Press of Glencoe.

Civitella, M. (2019) 'Fake Polls and Faux Influence: The Illusions in Political Polls and Advocacy Research', in M.J. Sheehan (ed.) *Advocates and persuaders*. North Melbourne, Vic: Australian Scholarly Publishing.

Cohn, N. (2016) 'We Gave Four Good Pollsters the Same Raw Data. They Had Four Different Results.', *The New York Times*, 10 September. Available at: https://www.nytimes.com/interactive/2016/09/20/upshot/the-error-the-polling-world-rarely-talks-about.html.

Cohn, N. (2022) 'Who in the World Is Still Answering Pollsters' Phone Calls?', *The New York Times*, 12 October. Available at: https://www.nytimes.com/2022/10/12/upshot/midterms-polling-phone-calls.html.

Collins, H.M. (2015) 'Scientific Knowledge, Sociology of', in *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, pp. 308–312. Available at: https://doi.org/10.1016/B978-0-08-097086-8.85029-1.

Coopmans, C. (2021) 'Learning from Fakes: A Relational Approach', in S. Woolgar et al. (eds) *The imposter as social theory: thinking with gatecrashers, cheats and charlatans*. Bristol: Bristol University Press, pp. 77–102.

Cornesse, C. *et al.* (2020) 'A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research', *Journal of Survey Statistics and Methodology*, 8(1), pp. 4–36. Available at: https://doi.org/10.1093/jssam/smz041.

Couper, M.P. (2000) 'Web Surveys : A Review of Issues and Approaches', *Public Opinion Quarterly*, 64(4), pp. 464–494. Available at: https://doi.org/10.1086/318641.

Couper, M.P. (2013) 'Is the sky falling? New Technology, Changing Media, and the Future of Surveys', *Survey Research Methods*, 7(3), pp. 145–156. Available at: https://doi.org/10.18148/srm/2013.v7i3.5751.

Crossley, A.M. (1937) 'Straw Polls in 1936', *Public Opinion Quarterly*, 1(1), p. 24. Available at: https://doi.org/10.1086/265035.

Daston, L. (1988) *Classical Probability in the Enlightenment*. Princeton, N.J: Princeton University Press.

Davies, W. (2017) 'How statistics lost their power and why we should fear what comes next', *The Guardian*, 19 January. Available at: https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy.

Davies, W. (2018) *Nervous states: how feeling took over the world*. London: Jonathan Cape.

Davies, W. (2020) *This is not normal: the collapse of liberal Britain*. London New York: Verso.

Delphi Analytica (2017) 'About Us'. Available at: https://web.archive.org/web/20170821190701/https://delphianalytica.org/about-us/.

Derrida, J. (1973) *Speech and phenomena, and other essays on Husserl's theory of signs*. Evanston: Northwestern University Press (Northwestern University studies in phenomenology & existential philosophy).

Desrosières, A. (1998) *The politics of large numbers: a history of statistical reasoning*. Cambridge, Mass: Harvard University Press.

Desrosières, A. (2005) 'Décrire l'État ou explorer la société : les deux sources de la statistique publique', *Genèses*, 58(1), p. 4. Available at: https://doi.org/10.3917/gen.058.0004.

Detel, W. (2001) 'Social Constructivism', in N.J. Smelser and P.B. Baltes (eds) *International encyclopedia of the social & behavioral sciences*. 1st ed. Amsterdam ; New York: Elsevier, pp. 14264–14267.

Dever, J.A. *et al.* (2021) 'Fit for Purpose in Action: Design, Implementation, and Evaluation of the National Internet Flu Survey', *Journal of Survey Statistics and Methodology*, 9(3), pp. 449–476. Available at: https://doi.org/10.1093/jssam/smz050.

Diaz-Bone, R. (2019) 'Convention Theory, Surveys and Moral Collectives', in S. Joller and M. Stanisavljevic (eds) *Moralische Kollektive*. Wiesbaden: Springer Fachmedien Wiesbaden

(Wissen, Kommunikation und Gesellschaft), pp. 115–135. Available at: https://doi.org/10.1007/978-3-658-22978-8_7.

Diaz-Bone, R. and Didier, E. (2016) 'The sociology of quantification - Perspectives on an emerging field in the social sciences', *Historical Social Research*, 41(2), pp. 7–26. Available at: https://doi.org/10.12759/hsr.41.2016.2.7-26.

Diaz-Bone, R. and Horvath, K. (2021) 'Official statistics, big data and civil society. Introducing the approach of "economics of convention" for understanding the rise of new data worlds and their implications', *Statistical Journal of the IAOS*, 37(1), pp. 219–228. Available at: https://doi.org/10.3233/SJI-200733.

Didier, E. (2002) 'Sampling and Democracy: Representativeness in the First United States Surveys', *Science in Context*, 15(3), pp. 427–445. Available at: https://doi.org/10.1017/S0269889702000558.

Dillman, D.A. (2002) 'Presidential Address: Navigating the Rapids of Change: Some Observations on Survey Methodology in the Early Twenty-First Century', *The Public Opinion Quarterly*, 66(3), pp. 473–494.

Dillman, D.A. (2017) 'The promise and challenge of pushing respondents to the web in mixed-mode surveys', *Survey Methodology*, 43(1), pp. 3–31.

Disch, L. (2011) 'Toward a Mobilization Conception of Democratic Representation', *The American Political Science Review*, 105(1), pp. 100–114.

Disch, L. (2019) 'Introduction: the end of representative politics?', in L. Disch, N. Urbinati, and M. van de Sande (eds) *The Constructivist Turn in Political Representation*. Edinburgh: Edinburgh University Press, pp. 1–18. Available at: https://doi.org/10.3366/edinburgh/9781474442602.003.0001.

Disch, L. (2021) *Making constituencies: representation as mobilization in mass democracy*. Chicago ; London: The University of Chicago Press.

Disch, L., Sande, M. van de and Urbinati, N. (eds) (2019) *The constructivist turn in political representation*. Edinburgh: Edinburgh University Press.

Dorling, D.F.L. and Simpson, S. (1999) *Statistics in society: The arithmetic of politics*. Edited by D.F.L. Dorling and S. Simpson. Arnold Publishers.

Douglas, M. (2005) *Purity and danger: an analysis of concept of pollution and taboo*. London ; New York: Routledge (Routledge classics).

Droesbeke, J.-J., Fichet, B. and Tassi, P. (eds) (1987) *Les Sondages*. Paris: Economica.

Duncan, O.D. (1984) *Notes on social measurement : historical and critical*. New York (N.Y.) : Russell Sage Foundation. Available at: http://lib.ugent.be/catalog/rug01:000116146.

Durkheim, E. and Mauss, M. (2009) *Primitive Classification (Routledge Revivals)*. London: Routledge. Available at: https://doi.org/10.4324/9780203092828.

Ellwanger, A. (2017) 'Reinventing *doxa* : public opinion polling as deliberative discourse', *Argumentation and Advocacy*, 53(3), pp. 181–198. Available at: https://doi.org/10.1080/00028533.2017.1337330.

Enns, P.K. and Rothschild, J. (2020) 'Revisiting the "gold standard" of polling : new methods outperformed traditional ones in 2020', *medium.com*. Available at: https://medium.com/3streams/revisiting-the-gold-standard-of-polling-new-methods-outperformed-traditional-ones-in-2020-451650a9ba5b.

Enten, H. (2017) 'Fake Polls Are A Real Problem', 22 August. Available at: https://fivethirtyeight.com/features/fake-polls-are-a-real-problem/.

Erikson, R.S. and Tedin, K.L. (1981) 'The 1928–1936 Partisan Realignment: The Case for the Conversion Hypothesis', *American Political Science Review*, 75(4), pp. 951–962. Available at: https://doi.org/10.2307/1962295.

ESOMAR (2017) *Global Market Research 2017*. Amsterdam.

Espeland, W.N. and Stevens, M.L. (2008) 'A sociology of quantification', *Archives Europeennes de Sociologie*, 49(3), pp. 401–436. Available at: https://doi.org/10.1017/S0003975609000150.

Evans, J., Ruane, S. and Southall, H. (eds) (2019) *Data in society: challenging statistics in an age of globalisation*. Bristol Chicago, Il: Policy Press.

Felt, U. *et al.* (eds) (2017) *The handbook of science and technology studies*. Fourth edition. Cambridge, Massachusetts: The MIT Press.

Fiorina, M. and Krosnick, J.A. (2004) *The Economist/YouGov Internet Presidential Poll*. Available at: https://www.economist.com/media/pdf/Paper.pdf.

Fishkin, J.S. (1988) *Deliberative polling: Executive summary*, *Center for Deliberative Democracy at Stanford University*. Available at: https://web.archive.org/web/20150210030843/http://cdd.stanford.edu/polls/docs/summary/.

Fishkin, J.S. (1997) *The voice of the people: public opinion and democracy*. New ed. New Haven: Yale University Press.

Fleck, L. (1981) *Genesis and development of a scientific fact*. Edited by T.J. Trenn and R.K. Merton. Translated by F. Bradley and T.J. Trenn. Chicago: Univ. of Chicago Press (Sociology of Science).

Foucault, M. (2009) *Security, territory, population: lectures at the Collège de France, 1977-1978*. 1. Picador ed. Edited by M. Senellart. Translated by G. Burchell. New York, NY: Picador (Lectures at the Collège de France).

Freedman, D., Pisani, R. and Purves, R. (2007) *Statistics*. 4th edn. New York: W.W. Norton & Co.

Gallup, G. (1936) 'Measuring public opinion', *Vital Speeches of the Day*, 2, pp. 370–372.

Gallup, G. (1948) *A guide to public opinion polls.* 2nd, Revised edition edn. Princeton university press.

Gallup, G. (1955) 'The Role of the public opinion poll', in *Free Press and Society/ed. by G. Bird, F. Merwin. New York: Prentice-Hall*, p. 20.

Gallup, G. (1957) 'The changing climate for public opinion research', *Public Opinion Quarterly*, 21(1), pp. 23–27.

Gallup, G. (1976) *The sophisticated poll watcher's guide*. Rev. ed. Princeton, N.J: Princeton Opinion Press.

Gallup, G. and Rae, S.F. (1968) *The pulse of democracy: the public-opinion poll and how it works*. 2. Repr. Westport, Conn: Greenwood Pr.

Garfinkel, Harold. (1967) *Studies in ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.

Gelman, A. (2014) 'President of American Association of Buggy-Whip Manufacturers takes a strong stand against internal combustion engine, argues that the so-called "automobile" has "little grounding in theory" and that "results can vary widely based on the particular fuel that is used."', *Statistical modeling, causal inference, and social science blog*. Available at: https://statmodeling.stat.columbia.edu/2014/08/06/president-american-association-buggy-whip-manufacturers-takes-strong-stand-internal-combustion-engine-argues-called-automobile-little-grounding-theory/.

Gelman, A. and Rothschild, D. (2014) 'Modern polling needs innovation, not traditionalism', *Washington Post*. Available at: https://www.washingtonpost.com/news/monkey-cage/wp/2014/08/04/modern-polling-requires-both-sampling-and-adjustment/.

Gieryn, T.F. (1983) 'Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists', *American Sociological Review*, 48(6), p. 781. Available at: https://doi.org/10.2307/2095325.

Gieryn, T.F. (1999) *Cultural boundaries of science: credibility on the line*. Chicago: University of Chicago Press.

Ginsberg, B. (1989) 'How polling transforms public opinion', *Manipulating public opinion: Essays on public opinion as a dependent variable*, pp. 271–293.

Godfrey-Smith, P. (2003) *Theory and reality: an introduction to the philosophy of science*. Chicago: University of Chicago Press (Science and its conceptual foundations).

Goffman, E. (1956) *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh.

Goldthorpe, J.H. (2021) *Pioneers of Sociological Science: Statistical Foundations and the Theory of Action*. 1st edn. Cambridge New York (N.Y.): Cambridge University Press. Available at: https://doi.org/10.1017/9781108933254.

Graunt, J. (1977) 'Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality', in, pp. 11–20. Available at: https://doi.org/10.1007/978-3-642-81046-6_2.

Griessl, L. (2022) '(Dis)Assembling Predictive Stability: On the History and Culture of Survey Sampling for Election Forecasts', *Conference Proceedings of the STS Conference Graz 2022*.

Edited by G. Gethinger, M. Jahrmacher, and F. Häller, pp. 109–127. Available at: https://doi.org/10.3217/978-3-85125-932-2.

Griessl, L. (2023) 'From Skopein to Scraping: Probability, Agency, and the Politics of Public Opinion Research', *Parallax*, 28(04), pp. 396–410. Available at: https://doi.org/10.1080/13534645.2023.2206239.

Grommé, F. and Scheel, S. (2020) 'Doing statistics, enacting the nation: The performative powers of categories', *Nations and Nationalism*, (October 2018), pp. 1–18. Available at: https://doi.org/10.1111/nana.12596.

Gross, M. (2021) 'The Largest Possible Experiment: The Corona Pandemic as Nonknowledge Transfer', *The European Sociologist* [Preprint], (46(2)). Available at: https://europeansociology.org/european-sociologist/issue/46/discussion/fe72fec8-ea0b-4089-a610-1eae552ed78c.

Gross, M. and McGoey, L. (eds) (2015) *Routledge international handbook of ignorance studies*. London ; New York: Routledge, Taylor & Francis Group (Routledge international handbooks).

Gross, M. and McGoey, L. (eds) (2022) *Routledge handbook of ignorance studies*. Second Edition. London ; New York: Routledge/Taylor & Francis Group (Routledge international handbooks).

Groves, R.M. (ed.) (2009) *Survey methodology*. 2nd ed. Hoboken, N.J: Wiley (Wiley series in survey methodology).

Groves, R.M. (2011) 'Three Eras of Survey Research', *Public Opinion Quarterly*, 75(5), pp. 861–871. Available at: https://doi.org/10.1093/poq/nfr057.

Groves, R.M. and Lyberg, L. (2010) 'Total Survey Error: Past, Present, and Future', *Public Opinion Quarterly*, 74(5), pp. 849–879. Available at: https://doi.org/10.1093/poq/nfq065.

Habermas, J. (1999) *The Structural transformation of the public sphere: an inquiry into a category of bourgeois society*. 10. print. Cambridge, Mass: MIT Press (Studies in contemporary German social thought).

Hacking, I. (1983) *Representing and intervening: introductory topics in the philosophy of natural science*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Hacking, I. (1991) 'How should we do the history of statistics?', in Graham. Burchell et al. (eds) *The Foucault Effect: Studies in Governmentality. With Two Lectures By and an Interview with Michel Foucault*. Chicago: Chicago University Press Chicago, pp. 181–197.

Hacking, I. (1996) 'The looping effects of human kinds', in D. Sperber, D. Premack, and A.J. Premack (eds) *Causal Cognition*. Oxford University Press, pp. 351–383. Available at: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198524021.001.0001/acprof-9780198524021.

Hacking, I. (2013) 'Making up people', *Forms of Desire: Sexual Orientation and the Social Constructionist Controversy*. Taylor and Francis. Available at: https://doi.org/10.4324/9780203760130.

Hamersveld, M. van, Bont, C. de and Hamersveld, M. van (eds) (2007) *Market research handbook*. 5. ed. Chichester: Wiley (ESOMAR world research publication).

Hannah, M.G. (2001) 'Sampling and the politics of representation in US Census 2000', *Environment and Planning D: Society and Space*, 19(5), pp. 515–534. Available at: https://doi.org/10.1068/d289.

Hansen, M.H., Dalenius, T. and Tepping, B.J. (1985) 'The Development of Sample Surveys of Finite Populations', in A.C. Atkinson and S.E. Fienberg (eds) *A Celebration of Statistics*. New York, NY: Springer New York, pp. 327–354. Available at: https://doi.org/10.1007/978-1-4613-8560-8_13.

Hays, R.D., Liu, H. and Kapteyn, A. (2015) 'Use of Internet panels to conduct surveys', *Behavior Research Methods*, 47(3), pp. 685–690. Available at: https://doi.org/10.3758/s13428-015-0617-9.

Heide, L. (2009) *Punched-card systems and the early information explosion, 1880-1945*. Baltimore: Johns Hopkins University Press (Studies in industry and society).

Heintz, B. (1993) 'Wissenschaft im Kontext: neuere Entwicklungstendenzen der Wissenschaftssoziologie', *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 45(3).

Herbst, S. (1992) 'Surveys in the Public Sphere: Applying Bourdieu's Critique of Opinion Polls', *International Journal of Public Opinion Research*, 4(3), pp. 220–229. Available at: https://doi.org/10.1093/ijpor/4.3.220.

Herbst, S. (1993) *Numbered voices: how opinion polling has shaped American politics*. Chicago: University of Chicago Press (American politics and political economy series).

Herbst, S. (2011) '(Un)Numbered Voices? Reconsidering the Meaning of Public Opinion in a Digital Age', in R.K. Goidel (ed.) *Political polling in the digital age: the challenge of measuring and understanding public opinion*. Baton Rouge: Louisiana State Univ. Press, pp. 86–98.

Hicks, W. (1949) *World War I and American public opinion, 1914-1917*. University of Louisville. Available at: https://doi.org/10.18297/etd/1973.

Igo, S.E. (2007) *The averaged American: surveys, citizens, and the making of a mass public*. Cambridge, Mass: Harvard University Press.

IIS (1895) *Bulletin de l'Institut international de statistique. TOME IX*. Den Haag: Institut international de statistique. Available at: https://doi.org/ark:/12148/bpt6k61560p.

Ippen, D. (2016) 'Zeit und Zeitung – Erlebtes aus 8 deutschen Jahrzehnten'. Evangelische Akademie Tutzing. Available at: https://www.ev-akademie-tutzing.de/wp-content/uploads/2016/03/Kanzelrede_Dirk-Ippen_formatiert1.pdf (Accessed: 6 March 2023).

Irvine, J., Miles, I. and Evans, J. (1979) *Demystifying social statistics*. Pluto Press.

Jacobs, L.R. and Shapiro, R.Y. (2000) *Politicians don't pander: political manipulation and the loss of democratic responsiveness*. Chicago, IL: University of Chicago Press (Studies in communication, media, and public opinion).

Jasanoff, S. (1995) *Science at the Bar: Law, Science, and Technology in America*. Cambridge, MA: Harvard University Press.

Jasanoff, S. (ed.) (2010) *States of knowledge: the co-production of science and social order*. transferred to digital print. London: Routledge (International library of sociology).

Jensen, A., Saenger, K. and Bowley, A.L. (1926) 'Rapport de M. Jensen sur l'Application de la méthode représentative', *Bulletin de l'Institut international de statistique*, 22, pp. 58–61.

Jones, M. *et al.* (eds) (1990) *Fake? the art of deception*. Berkeley: University of California Press.

Kalton, G. (2019) 'Developments in Survey Research over the Past 60 Years: A Personal Perspective', *International Statistical Review*, 87(S1), pp. S10–S30. Available at: https://doi.org/10.1111/insr.12287.

Katz, D. and Cantril, H. (1937) 'Public Opinion Polls', *Sociometry*, 1(1/2), pp. 155–179. Available at: https://doi.org/10.2307/2785264.

Keller, F. (2001) *Archäologie der Meinungsforschung*. Konstanz: UVK Verlagsgesellschaft mbH.

Kennedy, C. and Hartig, H. (2019) 'Response rates in telephone surveys have resumed their decline'. Available at: https://policycommons.net/artifacts/616911/response-rates-in-telephone-surveys-have-resumed-their-decline/1597639/.

Kennedy, C., Popky, D. and Keeter, S. (2023) 'How Public Polling Has Changed in the 21st Century', 19 April. Available at: https://www.pewresearch.org/methods/2023/04/19/how-public-polling-has-changed-in-the-21st-century/.

Keverberg, B. de (1827) 'Notes sur Quetelet', *Nouveaux Mémoires de l'Academie royal des sciences et belles-lettres de Bruxelles*, 4, pp. 175–192.

Kiær, A. (1899) 'Sur les méthodes représentatives ou typologiques appliquées à la statistique', *Bulletin de l'Institut International de Statistique*, 11(1), pp. 180–185.

Knorr-Cetina, K. (1999) *Epistemic cultures: how the sciences make knowledge*. Cambridge, Mass: Harvard University Press.

Krippendorff, K. (2005) 'The Social Construction of Public Opinion', in *Kommunikation über Kommunikation*. Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 129–149. Available at: https://doi.org/10.1007/978-3-322-80821-9_10.

Kruger, L., Daston, L.J. and Heidelberger, M. (eds) (1987) *The probabilistic revolution, Vol. 1: Ideas in history; Vol. 2: Ideas in the sciences.*, *The probabilistic revolution, Vol. 1: Ideas in history; Vol. 2: Ideas in the sciences.* Cambridge, MA, US: The MIT Press.

Kruskal, W. and Mosteller, F. (1979a) 'Representative Sampling, I: Non-Scientific Literature', *International Statistical Review / Revue Internationale de Statistique*, 47(1), p. 13. Available at: https://doi.org/10.2307/1403202.

Kruskal, W. and Mosteller, F. (1979b) 'Representative Sampling, II: Scientific Literature, Excluding Statistics', *International Statistical Review / Revue Internationale de Statistique*, 47(2), p. 111. Available at: https://doi.org/10.2307/1402564.

Kruskal, W. and Mosteller, F. (1979c) 'Representative Sampling, III: The Current Statistical Literature', *International Statistical Review / Revue Internationale de Statistique*, 47(3), p. 245. Available at: https://doi.org/10.2307/1402647.

Kruskal, W. and Mosteller, F. (1980) 'Representative Sampling, IV: The History of the Concept in Statistics, 1895-1939', *International Statistical Review / Revue Internationale de Statistique*, 48(2), p. 169. Available at: https://doi.org/10.2307/1403151.

Kuhn, T.S. (1996) *The structure of scientific revolutions*. Chicago; London: University of Chicago Press.

Kuhn, T.S. (2012) *The structure of scientific revolutions*. Fourth edition. Chicago ; London: The University of Chicago Press.

Kuru, O., Pasek, J. and Traugott, M.W. (2020) 'When Polls Disagree: How Competitive Results and Methodological Quality Shape Partisan Perceptions of Polls and Electoral Predictions', *International Journal of Public Opinion Research*, 32(3), pp. 586–603. Available at: https://doi.org/10.1093/ijpor/edz035.

Kuusela, V. (2011) *Paradigms in Statistical Inference for Finite Populations Up to the 1950s*. Statistics Finland.

Laclau, E. (2007) *Emancipation(s)*. London: Verso (Radical thinkers, 19).

Lakatos, I. (1978) *The Methodology of Scientific Research Programmes: Philosophical Papers*. 1st edn. Edited by J. Worrall and G. Currie. Cambridge ; New York: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511621123.

Lakatos, I. (2015) *Proofs and Refutations*, *Proofs and Refutations*. Edited by J. Worrall and E. Zahar. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781316286425.

Laplace, P.S. (1786) 'Sur les naissances, les mariages et les mortsa Paris depuis 1771 jusqu'a 1784 et dans toute l'étendue de la France, pendant les années 1781 et 1782', *Mémoires de l'Académie Royale des Sciences présentés par divers savans*, pp. 35–46.

Latour, B. (1984) 'The Powers of Association', *The Sociological Review*, 32(1_suppl), pp. 264–280. Available at: https://doi.org/10.1111/j.1467-954X.1984.tb00115.x.

Latour, B. (1986) '« Visualisation and Cognition: Drawing Things Together »', in H. Kuklick (ed.) *Knowledge and Society - Studies in the Sociology of Culture Past and Present*. Greenwich, Conn.: Jai Press, pp. 1–40.

Latour, B. (1993) *We have never been modern*. Cambridge, Mass: Harvard University Press.

Latour, B. (1996) 'On actor-network theory: A few clarifications', *Soziale Welt*, 47(4), pp. 369–381.

Latour, B. (1999) *Pandora's hope: essays on the reality of science studies*. Cambridge, Mass: Harvard University Press.

Latour, B. (2004) 'Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern', *Critical Inquiry*, 30(2), pp. 225–248. Available at: https://doi.org/10.1086/421123.

Latour, B. (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press. Available at: http://www.amazon.com/Reassembling-Social-Introduction-Actor-Network-Theory-Management/dp/0199256047.

Latour, B. (2014) 'What Is the Style of Matters of Concern?', in N. Gaskill and A.J. Nocek (eds) *The Lure of Whitehead*. University of Minnesota Press, pp. 92–126. Available at: https://doi.org/10.5749/minnesota/9780816679959.003.0004.

Latour, B. and Woolgar, S. (1979) *Laboratory life : the social construction of the scientific facts*. Beverly Hills, CA: Sage Publications.

Latour, B. and Woolgar, S. (1986) *Laboratory life the construction of scientific facts*. Princeton: Princeton Univ. Press.

Law, J. (2002) *Aircraft stories: decentering the object in technoscience*. Durham, NC: Duke University Press.

Law, J. (2004) *After method: mess in social science research*. 1st ed. London: Routledge.

Law, J. (2009) 'Seeing like a survey', *Cultural Sociology*, 3(2), pp. 239–256. Available at: https://doi.org/10.1177/1749975509105533.

Law, J. and Lien, M.E. (2013) 'Slippery: Field notes in empirical ontology', *Social Studies of Science*, 43(3), pp. 363–378. Available at: https://doi.org/10.1177/0306312712456947.

Law, J. and Urry, J. (2004) 'Enacting the social', *Economy and Society*, 33(3), pp. 390–410. Available at: https://doi.org/10.1080/0308514042000225716.

Lazarsfeld, P.F. (1961) 'Notes on the History of Quantification in Sociology--Trends, Sources and Problems', *Isis*, 52(2), pp. 277–333.

Legg, S. (2005) 'Foucault's population geographies: classifications, biopolitics and governmental spaces', *Population, Space and Place*, 11(3), pp. 137–156. Available at: https://doi.org/10.1002/psp.357.

Leiser, E. (1948) 'Blames pollsters for Dewey 'scoop'', *New York Post*. 42nd edn, 4 November.

Likert, R. (1948a) 'Opinion Studies and Government Policy', *Proceedings of the American Philosophical Society*, 92(5), pp. 341–350.

Likert, R. (1948b) 'Public Opinion Polls', *Scientific American*, 179(6), pp. 7–11.

Lippmann, W. (1993) *The phantom public*. New Brunswick, N.J., U.S.A: Transaction Publishers (The Library of conservative thought).

Lowenthal, L. and Guterman, N. (2021) *Prophets of deceit: a study of the techniques of the American agitator*. London: Verso.

Lusinchi, D. (2012) '"President" Landon and the 1936 Literary Digest Poll: Were Automobile and Telephone Owners to Blame?', *Social Science History*, 36(1), pp. 23–54. Available at: https://doi.org/10.1215/01455532-1461650.

Lusinchi, D. (2016) 'Using online panels for election polls'. Available at: http://sociological-improvisations.farwestresearch.com/?p=2.

Lusinchi, D. (2017a) 'Are online polls "junk science"?' Available at: http://sociological-improvisations.farwestresearch.com/?p=4.

Lusinchi, D. (2017b) 'The Rhetorical use of Random Sampling: Crafting and Communicating the Public Image of Polls as a Science (1935-1948)', *Journal of the History of the Behavioral Sciences*, 53(2), pp. 113–132. Available at: https://doi.org/10.1002/jhbs.21836.

Lusinchi, D. (2018) '"The Great Fiasco" of the 1948 presidential election polls: status recognition and norms conflict in social science', *Annals of Science*, 75(2), pp. 120–144. Available at: https://doi.org/10.1080/00033790.2018.1466194.

Lusinchi, D. (2021) 'Kiær and the rebirth of the representative method: A case-study in controversy management at the International Statistical Institute (1895–1903)', *Journal of the History of the Behavioral Sciences*, (June), pp. 1–20. Available at: https://doi.org/10.1002/jhbs.22145.

Lyberg, L.E. and Weisberg, H.F. (2016) 'Total Survey Error: A Paradigm for Survey Methodology', in C. Wolf et al. (eds) *The SAGE Handbook of Survey Methodology*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd, pp. 27–40. Available at: https://doi.org/10.4135/9781473957893.

Lynch, M. (2012) 'Self-exemplifying revolutions? Notes on Kuhn and Latour', *Social Studies of Science*, 42(3), pp. 449–455. Available at: https://doi.org/10.1177/0306312712439120.

Lynch, M. (2020) 'We Have Never Been Anti-Science: Reflections on Science Wars and Post-Truth', *Engaging Science, Technology, and Society*, 6, p. 49. Available at: https://doi.org/10.17351/ests2020.309.

Lynn, P. (2020) 'Evaluating push-to-web methodology for mixed-mode surveys using address-based samples', *Survey Research Methods*, pp. 19–30. Available at: https://doi.org/10.18148/SRM/2020.V14I1.7591.

Mahfoud, T. *et al.* (2018) 'The limits of dual use', *Issues in Science and Technology*, 34(4), pp. 73–78.

Mahfoud, T., McLean, S. and Rose, N. (2017) 'Preface', in *Progress in Brain Research*. Elsevier, pp. xi–xx. Available at: https://doi.org/10.1016/S0079-6123(17)30098-5.

Mannheim, K. (1998) *Ideology and utopia: an introduction to the sociology of knowledge*. Repr. London and Henley: Routledge & K. Paul.

Manza, J. and Cook, F.L. (2002) 'A Democratic Polity?: Three Views of Policy Responsiveness to Public Opinion in the United States', *American Politics Research*, 30(6), pp. 630–667. Available at: https://doi.org/10.1177/153267302237231.

Margolis, M. and Mauser, G.A. (eds) (1989) *Manipulating public opinion: essays on public opinion as a dependent variable*. Pacific Grove, Calif: Brooks/Cole Pub. Co.

marktforschung.de (2018) 'Beschwerde beim Presserat wegen Civey-Umfrage', *marktforschung.de*, 1 October. Available at: https://www.marktforschung.de/aktuelles/marktforschung/beschwerde-beim-presserat-wegen-civey-umfrage/.

Marsh, Catherine. (1982) 'The survey method : the contribution of surveys to sociological explanation'. London; Boston: Allen & Unwin.

Martens, M. (2018) 'Deutsche und Passdeutsche', *FAZ*, 29 July. Available at: https://www.faz.net/aktuell/politik/inland/warum-die-rede-von-passdeutschen-unangemessen-ist-15712646-p2.html (Accessed: 30 May 2023).

Marx, K. (1993) *Grundrisse: foundations of the critique of political economy*. London: Penguin books (Penguin Classics).

McGoey, L. (2019) *The unknowers: how strategic ignorance rules the world*. London: Zed Books Ltd.

Mermin, N.D. (2008) 'Science wars revisited', *Nature*, 454(7202), pp. 276–277. Available at: https://doi.org/10.1038/454276a.

Merton, R.K. (1974) *The sociology of science: theoretical and empirical incvestigations*. Chicago: Univ. of Chicago Pr.

Mohr, M. (2021) 'Politische Zukunft von Sebastian Kurz in Österreich 2021', *Statista*, 13 December. Available at: https://de.statista.com/statistik/daten/studie/1272248/umfrage/politische-zukunft-von-sebastian-kurz-in-oesterreich/.

Mol, A. (1999) 'Ontological Politics. A Word and Some Questions', *The Sociological Review*, 47(1_suppl), pp. 74–89. Available at: https://doi.org/10.1111/j.1467-954X.1999.tb03483.x.

Mondon, A. (2022) 'Populism, public opinion, and the mainstreaming of the far right: The "immigration issue" and the construction of a reactionary "people"', *Politics*, pp. 1–18. Available at: https://doi.org/10.1177/02633957221104726.

Mouat, F.J. (1885) 'History of the Statistical Society of London', *Journal of the Statistical Society of London*, pp. 14–371.

mrweb (2007) 'Two More Buys and £27m Share Placing for YouGov', *mrweb*, 7 August. Available at: https://www.mrweb.com/drno/news7148.htm.

Newsome, A.R. (1939) *The Presidential Election of 1824 in North Carolina*. Chapel Hill: University of North Carolina Press (James Sprunt studies in history and political science).

Neyman, J. (1934) 'On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection', *Journal of the Royal Statistical Society*, 97(4), pp. 558–625. Available at: https://doi.org/10.2307/2342192.

Osborne, T. and Rose, N. (1999) 'Do the social sciences create phenomena?: the example of public opinion research', *The British Journal of Sociology*, 50(3), pp. 367–396. Available at: https://doi.org/10.1111/j.1468-4446.1999.00367.x.

Özil, M., @M10 (2018) 'Meeting President Erdogan', 22 July. Available at: https://twitter.com/M10/status/1020984884431638528.

Pack, M. (2022) *Polling unpacked: the history, uses and abuses of political opinion polls*. London: Reaktion Books.

Page, B.I. and Shapiro, R.Y. (1992) *The rational public: fifty years of trends in Americans' policy preferences*. Chicago: University of Chicago Press (American politics and political economy series).

Pausch, R. and Zimmermann, F. (2020) 'Kampf der Torten', *Die ZEIT*, 5 February, pp. 1–7.

Perrin, A.J. and McFarland, K. (2011) 'Social Theory and Public Opinion', *Annual Review of Sociology*, 37(1), pp. 87–107. Available at: https://doi.org/10.1146/annurev.soc.012809.102659.

Pinch, T.J. and Bijker, W.E. (1984) 'The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other', *Social Studies of Science*, 14(3), pp. 399–441. Available at: https://doi.org/10.1177/030631284014003004.

Pitkin, H. (1967) *The Concept of Representation*. Reprint 2019. Berkeley, CA: University of California Press.

Planck, M. (1949) *Scientific autobiography and other papers*. Translated by F. Gaynor. New York: Philosophical Library.

Poovey, M. (1998) *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society*. Chicago: University of Chicago Press.

Porter, T.M. (1986) *The rise of statistical thinking, 1820-1900*. Princeton, N.J: Princeton University Press.

Porter, T.M. (1995) *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton (NJ): Princeton university press.

Prince, R. (2019) 'The geography of statistics: Social statistics from moral science to big data', *Progress in Human Geography*, 44(6), pp. 1047–1065. Available at: https://doi.org/10.1177/0309132519873421.

Proctor, R. and Schiebinger, L.L. (eds) (2008) *Agnotology: the making and unmaking of ignorance*. Stanford, Calif: Stanford University Press.

Quetelet, A. (1846) *Lettres à SAR le duc régnant de Saxe-Coburg et Gotha: sur la théorie des probabilités, appliquée aux sciences morales et politiques*. M. Hayez.

Raben, F. and Francovic, K. (2019) 'Fake polls – how to spot them?', *Research World*, 21 November. Available at: https://archive.researchworld.com/fake-polls/.

Radermacher, W. (2020) *Official statistics 4.0: verified facts for people in the 21st century*. Cham, Switzerland: Springer.

Rancière, J. (1999) *Disagreement: politics and philosophy*. Minneapolis, Minn.: Univ. of Minnesota Press.

Rancière, J., Panagia, D. and Bowlby, R. (2001) 'Ten Theses on Politics', *Theory & Event*, 5(3). Available at: https://doi.org/10.1353/tae.2001.0028.

Randeraad, N. (2011) 'The International Statistical Congress (1853—1876): Knowledge Transfers and their Limits', *European History Quarterly*, 41(1), pp. 50–65. Available at: https://doi.org/10.1177/0265691410385759.

Raphael, L. (1996) 'Die Verwissenschaftlichung des Sozialen als methodische und konzeptionelle Herausforderung für eine Sozialgeschichte des 20. Jahrhunderts', *Geschichte und Gesellschaft*, 22(2), pp. 165–193.

Ravetz, J. (2015) 'From Descartes to Rumsfeld', in M. Gross and L. McGoey (eds) *Routledge International Handbook of Ignorance Studies*. London ; New York: Routledge, Taylor & Francis Group, pp. 53–60.

Rendtel, U. (2018) *Gutachten zur Repräsentativität von Online-Umfragen*, pp. 1–7. Available at: https://assets.ctfassets.net/ublc0iceiwck/3JEuIxMxJC2WEG2A82aKQ6/e1567340430a3b4caee 8515ff000775d/Gutachten_Repraesentativitaet_Online_Umfragen_Rendtel.pdf (Accessed: 24 September 2023).

Rivers, D. (2007) 'Sampling for web surveys', in. *Joint Statistical Meetings*, American Statistical Association Alexandria, VA.

Roper, E. and Woodward, J.L. (1948) 'Democracy's auxiliary ballot box', *New York Herald Tribune*, 4.

Rosenbaum, P.R. and Rubin, D.B. (1983) 'The central role of the propensity score in observational studies for causal effects', *Biometrika*, 70(1), pp. 41–55. Available at: https://doi.org/10.1093/biomet/70.1.41.

Rousseau, J.-J. (2002) *The social contract: and, The first and second discourses*. New Haven: Yale University Press (Rethinking the Western tradition).

Saris, W.E. (1998a) 'Ten Years of Interviewing Without Interviewers: The Telepanel', in C. Mick P. et al. (eds) *Computer Assisted Survey Information Collection*. John Wiley & Sons.

Saris, W.E. (1998b) 'The public opinion about the EU can easily be swayed in different directions', *Acta politica*, 33(4), pp. 406–435.

Saris, W.E. and De Pijper, W.M. (1986) 'Computer assisted interviewing using home computers', *European Research*, 14(3), pp. 144–150.

Saussure, F. de (2011) *Course in general linguistics*. Edited by P. Meisel and H. Saussy. Translated by W. Baskin. New York: Columbia University Press.

Savage, M. (2010) *Identities and social change in Britain since 1940: the politics of method*. Oxford ; New York, NY: Oxford University Press.

Savage, M. and Burrows, R. (2007) 'The Coming Crisis of Empirical Sociology', *Sociology*, 41(5), pp. 885–899. Available at: https://doi.org/10.1177/0038038507080443.

Saward, M. (2006) 'The Representative Claim', *Contemporary Political Theory*, 5(3), pp. 297–318. Available at: https://doi.org/10.1057/palgrave.cpt.9300234.

Scheppe, M. (2023) 'Zwei Umfrageinstitute führen etliche Prozesse in vier Jahren', *Handelsblatt*, 10 March. Available at: https://www.handelsblatt.com/unternehmen/it-medien/forsa-vs-civey-zwei-umfrageinstitute-fuehren-etliche-prozesse-in-vier-jahren-/29019918.html.

Schmitt, L.H. (2004) 'Die Messung des Konsumentenverhaltens der Internetnutzer—ein Methodenvergleich', in K.-P. Wiedmann et al. (eds) *Konsumentenverhalten im Internet: Konzepte—Erfahrungen—Methoden*. Springer, pp. 273–290.

Schonlau, M., Fricker, R.D. and Elliott, M.N. (2002) 'Internet Survey Case Studies', in *Conducting Research Surveys via E-mail and the Web*. RAND Corporation, pp. 55–72.

Schüller, K., Koch, H. and Rampelt, F. (2021) *Data-Literacy-Charta*. Stifterverband Berlin. Available at: https://www.stifterverband.org/sites/default/files/data-literacy-charter.pdf.

Schuman, H. and Presser, S. (1980) 'Public opinion and public ignorance: The fine line between attitudes and nonattitudes', *American Journal of Sociology*, 85(5), pp. 1214–1225.

Sciortino, L. (2017) 'On Ian Hacking's Notion of Style of Reasoning', *Erkenntnis*, 82(2), pp. 243–264. Available at: https://doi.org/10.1007/s10670-016-9815-9.

Seng, Y.P. (1951) 'Historical Survey of the Development of Sampling Theories and Practice', *Journal of the Royal Statistical Society. Series A (General)*, 114(2), p. 214. Available at: https://doi.org/10.2307/2980977.

Seymour, G. *et al.* (1949) 'Should Public Opinion Polls Make Election Forecasts?: A Symposium', *Journalism Quarterly*, 26(2), pp. 131–144. Available at: https://doi.org/10.1177/107769904902600201.

Shahrokni, N. (2012) 'The politics of polling: Polling and the constitution of counter-publics during "reform" in Iran', *Current Sociology*, 60(2), pp. 202–221. Available at: https://doi.org/10.1177/0011392111429222.

Shapin, S. and Schaffer, S. (2011) *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Princeton, N.J: Princeton University Press.

Simmel, G. (1950) *The sociology of Georg Simmel*. Edited and translated by K.H. Wolff. New York London: Free press C. Macmillan.

Simon, P. (2012) 'Collecting ethnic statistics in Europe: a review', *Ethnic and Racial Studies*, 35(8), pp. 1366–1391. Available at: https://doi.org/10.1080/01419870.2011.607507.

Simons, J. (1999) 'Are Political Polls Via Internet Reliable ? Yes ? No ? Maybe ?', *Wall Street Journal*, 13 April. Available at: https://www.wsj.com/articles/SB923957499521692.

Sintomer, Y. (2013) 'The Meanings of Political Representation: Uses and Misuses of a Notion', *Raisons politiques*, 50(2), pp. 13–34. Available at: https://doi.org/10.3917/rai.050.0013.

Sismondo, S. and Chrisman, N. (2001) 'Deflationary Metaphysics and the Natures of Maps', *Philosophy of Science*, 68(S3), pp. S38–S49. Available at: https://doi.org/10.1086/392896.

Sismondo, Sergio. (2011) *An Introduction to Science and Technology Studies.* Somerset: Wiley.

Smith, R.N. (1982) *Thomas E. Dewey and his times*. New York: Simon and Schuster.

Smith, T.W. (1990) 'The First Straw?: A Study of the Origins of Election Polls', *Public Opinion Quarterly*, 54(1), p. 21. Available at: https://doi.org/10.1086/269181.

Sokal, A.D. (1996) 'Transgressing the boundaries: Toward a transformative hermeneutics of quantum gravity', *Social text*, (46/47), pp. 217–252.

Sproull, L. and Kiesler, S. (1986) 'Reducing Social Context Cues : Electronic Mail in Organizational Communications', *Management Science*, 32(11), pp. 1492–1512.

Statactivisme: comment lutter avec des nombres (2014). Paris: Zones, Éditions La Découverte.

Statista (2020) *Industry revenue of "marketing research and public opinion polling" in the U.S. from 2012 to 2024*. Statista; US Census Bureau. Available at: https://www.statista.com/forecasts/409757/marketing-research-and-public-opinion-polling-revenue-in-the-us.

Stigler, S.M. (1986) *The history of statistics : the measurement of uncertainty before 1900.* Cambridge, Mass.: Belknap Press of Harvard University Press.

Stilgoe, J. (2020) *Who's Driving Innovation?: New Technologies and the Collaborative State*. Cham: Springer International Publishing. Available at: https://doi.org/10.1007/978-3-030-32320-2.

Sturgis, P. *et al.* (2018) 'An assessment of the causes of the errors in the 2015 UK general election opinion polls', *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(3), pp. 757–781. Available at: https://doi.org/10.1111/rssa.12329.

Sturgis, P. and Smith, P. (2010) 'Fictitious Issues Revisited: Political Interest, Knowledge and the Generation of Nonattitudes', *Political Studies*, 58(1), pp. 66–84. Available at: https://doi.org/10.1111/j.1467-9248.2008.00773.x.

Swedlow, B. (2007) 'Using the boundaries of science to do boundary-work among scientists: Pollution and purity claims', *Science and Public Policy*, 34(9), pp. 633–643. Available at: https://doi.org/10.3152/030234207X264953.

Taylor, H.C. and Taylor, A.D. (1952) *The story of agricultural economics in the United States, 1840-1932: men, services, ideas*. Ames: Iowa State College Press.

Taylor, H. and Terhanian, G. (1999) 'No Witchcraft Here', *Public Perspective*, 10, p. 42.

Terhanian, G. *et al.* (2000) 'Correcting data from online surveys for the effects of nonrandom selection and nonrandom assignment', *Harris Interactive White Paper*, pp. 1–13.

The New York Times (2008) *The New York Times Polling Standards*. Available at: https://archive.nytimes.com/www.nytimes.com/ref/us/politics/10_polling_standards.html (Accessed: 24 September 2023).

Thomas, D. (2023) 'YouGov considers US listing as business expands', *Financial Times*, 14 August. Available at: https://www.ft.com/content/31a4e40d-1153-4573-a9b5-460bc212f47a (Accessed: 24 September 2023).

Tillé, Y. (2020) *Sampling and estimation from finite population*. Hoboken, NJ: Wiley (Wiley series in probability and statistics applied. Probability and statistics section).

Tilly, C. (1983) 'Speaking Your Mind Without Elections, Surveys, or Social Movements', *The Public Opinion Quarterly*, 47(4), pp. 461–478.

Ulmer, B. (1988) 'Konversionserzählungen als rekonstruktive Gattung: Erzählerische Mittel und Strategien bei der Rekonstruktion eines Bekehrungserlebnisses', *Zeitschrift für Soziologie*, 17(1), pp. 19–33. Available at: https://doi.org/10.1515/zfsoz-1988-0102.

Urbinati, N. (2006) *Representative democracy: principles and genealogy*. Chicago: University of Chicago Press.

Van Belle, G. (2008) *Statistical rules of thumb*. 2nd ed. Hoboken, N.J: Wiley (Wiley series in probability and statistics).

Verba, S. (1996) 'The Citizen as Respondent: Sample Surveys and American Democracy Presidential Address, American Political Science Association, 1995', *American Political Science Review*, 90(1), pp. 1–7. Available at: https://doi.org/10.2307/2082793.

Vogel, R. (2019) *Survey-Welten*. Wiesbaden: Springer Fachmedien Wiesbaden. Available at: https://doi.org/10.1007/978-3-658-25437-7.

Walton, D. (1997) 'What Is Propaganda, and What Exactly Is Wrong with It', *Public Affairs Quarterly*, 11(4), pp. 383–413.

Wang, W. *et al.* (2015) 'Forecasting elections with non-representative polls', *International Journal of Forecasting*, 31(3), pp. 980–991. Available at: https://doi.org/10.1016/j.ijforecast.2014.06.001.

Warner, M. (2005) *Publics and counterpublics*. New York, NY: Zone Books.

Weber, M. (2019) *Economy and society: a new translation*. Translated by K. Tribe. Cambridge, Massachusetts: Harvard University Press.

Wenger, E. (1998) *Communities of practice: learning, meaning, and identity*. Cambridge, U.K. ; New York, N.Y: Cambridge University Press (Learning in doing).

Westergaard, H. (1916) 'Scope and Method of Statistics', *Publications of the American Statistical Association*, 15(115), p. 229. Available at: https://doi.org/10.2307/2964849.

Wikipedia contributors (2023) 'YouGov — Wikipedia, The Free Encyclopedia'. Available at: https://en.wikipedia.org/w/index.php?title=YouGov&oldid=1174073052.

Wildner, R. (2020) 'Herausforderungen bei der Stichprobenziehung in der Markt- und Sozialforschung', *planung&analyse*, November. Available at: https://www.horizont.net/planung-analyse/nachrichten/-herausforderungen-bei-der-stichprobenziehung-in-der-markt--und-sozialforschung-186468.

Wodak, R. (2022) 'Shameless normalization as a result of media control: The case of Austria', *Discourse & Society*, pp. 1–17. Available at: https://doi.org/10.1177/09579265221095419.

Woolgar, S. and Lezaun, J. (2013) 'The wrong bin bag: A turn to ontology in science and technology studies?', *Social Studies of Science*, 43(3), pp. 321–340. Available at: https://doi.org/10.1177/0306312713488820.

Wright, L.A. (2020) *Star power: American democracy in the age of the celebrity candidate*. New York, NY: Routledge (Media and power).

Yeager, D.S. *et al.* (2011) 'Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples', *Public Opinion Quarterly*, 75(4), pp. 709–747. Available at: https://doi.org/10.1093/poq/nfr020.

Young, I.M. (2000) *Inclusion and democracy*. Oxford ; New York: Oxford University Press (Oxford political theory).

Yudin, G. (2020) 'Governing Through Polls: Politics of Representation and Presidential Support in Putin's Russia', *Javnost*, 27(1), pp. 2–16. Available at: https://doi.org/10.1080/13183222.2020.1675434.

11 Appendix

Table 1: *Initial Codes*

Initial Code	Sample Quote(s)
Positioning in Favor of Probability Sampling	And the and the the discovery that randomization could do that, I mean, there's a there's a moment in history, in scientific history when that discovery was made and that the application of that brilliant insight to experiments is obviously well understood, but not so well implemented. So in in medical research and other fields, sometimes the idea that we would randomly assigned people to the treatment group versus the control group, that the presumption in many studies that got published was, oh, as long as there there are two different groups of people, you don't need to actually flip a coin or truly randomize. But we've now come to recognize that that's the brilliance of randomization, is that you really need to do it. (Krosnick, Interview)
Positioning Against Non-Probability Sampling	There is no empirical evidence that it works. The procedure is not documented. It is not comprehensible. And the fascinating thing is that they then find lunatics who believe it. It's just like homeopathy. Yes, you have someone there who is fooling you. And of course, there are people who believe it. It's cheap. It's fast. It's sexy. But the problem is, they don't explain it to you so that it's comprehensible. There's no reproductive proof. There's no prognostic power. And the interesting thing, the really sociologically interesting thing is that there are people who believe them, so it's really believing. There is no mathematics behind it, nor is there empirically convincing evidence that it works. And you see this persistence despite the absence of evidence with homeopaths for, I don't know, 120 years, and why should it go faster in statistics? (Schnell, Interview, Translation).

	But there are really two fundamental elements of any new approach, new method or technology. One is the theoretisism and the other is empiricist. We can try to assess whether there is a sound and meaningful theoretical basis for an approach to work. That's important. And in my evaluation, find it lacking in nonprobability research. (Langer, Interview)
Positioning in Favor of Non-Probability Sampling	But ultimately, they are sitting in a glass house themselves. With a response rate of 5%, that's a laugh (Rendtel, Interview, Translation)
Positioning Against Probability Sampling	So, there was [someone, who] said that you have to make random samples and a random sample is characterized by the fact that I know exactly beforehand the probability with which each participant will be included in the sample. Sorry, there is no such thing as random sampling. And they should say that honestly and not just claim things like that, following some chimera (Wildner, Interview, Translation).
Doubting previously held assumptions	And she wrote me, you know, basically saying, Jon, how could you, this is terrible, non-random sampling is crazy. You seem to be endorsing it. It's a really bad idea. And so, wow, that was pretty striking to me because I respect her, and I trusted her, and she thought this was a dangerous pathway to go down and that we would be giving a spotlight to a methodology that would cause problems. Turns out she was right. She knew this long before I did. But what she did was basically, I mean, it's kind of like too late for me to get out. So, what she inspired me to do was to evaluate the data, to evaluate the accuracy of all the data by various different methods. So, in many ways I credit her as the person who inspired me to write all those papers […] and once I did one of them […] I kind of was stuck, I have to add to keep doing it (Krosnick, Interview).

	And, you know, so I realized what I was saying actually wasn't correct. It doesn't mean that, you know, you shouldn't do probability sampling. I think where you can and can control it is great. But there are these applications, where people are calling stuff probability sampling, but it, you know, I mean something with a five percent response rate can by no measure be called a probability sample. Calling it probability-based is marketing bs. You know, it's a self-selected sample (Rivers, Interview).
Description of the Nature of the Controversy	And then there was a discussion between Civey, Ms Mütze, Rainer Schnell and one […] from Forsa [...]. And the only thing missing was someone selling popcorn. I sat inside at the end and thought that the most professional one was the youngest, namely Ms. Mütze, who argued everything most professionally and also asked questions, while the others just mocked around and I thought that was really bad. I was a bit ashamed because that's not a culture we should have as scientists (Zinn, Interview, Translation).
The Need for good surveys	Huge numbers of decision makers will be handicapped if surveys disappear, because […] the unemployment rate and lots of other economic statistics that influence investing and decision making by businesses and government come from surveys [and that] if a democratic government wants to at least consider what the public wants it to do, what the public is willing to pay for, what the public approves of and doesn't approve of, the only way I k04/10/2023 13:01:00now to find out is surveys (Krosnick, Interview).
Changing Landscape	Actually, there's one difference with the general procedure nowadays, that is that we tried to get a random sample of the population. We wanted it to be representative and that is nowadays not the case. So, they are in general, with few

	exceptions, they are just collecting data from people who are willing to participate (Willem Saris, Interview). At the moment we don't have a reasonable science of sampling (Wildner, Interview, Translation).
Reasons for the Changing Landscape	And that's true that these days that if you draw a random sample of telephone numbers, like by random digit dialling, and then you call those telephone numbers, there are lots of phone numbers where probably there's a person at the other end who you will not interview. And that was true from the very beginning of our accuracy papers, that response rates were not 100 percent (Krosnick, Interview).
Confusion and Disbelief	There are all empirical studies on the selectivity of web surveys show the same results. They all show the same thing. And it's a mystery to me how anyone can ignore that. I stand there and marvel (Schnell, Interview, Translation). Yes, that is a good question. Why are there charlatans and why are they not always discovered (Thierhoff, Interview, Translation).
Advocacy for Balanced Perspectives	I don't know why we have to demonise it, it's not to be demonised. […] I don't know where that comes from. Maybe it's some kind of old professorial textbook attitude. I don't know. You don't always have to demonise things; you also have to see the advantages of them. And that's exactly it. And the people who collect such non-probabilistic samples should perhaps be told that they shouldn't claim that they can now explain the whole world. It is a kind of learning, a mutual learning, and that also takes place (Zinn, Interview, Translation).

Public and Published Opinion	If an institute says that something is representative, then the press is entitled to accept that and does not have to verify it further. (Güllner, Interview, Interview).
Transparency is Crucial	This is the point at which I personally, yes, I am willing to go and do battle, transparency, transparency is essential. And to claim some proprietary method of producing data that I can't tell you about, you have to trust me and believe, is antithetical to the cause of science and good research. […] And any claims of proprietary methods that cannot be disclosed, to me, is fundamentally disqualifying, for taking any data seriously. It should be thrown out the window, if we don't have a full and fair description of every piece of the methods involved (Langer, Interview).
What is Representativity	The other is the question of representativeness, and here we have to say that yes, there is no scientific definition of representativeness. But there is […] an expectation […]. So […] people have an idea of what representative means. And it is quite clear. Because representative means that the statement I make is transferable to the population. (Thierhoff, Interview)
In addition, I coded the whole interview material regarding the reference to or mentioning of years / time periods, statistical terms, institutions / companies, actors. This allowed me to, for instance, quickly turn to statements about certain periods in the history of polling and survey research.	

Table 2: *Grouping of Initial Codes to Form Themes*

Theme	Initial Code

A Conflictual Controversy	Positioning Against Non-Probability Sampling
	Positioning in Favor of Non-Probability Sampling
	Positioning Against Probability Sampling
	Positioning in Favor of Probability Sampling
	Description of the Nature of the Controversy
Past and Future Developments in the Field	Changing Landscape
	Reasons for the Changing Landscape
Boundary Work	Doubting previously held assumptions.
	Advocacy for Balanced Perspectives
	Confusion and Disbelief
	Transparency is Crucial
Narratives of Conversion	Doubting previously held assumptions.
	Reasons for the Changing Landscape

Table 3: *Relation of Themes to Research Questions*

Research Question	Themes that address question
How has the shift towards non-probability sampling occurred and which discursive positions developed in the field?	Development of the field
	Narratives of Conversion
	Nature of the Controversy
How do (power) struggles over the epistemology of polling shape competing statistical methods?	Boundary Work
	Narratives of Conversion

Ethics ETH1920-1248: Mr Lukas Griessl

Date Created	08 Apr 2020
Date Submitted	01 May 2020
Date of last resubmission	08 Jun 2020
Academic Staff	Mr Lukas Griessl
Category	Postgraduate Research Student
Supervisor	Prof Linsey Mcgoey
Project	Modelling the Unknown
Faculty	Social Sciences
Department	Sociology
Current status	Signed off under Annex B

Ethics application

Project overview

Title of project
Modelling the Unknown

Do you object to the title of your project being published?
No

Applicant(s)
Mr Lukas Griessl

Supervisor(s)
Prof Linsey Mcgoey

Prof Nick Allum

Proposed start date of research
01 Jul 2020

Expected end date
01 Jul 2022

Will this project be externally funded?
Yes

Will the research involve human participants?
Yes

Will the research use collected or generated personal data?
No

Will the research involve the use of animals?
No

Will any of the research take place outside the UK?
Yes

Project details

Summary of the project

Numbers and statistics are usually taken for granted and the conditions of their production are oftentimes not further questioned. Central to the field of science and technology studies (STS) is the assumption that we need to look at the way how science works in practice and thus on the (social) conditions under which scientific knowledge is constructed. This research is thus interested in methods of statistical measurement and its social, historical, and political embeddedness. The overarching interest of this research is on the topic of uncertainty and non-knowledge, which is explored in drawing on case studies in the field of statistics.

The rationale for choosing statistics as an object of inquiry is that statistics always deals with uncertainty and it is because of this, that methodological assumptions and decisions are made and disputed. When looking at such controversies, foundational issues of statistical practice can be brought to the forth and be looked at.

The project is of a rather theoretical nature but will be informed and supported through case studies. The study, therefore, proposes to conduct semi-structured interviews with around 40 statisticians and historians of statistics and a re-examination of historical literature on key shifts in statistical thought. Initially, I will look at two German Polling Institutes, 'Forsa' and 'Civey', between which there exists a strong controversy regarding the validity of the used methodology. The controversy touches the heart of statistical reasoning, as to how inferences that extrapolate from data to general facts are justified. The interviewees will be professional statisticians or academicians and will be asked questions on the nature of their work, their educational backgrounds, and other related matters. The interviews may also be about questions beyond the particular case studies to better understand how research institutes respond methodologically to new changes coming e.g. through Big Data and new forms of data scraping and extraction.

Research project proposal
Will the participants, either the subjects or the investigators, be involved in any activities that could be considered to be unlawful in the UK?

If the project is being undertaken outside the UK, will the participants, either the subjects or the investigators, be involved in any activities that could be considered to be unlawful in the country overseas?

Funding

Funder
Studienstiftung des deutschen Volkes (German Academic Scholarship Foundation)

Grant type

Other

Funder's award/grant reference or ID
Doctoral Scholarship

RCP project ID

Participant details

Who are the potential participants?
The involved participants will be professional researchers, working for national statistical institutes, polling organisations or other research institutes, including universities. Initially, statisticians from two German Polling Institutes, 'Forsa' and 'Civey' will be interviewed regarding the adaption and justification of their methods, the nature of their everyday work, or questions regarding the methods of their competitors.

Other interviewees will most likely be from the UK. The kinds of institutes where these people will come from will be similar to those in Germany. Which people to interview, however, will be determined after the initial interviews took place.

How will they be recruited?
Participants will most likely be from Germany and the UK and will be recruited via E-Mail. They will be chosen based on the particular case studies, their expertise and their position within existing controversies. Depending on the participants, interviewees may also recruit or recommend new interviewees from their colleagues or associated researchers.

Recruiting materials

Will participants be paid or reimbursed?
No

If yes, please provide details and justification for this payment.

How much will the participants be paid?

Could potential participants be considered vulnerable?
No

If yes, please explain how the participants could be considered vulnerable and why vulnerable participants are necessary for the research.

Could potential participants be considered to feel obliged to take part in the research?
No

If yes, please explain how the participants could feel obliged and how any possibility for coercion will be addressed.

Will the research involve individuals below the age of 18 or individuals of 18 years and over with a limited capacity to give informed consent?
No

Is a Disclosure and Barring Service (DBS) Check required?
No

If yes, has the DBS check been completed?

If your project involves children or vulnerable adults but does not require a DBS check, please explain why.

Informed consent

How will consent be obtained?
Written

If consent will be obtained in writing, please upload the written consent form for review and approval.
If consent will be obtained orally, please explain why.

Please upload a copy of the script that will be used to obtain oral consent.

If no script is available to upload please explain why.

Who will be obtaining and recording consent?
Consent will be obtained and recorded by myself (Lukas Griessl)

Please indicate at what stage in the data collection process consent will be obtained.
Consent will be obtained prior to the conduction of the interview.

If informed consent will not be obtained, explain why.

Please upload a participant information sheet.
Have you reviewed the information provided by the REO on participant information and consent?
Yes

Confidentiality and anonymity

Will you be maintaining the confidentiality and anonymity of participants whose personal data will be used in your research?
Yes

If yes, describe the arrangements for maintaining anonymity and confidentiality.

Names and further details of the participants will be pseudonymized. Certain affiliations (institutes, universities etc.) will only be referred to in a generalised way and will be numbered (e.g. Opinion Poll Institute 1).

If you are not maintaining anonymity and confidentially, please explain your reasons for not doing so.

Data access, storage and security

Describe the arrangements for storing and maintaining the security of any personal data collected as part of the project.
Audio recordings and interview transcripts will be stored in a password-protected folder on the cloud storage Box. Personal data will be collected in a separate password-protected folder on the Box cloud storage.

Please provide details of all those who will have access to the data.
Only I and my supervisors will have access to the data. After study completion the data may, however, be deposited with a data archive, such as the University of the Essex Data Archive to allow appropriate and controllable access to and sharing of the data, while maintaining privacy and confidentiality.

Risk and risk management

Risk Assessment documents

Are there any potential risks (e.g. physical, psychological, social, legal or economic) to participants or subjects associated with the proposed research?
No

If yes, please provide full details and explain what risk management procedures will be put in place to minimise the risks.

Are there any potential risks (e.g. physical, psychological, social, legal or economic) to the researchers working on the proposed research?
No

If yes, please provide full details and explain what risk management procedures will be put in place to minimise the risks.

Are there any potential reputational risks to the University as a consequence of undertaking the proposed research?
No

If yes, please provide full details and explain what risk management procedures will be put in place to minimise the risks.

Are there any other ethical issues that have not been addressed which you would wish to bring to the attention of the reviewer(s) of your application?

I would like to note that due to the Covid-19 pandemic, I will not conduct face-to-face interviews until it is allowed again. If, however, face-to-face interviews will be possible again and I decide to travel to Germany or another country for research, I will obtain a university travel insurance.

Other documents

Research abroad

Country

[Germany](#)

Site:

Statistisches Bundesamt

Forsa

Civey

Forschungsgruppe Wahlen

Infras

Infratest dimap

Ludwig-Maximilians-Universität München (LMU)

Deutsche Statistische Gesellschaft

(This list is not exhaustive, but rather an indication of which kind of institutes the interviewees will come from.)

Is local approval required?

No

Approving body

Is approval:

Attached files

Research Proposal. Lukas Griessl.docx

consent-form.docx

participant-information-sheet.docx

08/06/2020

Mr Lukas Griessl

Sociology

University of Essex

Dear Lukas,

Ethics Committee Decision

I am writing to advise you that your research proposal entitled "Modelling the Unknown" has been reviewed by the Social Sciences Ethics Sub Committee.

The Committee is content to give a favourable ethical opinion of the research. I am pleased, therefore, to tell you that your application has been granted ethical approval by the Committee.

Best

Yasemin Soysal

Colchester Campus
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom

T 01206 873333

www.essex.ac.uk

@Uni_of_Essex   /uniofessex   /uniofessex

University of Essex

Participant Information Sheet

Project title: Modelling the Unknown

My name is Lukas Griessl and I am a Ph.D. candidate in the Sociology Department at the University of Essex. I would like to invite you to take part in a research study in form of an interview. Before you finally decide whether or not to take part in this study, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?

The study is part of my doctoral thesis in Sociology at the University of Essex. The research is interested in controversies surrounding statistical practice especially in regard to methodological decisions. The doctoral thesis develops around different case studies, which inform the overall analysis of the dissertation. The overarching interest of this research is on the topic of uncertainty and non-knowledge, which is explored in drawing on case studies in the field of statistics. The different case studies help to explore paradigms and perceptions underlying statistical controversies and are used as a means to get access to such controversies and to understand the perception regarding the choice, justification and implications of certain methods. The project is planned to last three years, of which about one year is planned for data gathering and analysis.

Why have I been invited to participate?

You have been invited to participate to this study, since you can speak for and represent your institute, which has a major stake in the controversy around one case study that this research wants to explore or can contribute to better understand the issue at stake due to your expertise. In total, a number of around 40 people, working in the same or other institutes and universities will be interviewed.

Do I have to take part?

The participation in this research is entirely voluntary. You can decide to not take part at any time, before, during or after the interview without the need to mention the reasons. Prior to the interview, you will be asked whether you agree to me recording or taking notes and will be explained that I will stop recording/ note taking at any point during the session when you ask me to. This will happen through a consent form, in which you will be presented with issues that are covered in this information sheet and which needs to be signed by you and me. If you decide to withdraw, you can contact me directly (Lukas Griessl / l.griessl@essex.ac.uk) and the interview won't take place or will be deleted if it was already conducted. If this is the case, the information that have already been provided will be removed and will not be used for the research. This does, however, not account for research that has already been published. In this sense, the data will be deleted from an already password protected folder on Box, the University of Essex cloud storage. This limits the possibility that confidential data does not reach anybody but me or the supervisory team.

What will happen to me if I take part?

If you take part in the study, an interview will be conducted via Skype or in person. The interview will last approximately one hour and will be recorded and transcribed. Additional interviews are not planned but may be conducted if needed and if agreed upon. The recording will be transcribed, to put it in a format appropriate for a subsequent analysis. The interview will be conducted by myself, as well as the transcription and the subsequent data analysis. The only person who has access to the material is me and the supervisory team. It is possible that another person will be involved in the transcription of the interview material. In this case, confidentiality and privacy will be maintained.

All files, including recordings, transcripts and personal information will be saved in different password protected folder on the University of Essex Box cloud storage. The audio files and the transcripts will be deleted 10 years after the research was conducted. Your personal and contact details will be deleted 3 years after the completion of the research. This is to ensure that I am able to have access during the research process and to go back to the material in the case that this is necessary after the research. As just indicated, however, you continue to be able to request a deletion before the expiration of the term. You will be provided with a full transcript and/ or the recording upon request.

Upon request, your name and affiliation will pseudonymised or anonymised in the dissertation. After study completion the data may be deposited with a data archive, such as the University of the Essex Data Archive to allow appropriate and controllable access to and sharing of the data, while maintaining privacy and confidentiality. This may include pseudonymised or anonymised interview transcripts, or context such as about the interview setting, the duration or other relevant information. This allows the data to be viewed or possibly re-used by other researchers but does not disclose personal or confidential information.

What are the possible disadvantages and risks of taking part?

Besides the time involved in the preparation and the execution of the interview, there are no other 'costs' involved in taking part in the study.

What are the possible benefits of taking part?

Taking part in this study does not guarantee any direct advantages or benefits. Taking part, however, will help to further the understanding of the topic and may shed light on issues regarding your work or your discipline that you may not have had otherwise.

What information will be collected?

Only information that you provide us with during the interview and those which are publicly available (e.g. on the website of your institute, publications etc.) will be collected and included in the research. The questions in the interview will be on topics such as perceptions of certain methods, working life, educational and professional background, as well as on the history and self-image of the institute and the discipline. Names and affiliations will be pseudonymized, but the exclusion of a possible identification cannot be guaranteed. As already indicated, an anonymisation of the data will take place upon request.

Will my information be kept confidential?

Your interview data will be held in accordance with the Data Protection Act. The information you provide will be used for research purposes and your personal data will be processed and stored in accordance with current data protection legislation. Any personal data you provide will be held in the strictest confidence and will not be disclosed to any unauthorised third parties. The results of the research will be submitted / published in a pseudonymised – or if requested anonymised form. Third parties will not be allowed access to interview recordings and transcripts. The only exception to this is if I am required to do so by law, or in the event that something disclosed during the interview causes concerns about possible harm to you or to someone else. As indicated before, data – including pseudonymised or anonymised transcripts - may be deposited with a data archive after completion of the research, while confidentiality will be maintained. Data thus may be accessible and re-used, but confidentiality will be secured.

Your personal and contact details will be stored separately from your interview transcript / digital recording and may be retained for up to 3 years after the completion of the research. The recordings and transcripts will be saved in a password protected folder on the University of Essex Box cloud storage and will only be used for this research and related research in the future. These data may be retained for up to 10 years after the completion of the research. If you request it, you will be supplied with a copy of the digital recording and your interview transcript. If, after review, you decide to withdraw, you can do so at any point. In this event, your interview data and contact details will be deleted immediately and will not be further used in my study. This does not, however, apply in retrospect to material that has already been published or presented in an academic setting.

What is the legal basis for using the data and who is the Data Controller?

The legal bases for the usage of data is the free consent of the participant, which is taken prior to the interview.

The Data Controller is the University of Essex and the contact is Sara Stock, University Information Assurance Manager (dpo@essex.ac.uk).

What will happen to the results of the research study?

Results of the study will be included in my doctoral dissertation, which will be submitted at the University of Essex. Results may also be published as academic publications, such as journal articles or be used in conference papers or presentations. Data may also be shared with a data archive after completion of the research. This means, the results will be made public in one way or another. In anyway, the confidentiality of personal information will be secured as outlined. Furthermore, if you want, a copy of the findings of this study will be made accessible for you.

Who is funding the research?

The research is funded by the Studienstiftung des deutschen Volkes (German Academic Scholarship Foundation) as part of its doctoral scholarship scheme. The foundation is politically and religiously independent and does not influence the study in any way.

Concerns and Complaints

If you have any concerns about any aspect of the study or you have a complaint, in the first instance please contact the principal investigator of the project, Lukas Griessl, using the contact details below. If you are still concerned or think your complaint has not been addressed to your satisfaction or you feel that you cannot approach the principal investigator, please contact the primary supervisor of this doctoral dissertation, Professor Linsey McGoey (lmcgoey@essex.ac.uk) or the University's Research Governance and Planning Manager, Sarah Manning-Press (e-mail sarahm@essex.ac.uk). Please include the ERAMS reference which can be found at the foot of this page.

Name of the Researcher/Research Team Members

Principal Investigator: Lukas Griessl; Ph.D. student, Department of Sociology; l.griessl@essex.ac.uk

First Supervisor: Linsey McGoey; Professor at the Department of Sociology; lmcgoey@essex.ac.uk

Second Supervisor: Nick Allum; Professor at the Department of Sociology; nallum@essex.ac.uk

Consent Form

Title of the Project: Modelling the Unknown

Researcher: Lukas Griessl

Please initial box

1. I confirm that I have read and understand the Information Sheet dated 18.09.2020 for the above study. I have had an opportunity to consider the information, ask questions and have had these questions answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw from the project at any time without giving any reason and without penalty. I understand that any data collected up to the point of my withdrawal e.g. will be destroyed.

3. I understand that the identifiable data provided will be securely stored and accessible only to the principal investigator and the supervisory team of the research, and that confidentiality will be maintained.

4. I understand that data collected during the interview will be used for the doctoral dissertation, for possible academic publication or conference presentations and that it may be made accessible through a data archive. I also understand that the data will be pseudonymised or anonymised upon my request.

5. I understand that the data collected about me may be used to support other related research in the future.

6. I agree to take part in the above study.

Participant Name   Date   Participant Signature

_____   _____   _____

Researcher Name   Date   Researcher Signature

_____   _____   _____