

Abductive Coding: Theory Building and Qualitative (Re)Analysis

Sociological Methods & Research

1–34

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00491241211067508

journals.sagepub.com/home/smr



Luis Vila-Henninger ^{1,3},
Claire Dupuy¹,
Virginie Van Ingelgom²,
Mauro Caprioli,
Ferdinand Teuber¹,
Damien Pennetreau²,
Margherita Bussi¹, and
Cal Le Gall^{1,4}

Abstract

Qualitative secondary analysis has generated heated debate regarding the epistemology of qualitative research. We argue that shifting to an abductive approach provides a fruitful avenue for qualitative secondary analysts who are oriented towards theory-building. However, the concrete implementation of abduction remains underdeveloped—especially for coding. We address this key gap by outlining a set of tactics for abductive analysis that

¹University of Louvain, ISPOLE

²F.R.S. – FNRS, University of Louvain, ISPOLE

³Aarhus University

⁴University of Vienna

This article reflects only the authors' views and the European Research Council is not responsible for any use that may be of the information it contains.

Corresponding Author:

Vila-Henninger Luis, Aarhus University, Department of Political Science; University of Louvain, ISPOLE.

Email: luis.vila@ps.au.dk

can be applied for qualitative analysis. Our approach applies Timmermans and Tavory's (Timmermans and Tavory 2012; Tavory and Timmermans 2014) three stages of abduction in three steps for qualitative (secondary) analysis: Generating an Abductive Codebook, Abductive Data Reduction through Code Equations, and In-Depth Abductive Qualitative Analysis. A key contribution of our article is the development of "code equations"—defined as the combination of codes to operationalize phenomena that span individual codes. Code equations are an important resource for abduction and other qualitative approaches that leverage qualitative data to build theory.

Keywords

Qualitative, Abduction, Secondary Qualitative Analysis, Coding, Team Research

Introduction

In the analysis of qualitative data, many researchers anchor their methodological approach—explicitly or not—in grounded theory as developed by Glaser and Strauss (1967) and elaborated further by Glaser (1992), Strauss (1987), and Charmaz (2000). While grounded theory offers a useful set of steps for conducting and analyzing qualitative data and has generated rich and insightful bodies of work, Deterding and Waters (2018) argue that grounded theory's "central prescription—theoretical sampling toward saturation, strongly inductive analysis, and full immersion in the research field—bear little resemblance to the actual methods used by many large-scale interview researchers" (Deterding and Waters 2018, 2).

From this observation arises the issue of what avenues are to be taken to analyze qualitative data when scholars' research practices do not match the central prescriptions of grounded theory. In this article, we take part in this ongoing conversation by focusing on qualitative secondary analysis of several datasets, one of many research practices that are misaligned with grounded theory. Qualitative secondary analysis entails the use of already produced qualitative data – interviews (face-to-face or focus groups), field and observations' notes – to develop new social scientific and/or methodological understandings (Irwin 2013).

While being developed as a new methodological approach for more than two decades (Bishop and Neale, 2011; Corti, 2006, 2012; Corti, Witzel and

Bishop 2005; Fiedling, 2004; Beck 2018; Hughes and Tarrant 2020; Duchesne 2017), qualitative secondary analysis has generated a heated debate related to the epistemology of qualitative research (Hammersley, 1997, 2010; Mauthner et al. 1998; Savage 2005; Duchesne 2017). Opponents have extensively discussed how qualitative secondary analysis contradicts the prescriptions of an inductive epistemology of qualitative research. In this article, rather than giving up on qualitative secondary analysis, we argue that shifting to an abductive approach (e.g. Timmermans and Tavory 2012; Tavory and Timmermans 2014) provides a fruitful avenue for qualitative secondary analysts who are oriented toward theory-building. Abduction thereby provides both an epistemology and a method of analysis that is adapted to some of the contemporary research practices of qualitative scholars. However, in the burgeoning literature on abduction, the concrete implementation of an abductive logic of inference remains underdeveloped (for a recent contribution, see Brandt and Timmermans 2021)—and especially for coding. In this article, we address a key gap in the abduction literature by outlining a set of tactics for abductive analysis that can be applied to secondary as well as to primary data analysis.

Our approach applies Timmermans and Tavory's (Timmermans and Tavory 2012; Tavory and Timmermans 2014) three stages of abduction—Revisiting the Phenomenon, Defamiliarization, and Alternative Casing—in three steps for qualitative (secondary) analysis: Generating an Abductive Codebook, Abductive Data Reduction through Code Equations, and In-Depth Abductive Qualitative Analysis. Concretely, researchers start with a deductive codebook and through the process of coding, build the codebook—and by extension build theory—by developing data-driven inductive codes that document theoretically anomalous cases (step 1). Our strategy uses Qualitative Data Analysis (QDA) software to combine codes to operationalize complex phenomena that span codes in the abductive codebook. We refer to this as a “code equation” (step 2). This step is crucial for data reduction and allows for further inductive coding that facilitates a final round of qualitative analysis (step 3).

We discuss the generally applicable features of each step, as well as their respective payoff in terms of the overarching abductive approach. Being explicitly directed towards theory-building, we argue that each of these coding tactics enriches the analysis and cultivates the conditions for generating creative and novel theoretical insights. In that respect, our intent is not to infringe upon the creative process of theory-building. Rather, it is to outline a set of strategies that foster the principles set out by Timmermans and Tavory (Timmermans and Tavory 2012; Tavory and Timmermans 2014) and can

help to facilitate this creative process—as well as to facilitate other scholars' development of similar applied tactics that can enrich abductive analysis and theory-building in ways to go beyond the scope of our article. We also echo Timmermans and Tavory's application of Peirce's (1934) philosophy to qualitative analysis as a strategy for building theory. In that regard, a key contribution of our article is the development of "code equations". We argue that by combining codes, qualitative researchers can operationalize increasingly complex phenomena and answer increasingly complex research questions. Code equations are then a key resource for abduction and other qualitative approaches that leverage qualitative data to build theory.

While we illustrate the three steps of abductive analysis using our experiences with abduction in ATLAS.ti, the tactics we develop in the article can be applied with any QDA software on primary or secondary data. Drawing from our research, we use existing qualitative data on citizens' perceptions of democracy from three qualitative data sets that consist of interviews and focus groups collected by different researchers from the mid-1990s until 2017 in the United Kingdom¹.

We proceed as follows: After setting up the terms of the methodological debates surrounding the secondary qualitative analysis, we present the abduction literature as a strategy for tackling some of these issues and highlight how our approach responds to the gap in this scholarship concerning tactics for abductive coding. In the next section, we discuss our framework for coding as it applies these steps, outline our corresponding tactics, and provide examples from our own research. We conclude by discussing the implications of our method, as well as its limitations beyond the illustrative case of secondary analysis that we present here.

Secondary Qualitative Analysis: From Induction to Abduction

In this section, we show that based on an abductive approach, the key features of secondary qualitative analysis can be leveraged to support theory building through abductive coding. Thus, the perceived limits of secondary qualitative analysis that ground our discussion—and/or primary qualitative analysis that challenges the prerequisites of grounded theory and induction—constitute an opportunity for theory construction and abductive analysis. Building on Timmermans and Tavory's three steps of an abductive approach, we then discuss how we apply them to build abductive coding of qualitative secondary data and, thereby, fill the related gap in the scholarship on abduction.

Resistance to Qualitative Secondary Analysis from an Inductive Epistemological Perspective

In the last decades, qualitative repositories such as the Henry A. Murray Research Centre (Harvard), the Qualitative Data Repository (Syracuse), the Australian Data Archive, the UK Data Archive, or BeQuali (Sciences Po Paris) have flourished.² These data repositories offer new opportunities for researchers to reanalyze qualitative data sets and thus perform qualitative secondary analysis (Duchesne 2017; Hughes and Tarrant 2020). Secondary analysis is understood as “a research strategy which makes use of pre-existing quantitative or pre-existing qualitative research data to investigate new questions or verify previous studies” (Heaton 2004, 16). In sum, qualitative secondary analysis entails the use of what Salganik (2018) calls “ready-made data” in reference to Marcel Duchamp’s art in his discussion of big data in social sciences.

The development of qualitative secondary analysis has generated large debates on the new research possibilities it may offer (Hughes and Tarrant 2020). Specifically, three arguments have been put forward to oppose qualitative secondary analysis. Opponents insist that a relevant analysis of qualitative data is conditional upon the participation of the analyst in data collection (Mauthner et al. 1998; Mauthner and Parry 2009, 2013; Parry and Mauthner 2005). In their landmark article, Mauthner et al. (1998), highlight that while researchers’ role in interpreting and theorizing the data is crucial, so is their role in data construction (ibid 1998: 733). They argue that there is a fundamental epistemological issue with qualitative secondary analysis relating to the reflexive and interpretive nature of the research paradigm within which qualitative researchers reside (ibid 1998). Failing to acknowledge this epistemological issue would result in researchers unwittingly adopting a “naively realist” position. Relatedly, another issue of qualitative secondary analysis pertains to whether it is possible to analyze data outside of the original context in which it was collected (Moore 2007; Heaton 2004). Parry and Mauthner (2005) consider that a key distinctive feature of qualitative data analysis in comparison to quantitative data is the paramount importance of the context of production. Because secondary analysts who were not involved in the data collection are likely not to be privy to primary researchers’ experience (Beck 2018: 13), Parry and Mauthner doubt that qualitative data can be reused to generate new substantive findings and theories in an equivalent way to quantitative data (see also Thorne 2013). Moreover, as qualitative secondary analysis requires assembling existing qualitative data sets into a new corpus, critics point that this reconfiguration of

data challenges primary research's theoretical sampling, thereby bringing back the usual questions about representativeness and generalizability of qualitative research. Some argue that secondary qualitative analysis could exacerbate the representational problems of the primary research (Beck 2018: 14).

These opposing arguments to qualitative secondary analysis raise fair concerns that should certainly be considered with the necessary reflexivity by secondary analysts but which we do not have space to address directly. Instead, we want to emphasize that they reflect and are embedded in a specific epistemological perspective which, essentially, is inductive and closely echoes grounded theory's "central prescriptions" (see Deterding and Waters 2018, 2). Arguments about the primacy of researchers' experiences with interviewees and their close knowledge of the context of data production during data analysis relates to grounded theory's idea that scholars should be fully immersed in their research fields and that data analysis should be "strongly inductive" (Deterding and Waters 2018, 2). This prescription runs contrary to practices of secondary analysis, where the amount of data is typically larger than in primary research as secondary analyses respond to research questions that usually extend beyond one sample and rely on the use of multiple data sources. Thus, a main issue is data reduction more so than immersion in the research field.

More importantly, induction is practically contradicted in any secondary analysis. To become familiar with the secondary data they reanalyze, researchers have to be aware of—and sensitized to—the research questions, theories, and concepts that drove primary researchers. Similarly, the argument about how theoretical sampling reflects a conceptualization where saturation is the main objective. In qualitative secondary analysis, sampling is theoretical but does not aim necessarily for saturation. Instead, comparison is key and sampling rather relies on the principle of diversity. Scholars' research questions drive which data sets they select when creating a secondary qualitative data corpus—thus creating opportunities for comparative secondary research (Halford and Savage 2017). Sampling data based on the goal of diversity, thereby, is instrumental in gaining theoretical traction more so than saturation.

Because the heated debate on the epistemology of qualitative research generated by qualitative secondary analysis has mainly remained anchored in an inductive epistemology, the potential of secondary qualitative methodology has been poorly evaluated. However, acknowledging that qualitative secondary analysis does not match the requirements of an inductive approach of data analysis does not mean that this type of analysis is a dead-end. Rather, we suggest that a shift to an abductive approach is an avenue forward as it

allows us to leverage the defining features of qualitative secondary analysis in order to build theory.

Abduction as an Avenue Forward

The application of abduction to qualitative research as a technique for theory-building (Timmermans and Tavory 2012; Tavory and Timmermans 2014) arose in the context of intense debate over the use of induction versus deduction (Jerolmack and Khan 2017). Glaser and Strauss (1967) have been hugely influential in advancing inductive approaches, while scholars such as Michael Burawoy (1998, 2009) have shaped the discourse around deductive analysis in qualitative methods via the Extended Case Method. In this context, Timmermans and Tavory (Timmermans and Tavory 2012; Tavory and Timmermans 2014; Tavory 2016; Tavory and Timmermans 2019) advanced the strategy of “abduction” to move beyond the induction versus deduction debate. Abduction is then part of a movement that highlights what its proponents see as a false dichotomy between inductive and deductive analysis (Goldberg 2015; Wagner-Pacifici, Mohr, and Breiger 2015).³

Abduction as a method of data analysis was initially developed by Peirce (1934) as a way to draw inferences that are oriented towards theory-building. It refers to the iterative process between theoretically surprising cases⁴ and tentative explanations (Coffey and Atkinson 1996; Kelle 2007; Reichertz 2007; Strübing 2007; Timmermans and Tavory 2012; Tavory and Timmermans 2014). As such, abduction is distinct both from deduction and induction but combines features of both types of inference. What sets abduction apart from a purely (ideal-typical) inductive form of inference is that the observed phenomenon does not contain an explanation in itself (induction), neither does it constitute a new case of an already known general rule (deduction), but is rather a combination of both.

From theoretically surprising empirical cases, “[a]bduction is the process of forming an explanatory hypothesis” (Peirce 1934, 171). Once these cases are “discovered,” researchers then combine multiple explanations to formulate a theoretically sound account of it (Reichertz 2007, 159–164). Subsequently, researchers test these inductive, theoretically-grounded explanations against existing and additional empirical data, which may result in another round of theory building. For Peirce, abduction, induction, and deduction are thus not exclusive inferences but act at different stages of the research process—which is iterative in nature.

In particular, abduction starts with a set of theories and extends them by looking for theoretically anomalous empirical cases. Empirical observations

are anomalous, novel, or surprising only based on what is already theoretically established or what is expected based on existing theories—which therefore serve as a benchmark to identify unexpected empirical observations. Existing theories are also an essential part of forming new hypotheses—as the latter stem from the mismatch between empirical observations and the former. On a theoretical level, new hypotheses can be generated only when all other possible theoretical explanations have been exhausted and proved unable to account for the observations at hand. In an abductive analysis, theoretical breadth is thus necessary for theory-building (Timmermans and Tavory 2012; Tavory and Timmermans 2014).

Abduction has been largely influential in the past five years in qualitative sociological research—as is evidenced by its application by European (Vassenden and Jonvik 2019) and American (Carlson 2017, 2018, 2019; Timmermans 2017; Reilly 2018; Blee 2019; Vila-Henninger 2020) scholars. Abduction has also been embraced by computational methods. For example, Ignatow (2020) develops strategies for “digital abduction,” while Karell and Freedman (2019) develop “computational abduction.” Some analysts have even performed what could be seen as secondary qualitative analysis using abduction (Blee 2019) whereas Brandt and Timmermans (2021) propose an abductive logic of scientific inference for quantitative research.

Although some qualitative researchers have hinted that abduction could be useful for coding qualitative materials (Deterding and Waters 2018), there is however an important gap in the literature concerning how abduction and abductive coding can be used concretely by researchers as a strategy for theory-building. Carleheden (2016), who is critical of abduction, calls for an elaboration of how abduction is to be carried out. Despite this call to action, scholars have yet to outline strategies that can be used across researchers to guide abductive coding and analysis. For example, Blee (2019) and Rinehart (2021) both provide personal insights into using abduction—which by nature are useful but do not take the additional step to leverage these insights into a set of tactics that cultivate the abductive logic of theory discovery as innovative and creative.

Abductive Analysis and Qualitative Secondary Analysis

We start from Timmermans and Tavory’s three steps of an abductive approach—Revisiting the Phenomenon, Defamiliarization, and Alternative Casing—and discuss how we apply them to build abductive coding of qualitative secondary data. In parallel, we build on our own experience of

abductive secondary analysis in order to outline a set of tactics that facilitate abductive coding and cultivate these three steps.

“Revisiting the Phenomenon” is the first of these steps. Abduction depends on the ability to consider empirical observations in light of different theories—or to “revisit” them (Timmermans and Tavory 2012). Through revisiting the data, different aspects may become salient, leading to an empirically-based revision of the theory. This revisit then helps “to sensitize different theoretical approaches” (ibid: 176). Thus, Timmermans and Tavory submit that revisiting a set of empirical observations “is a way to harness temporality in the service of theory construction” (ibid: 176) in the sense that over time, initial descriptions and analyses evolve and additional theoretical lenses are incorporated. This process then brings to light empirical features of the data that may have initially been overlooked by the analyst(s). In this way, researchers use extant theory to discover theoretically anomalous or surprising cases in the data through their revisit.

In secondary qualitative analysis (Davidson et al. 2019), the revisit is a defining feature. However, because of the amount of data that could be mobilized, the sensitization to different theoretical approaches operates under sensibly distinct conditions. In *Abductive Coding*, we contend, the revisit has to be supported by the initial steps of data analysis. The initial codebook must be theoretically broad so that multiple theoretical perspectives are built-in (*see* step 1.a). Such a theoretically-encompassing codebook is instrumental in looking at the data from multiple theoretical lenses. Yet “Revisiting the Phenomenon” is not limited to the initial steps of coding—as it also occurs along with data reduction. We propose that data reduction takes the form of code equations (*see* step 2.a). These theory-driven combinations of codes allow for specific operationalizations of each theory—and the deductive test is conducted through the application of these code combinations to the interview data. It is through this revisit, in the form of abductive coding, that anomalous cases can be found that then help to extend the researcher’s broad initial theoretical framework (*see* step 3.a and step 3.b).

According to Timmermans and Tavory (Timmermans and Tavory 2012; Tavory and Timmermans 2014), defamiliarization of the data is another methodological stage that supports an abductive logic of data analysis. The idea is to question the taken for granted to see the data differently. Here also, time can work as a support of abductive analysis as the loss of immediacy with data collection results in defamiliarization with the data. This process of defamiliarization then facilitates the discovery of anomalous cases.

In the reanalysis of qualitative data, as in medium and large-N qualitative studies, it is unlikely that the researcher(s) initially have an extended

familiarity with the data on which they are working. In this context, it may be that it is through familiarization with the corpus, and the studies it includes, that new or unexpected empirical observations emerge. We argue that familiarization with the data occurs through deductive coding—and subsequent inductive coding allows for identifying theoretically anomalous observations (see step 1.b) and thus for defamiliarization. As highlighted by Deterding and Waters (2018), coding is an integral part of (de)familiarization with the data. However, familiarization with the data—which in secondary qualitative analysis takes the place of “Defamiliarization” for the primary individual researcher—also occurs after data reduction when researchers verify code equations to ensure that the phenomenon they aim to operationalize is present (see step 2.b). When this is not the case, it may be that the code equation is incomplete or partly only adjusted to the given phenomenon—which subsequently triggers a further reconsideration of the phenomenon and of its conceptualization. Furthermore, (de)familiarization happens in the final step of abductive data analysis (see step 3.a) when reduced cases are inductively recoded to account for the nuances of the phenomenon at stake. These reduced cases call researchers to look at data differently and operate as a trigger for theory-building. In that respect, (de)familiarization not only serves to discover the anomalous cases in the data but also to generate new theoretical insights to account for them.

Alternative casing is the third methodological stage (Timmermans and Tavory 2012; Tavory and Timmermans 2014). For Timmermans and Tavory, “[e]ach casing abstracts and highlights different aspects of the phenomenon, rendering it comparable to different phenomena and turning it into a generalization that then can be linked to other fields and theories” (ibid 2012: 177). In that sense, “the theoretical background is foregrounded as a way to set up empirical puzzles” (ibid: 177).

Alternative casing is then an in-depth case study that applies a theory, or set of theories, in great detail. Alternative casing is an inherent part of the process of analysis—especially when working in a team, as collaborators are likely to study different aspects of the main research question and thereby draw from different bodies of scholarship. In that sense, teamwork as such supports alternative casing. Also, alternative casing is supported through the step of data reduction when code equations are built based on existing theories to study the empirical data (see step 2.a). Finally, alternative casing takes place in extended qualitative analysis (see steps 3.a and 3.b). Thus, in our approach, “Alternative Casing” occurs through qualitative analysis, in which researchers apply different theoretical frameworks at length to key empirical cases.

Unpacking Abductive Analysis in Coding

To respond to the gap in the literature and help facilitate abductive theory-building, we outline in this section a set of tactics for abductive coding of secondary and primary data. This approach combines traditional approaches to abduction (Peirce 1934) with qualitative methodological strategies (Timmermans and Tavory 2012; Tavory and Timmermans 2014). We also build upon computational applications of abduction (Karell and Freedman 2019; Ignatow 2020), the “Breadth-and-Depth Method” (Davidson et al. 2019) for qualitative secondary analysis⁵ and the “Flexible Coding” approach (Deterding and Waters 2018). Furthermore, our approach develops out of—and contributes to—the insights provided by qualitative, non-computational applications of abduction (Carlson 2017a, 2017b, 2019; Timmermans 2017; Reilly 2018; Blee 2019; Vassenden and Jonvik 2019; Vila-Henninger 2020). Finally, our approach stems out of what Tavory and Timmermans (Tavory and Timmermans 2014; Fine and Tavory 2019) call a “community of inquiry,” in that we discuss how our research team collectively applied abduction in coding and analysis.

Our steps are as follows: 1) Generating an Abductive Codebook, 2) Abductive Data Reduction through Code Equations, and 3) In-Depth Abductive Qualitative Analysis. It is through these three steps that one applies abduction to build theory using our approach (*see* Table 1 below).

Abduction consists of key moments of discovery—both small moments in the process of analysis and larger moments that synthesize these smaller moments (e.g. Tavory and Timmermans 2014). In our process, Step 1b (*Generating inductive codes through group coding*) and Step 2a (*Text reduction through code equations*) are phases in which the analyst creatively engages with the data in “little” abductive movements. Step 1b consists of

Table 1. Stages of Abduction and Steps in Abductive Coding.

	Step 1— Generating an Abductive Codebook	Step 2—Abductive Data Reduction through Code Equations	Step 3—In-Depth Abductive Qualitative Analysis
Revisiting the Phenomenon	1.a	2.a	3.a and 3.b
Defamiliarization	1.b	2.b	3.a
Alternative Casing	NA	2.a	3.a and 3.b

smaller abductive discoveries relative to coding that then enable Step 2a's smaller discoveries in the development of the code equation and reduction of data. It is then in completing step 2b (*Code equation verification*) and moving to Step 3 (*In-Depth Abductive Qualitative Analysis*) that the analyst leverages these smaller movements to make larger discoveries. Thus, these three steps help to organize the overall logic and progression of abductive analysis. By moving through these three steps, analysts engage in the broader abductive discovery process and conclude (Step 2b and 3) by potentially arriving at larger movements of discovery of new theoretical insights based on surprising and anomalous research results. We therefore see each step and sub-step as cumulative in the overall process of an abductive logic of inquiry.

In what follows, we provide examples of each step of our abductive approach in our team research. We developed this approach throughout our analysis—which seeks to advance the understanding of the interaction between policy and politics. More precisely, we investigate whether certain observable trends in public policy (supranationalization and neoliberalism) influence citizens' democratic linkages—defined as “the various ways in which citizens are connected in a structural and durable way to their political system” (Kitschelt and Wilkinson 2007). The theoretical argument at the center of the project is “policy feedback” (e.g. Pierson 1993; Campbell 2003; Mettler 2002; Bussi et al., 2022). For this article, we will focus only on two dimensions of the project—democratic linkages and neoliberalism—and on three qualitative data sets (*see* Appendix for a detailed description of each data set): Céline Belot's semi-structured interviews in the UK collected from 1995–6 (Belot 2000), the 2005–2006 UK CITAE focus groups (Duchesne et al. 2013), and the semi-structured interviews from the Qualitative Election Study of Britain collected from 2016–7.

Step 1. Generating an Abductive Codebook

Our first methodological step is the creation and development of the abductive codebook. In our case, we divided this coding among team members. Generating an abductive codebook can be divided into two sub-steps: a) using an initial deductive, theoretically-based broad codebook and b) subsequently generating inductive codes through group coding. Through these two sub-steps, one creates an abductive equivalent to indexing and creating memos (Deterding and Waters 2018). However, the indexation cannot follow the interview grid as the primary researchers' goals are different from the research questions developed in the secondary qualitative analysis.

More generally, we can see these two sub-steps as an applied strategy for abductive coding and as a first tactic to cultivate innovative findings by taking advantage of the analytical potential of abduction. Step 1 is where the researcher engages with the data as an expert coder and generates the first round of discoveries via coding—and more specifically through revisit and (de)familiarization. By applying this strategy, researchers can engage in abduction through their own creativity as it best applies to their research, but do so efficiently by already having an applied strategy for how to approach abductive coding as a way to develop new theoretical insights based on surprising and anomalous results in light of existing theories (Brandt and Timmermans 2021).

Step 1.a. Using an initial deductive, theoretically-based broad codebook. The first sub-step is the generation of the initial deductive, theoretically broad codebook.⁶ It should be based on the theory and research used to build the research questions. The key, then, is to design the deductive codebook so that researchers can code for excerpts of text that allow them to address the research question empirically. This sub-step then entails the first round of coding of a smaller—yet strategically selected in particular for diversity purposes—number cases within the data set to generate indicative codes.

In our project, subgroups of researchers were tasked with coding sections of the data set in a QDA software. Our initial deductive codebook consisted of two sections: thematic and descriptive. Building on sociological knowledge, we established a descriptive codebook of factors that would be important to our analysis in particular for socio-demographic factors (e.g. age, education) likely to structure citizens' perceptions and reactions to politics and policies. This descriptive coding corresponds to attributes as highlighted by Deterding and Waters (2018, 17–18). Descriptive codes are not limited to individual characteristics as other levels of analysis may be pertinent for the analysis such as the period of time or the location where the data were collected. In parallel with building and applying this descriptive coding, we investigated the relevant literatures regarding the four main dimensions of the project: policies, democratic linkages, supranationalization, and neoliberalism. Broadly speaking, each dimension constituted the main research focus of one or more of the team members. By adding these dimensions in a single common codebook, the analysis gains theoretical breadth and fosters both the revisit of the phenomena and defamiliarization which are central to the abductive logic of inference. Following our qualitative secondary analysis, we needed to operationalize those concepts in a way that would be neither

country- nor time-specific. Our operationalization also had to travel across datasets and, thereby, not to be restricted to one of the datasets.

A major aim of our research is to contribute to our understanding of “Democratic Linkages.” Initially, the “democratic linkage” codebook was developed in an explicitly deductive manner that was intentionally broad enough to allow for inductive codes to emerge—but at the same time served to limit the scope of inductive coding. Our deductive codes for Democratic Linkages were based on Easton’s paradigmatic framework (Easton 1965). This decision was informed by our reading of the historical trajectory of post-war studies on the relationship between citizens and democracy (Dalton 2007; Hay 2007; Norris 2011). Many surveys contain items that were inspired by Easton’s theories and that were kept with minimal changes to ensure longitudinal comparison in the study of political support (Linde and Ekman 2003, 1; see also Torcal and Moncagatta 2011). We thus created five deductive Democratic Linkage codes: Political Community, Regime Principles, Regime Institutions, Political Actors, and Citizen Participation (*see* Appendix for a table of all of our initial deductive codes).

Step 1.b. Generating inductive codes through group coding. In our theory-building approach, through an initial round of coding, the researcher(s) generate(s) inductive codes—thus identifying and documenting anomalous or surprising cases. If analysts are working as a team, meeting throughout this process to discuss the progress of the inductive codes and update the codebook is helpful. After the researcher(s) determine(s) that inductive coding has reached saturation, the researcher(s) update(s) the codebook—thus creating a final abductive codebook. The analyst(s) can then proceed to code the entire corpus—which is often a compilation of data sets—with the full abductive codebook. During this process, the researcher(s) recode(s) data that was initially coded with the initial deductive codebook using the subsequent abductive codebook.

In our project, the first element of induction came about when testing and further developing the democratic linkages deductive codebook. Initially, we assumed the distinction between formal political actors (e.g. politicians) and other actors not formally endowed with political power (e.g. pressure groups) as potentially relevant to explaining citizens’ changing relationship with the political system. This was deduced from the literature on the suggested double shift in which traditional political actors seem to lose importance while “new” or other actors gain a form of (depoliticized) political power—e.g. technocrats and economists during the 2008 Eurocrisis (Hay 2004,

2007; Harmes 2012; Stoker and Hay 2017). However, the distinction between these two types of politically relevant actors did not seem to offer a relevant understanding of the changing nature of citizens' relationship to democracy. What we did find in the data were elements reminiscent of what can be found in the literature on populism (Kaltwasser 2017)—namely, a recurrent skepticism towards “the elite(s)” and an emphasis on democracy as “rule by the people” in which participation is considered primarily as a way for the people to express their will—e.g. by voting or demonstrating.

In our view, both grounded theory coding schemes and machine learning algorithms might have serious issues identifying anomalies in the data: the former for epistemological reasons as without preexisting theories, empirical observations cannot be deemed surprising; the latter for practical reasons. Anomalous cases are by definition unforeseen and their detection often warrants substantial expert knowledge. For example, we often coded entire sentences, paragraphs, or passages that did not use an obvious “keyword” to be expected in a given context. Furthermore, it is in the nature of secondary data analysis that the corpora to be analyzed are made up of independently collected and diverse data. Diversity comes, for example, in form of different languages and types of data, e.g. focus groups, semi-directive interviews, etc. While machine learning algorithms are increasingly able to handle some of these issues, they may be less suited to detect anomalous theoretical cases. However, in principle, we believe that the use of machine learning methods for the analysis of the remaining corpus need not be incompatible with an abductive logic. For instance, one might consider training a supervised machine learning algorithm based on an established abductive codebook and evaluate its validity via the application of a “test set” (Grimmer and Stewart 2013). We agree with Quinn et al. that “each method for analyzing textual content imposes its own particular set of assumptions and, as a result, has particular advantages and weaknesses for any given question or set of texts” (Quinn et al. 2010: 210). As always, the decision should be based on the specific research question—as well as in the case of textual data, the size and heterogeneity of the corpus and the complexity of the searched categories. We acknowledge that our approach may necessitate comparatively high human “costs” in particular in the coding phase itself (Quinn et al. 2010) and thereby sets limits to the amount of data analyzed.

Step 2. Abductive Data Reduction Through Code Equations

For our approach, abductive data reduction consists of two sub-steps: a) Text-reduction through code equations and b) Code equation verification.

There is a key divide between Steps 2a and 2b. Through the researcher's engagement with codes, which then facilitates further engagement with the data, the researcher creates an initial code equation in Step 2a. It is in Step 2b that the researcher uses the code equation to reengage with the data and make big abductive discoveries concerning a "final" code equation. This equation is made possible by the little discoveries made in Step 1 from coding. By using the code equation, the researcher can tie together codes in a way that allows her to understand a complex phenomenon. Additionally, code equations offer insight into how a phenomenon's components relate to one another. This step is analogous to the dimensionality reduction that many methodological approaches employ—for example when topic models reduce the many dimensions of a document formed by its words to document-topics (e.g. DiMaggio et al. 2013). Many other approaches reduce dimensionality by showing that components tend to "hang together", and thereby lead researchers to new conceptualizations of phenomena.

Step 2.a. Text reduction through code equations. Rather than performing a preliminary analysis of a set number of words surrounding a keyword (Davidson et al. 2019, 372), we propose an additional useful approach in which the analyst(s) proceed(s) by investigating passages that they coded using the full abductive codebook. The task at hand, then, is to reduce data that have already been qualitatively coded using a QDA software. For this process, we suggest using QDA software tools that allow for the operationalization of phenomena that span codes and allow (the) researcher(s) to address the secondary analysis research question(s) empirically.

We define "code equations" as the combination of codes to operationalize phenomena that span individual codes. Creating a code equation is important because it allows for two key processes. First, one code in a data set may include hundreds or thousands of excerpts. Creating code equations then allows for more precise operationalizations that reduce text through the application of a theory-driven combination of codes. Second, a key feature of secondary qualitative analysis is its potential to investigate a research question that spans individual data sets. Thus, as researchers are driven to compile data sets to investigate increasingly complex questions, code equations allow for corresponding increasingly complex operationalizations.

We proceed with our discussion and demonstration using ATLAS.ti because it is the QDA software we used in our research. A key element of our process was our use of ATLAS.ti's "Query Tool"—which is

similar in many ways to running a “Coding Query” in NVivo. Code equations can be created using functions such as “AND” or “OR” or “COOC” (i.e. co-occur of two or more codes) to combine codes using the ATLAS.ti Query Tool. Researchers can create deductive code equations and then investigate if these operationalizations appear in the data. This process then allows us to reduce vast amounts of text to excerpts that fit our code equation. If the code equation does not succeed in reducing data, then this is a sign that the researcher might need to adjust the code equation (*see* step 2.b).

This is useful in two cases that we argue can be shared by all qualitative researchers. First, this is useful for scholars whose research question focuses on a phenomenon that has been untheorized or under-theorized and thus there is insufficient information to create specific deductive codes. For example, one of the team’s first analyses was generated by the following research question stemming from the literature on Euroscepticism in the UK: *How do UK citizens combine neoliberal ideological reasoning in support of market competition with nationalist ideological reasoning to oppose EU membership?* At the heart of this question is the issue of UK citizens’ specific strand of Euroscepticism, that is, citizens’ rejection of the EU that combines nationalism (the national political community) and neoliberalism (fair market competition). Thus, we wanted to investigate a potential intersection of neoliberalism and nationalism. Such an intersection was largely untheorized. However, through code equations that combined different neoliberal and nationalist codes, we were able to engage in a process of abductive discovery by leveraging code equations that combined these two codes. Second, code equations are useful for codes that are more specific than those from the original codebook. For example, we only wanted to look at the intersection of neoliberalism and nationalism as applied to the UK by Brits. By adding a deductive code to the code equation that coded for the theme of the UK, we were able to use a code equation to reduce the data to focus only on cases in which respondents used neoliberal and nationalist ideology in conjunction with the context of discussing the UK.

More specifically, we began by theorizing that we would find both input and output legitimacy in the nationalist neoliberal legitimations that we were interested in through our research question. The initial operationalization resulted in our construction of two equations for input and output legitimacy (Scharpf 1997, 1999) that mobilize some of our codes:

1) *Input legitimacy*: National Level (UK) COOC Regime Institutions COOC Neo-Liberal Fair Market Competition COOC Fairness Judgement

Application: the UK national level appears together with regime institutions, fair market competition, and judgments of fairness

2) *Output legitimacy*: EU Level COOC Regime Institutions COOC Political Community COOC Unfairness Judgment.

Application: the EU level appears together with regime institutions, fair market competition, and judgments of unfairness

See *Appendix* for the explanation of each code.

The way our research leveraged code equations was to combine theoretically relevant codes. Thus, this approach would have been impossible using more interpretative approaches (Glaser and Strauss 1967; Nelson 2017), as our code equations combined identified codes of theoretical relevance. The abductive codes we used were created by researchers through interpretative analysis—which is then impossible for coding via machine learning algorithms. In comparison to Karell and Freedman’s approach (2019), who perform a “recursive movement between the computational results, a close reading of selected corpus material, further literature on social movements, and additional theories that engendered a novel conceptualization of radical rhetorics” (ibid: 730), we engaged in a recursive abductive coding approach that was driven by deductive and inductive codes. Our use of codes that were identified as theoretical anomalies by team members gives this approach a certain type of relevance for our research questions that inductive and machine learning approaches cannot provide.

Finally, it is important to note that this step performs a key task that is similar to Davidson et al.’s (2019) second step (“Recursive surface ‘thematic’ mapping”). However, rather than data reduction through keywords, our sub-step reduces data through code equations.

Step 2.b. Code equation verification. Once we used the Query Tool to create a code equation, we analyzed excerpts of the text specified by the code equation. To proceed, the analyst must revisit each excerpt to verify that the phenomenon that is supposed to be captured through the code equation operationalization can in fact be identified in the data excerpt. This then creates a process of identifying false positives. By “false positive” we mean a data excerpt that has been coded with all of the codes specified in

the code equation but in which the researcher cannot recognize the phenomenon that the code equation operationalizes. Researchers must keep a record with notes concerning each excerpt identified by the code equation and why each false positive is not a case of the phenomenon. This process then allows researchers to identify cases that are anomalous to the operationalization—thus providing a further step of applying abduction to identify theoretically anomalous cases. Subsequent analysis can then compare and contrast false positives with verified cases.

Two key elements allow us to know if any code equation is useful. The first is data reduction. If the code equation does not reduce the data to a degree that is manageable for timely human analysis, then this is a sign that the equation is overly broad—and therefore the operationalization is excessively broad. The second is the number of false positives. While the number of acceptable false positives will depend on the data and the research question, our general rule is that if there are more false positives than matches, then the code equation is not precise enough.

During this step, the researcher(s) may also realize that the code equation is incomplete and use the stage of code equation verification to adjust the code equation and verify excerpts identified by the updated equation. The process of equation development may take some time and involve multiple iterations of a code equation. For example, if a code equation does not reduce—or minimally reduces—text excerpts, this is a sign the code equation could require adjustment. As a general tool in qualitative research, code equation verification can then be used to refine concepts. By double-checking code equations with the corresponding data to check for false positives, one not only verifies one's coding but also refines the operationalization of the phenomenon at hand. Code equation verification then allows the researcher to gauge if the operationalization—and corresponding conceptualization—is overly narrow or broad.

In the context of the abductive process, code equations are thus a key turning point. Code equation verification involves both small discoveries and large discoveries. The researcher makes small discoveries through code equation verification and then adjustment the equation. Once the researcher has found a code equation that is not too narrow (too many false positives) and not too broad (too little data reduced), then the analyst engages with the matching cases and can assess if the identified data from the code equation present a coherent and theoretically relevant discovery. Once the researcher has done this, she will know that she has made a “big” abductive discovery and therefore found a “final” code equation.

In our research, these equations led us to analyze all of the quotations that matched each equation. Through this process, two things became clear: First, the COOC (coocur) function was not specific enough and produced too many quotes that did not match the description of input or output legitimacy—or “false positives.” Second, rather than two different types of legitimacy, our analysis of the quotes revealed that respondents used *one* logic of legitimacy both to legitimate the UK and delegitimize the EU—which we labeled “Nationalist Neoliberal Euroscepticism.” (Vila-Henninger et al. 2019a, 2019b). Thus, through developing several iterations of the code equation, we found our final code equation:

Nationalist Neoliberal Euroscepticism: “Nationalist Neoliberal Euroscepticism” = “Neo-Liberal Fair Market Competition” AND “Political Community” AND “National Level” AND “In-Group/Out-Group” AND “Fairness” AND “Unfairness” AND “EU/European Level.”

See *Appendix* for the explanation of each code.

Application: We operationalized Nationalist Neoliberal Euroscepticism by combining two deductive codes sub-equations. First, UK respondents defined fair market competition relative to the interests of the UK or a subset of UK citizens. This means that respondents used the wellbeing of the UK, or a subset of its citizens, to define the rules for fair market competition—which is inherently nationalist. The result is that UK respondents then use this criterion to evaluate (a) UK actor(s) as behaving fairly (Sub-Equation 1). Next, UK participants delegitimized an EU-linked actor as unfair for not following the aforementioned nationalist market rules. By doing so, they define that actor as a member of an out-group (Sub-Equation 2).

Here, it is important to note our shift in theoretical frameworks. From our analysis of the quotes, we realized that the input/output distinction was not relevant to our analysis, and instead that it was more useful to understand these fairness legitimations as justifications of political power using widely accepted beliefs about “rightful standards of governance” (Beetham 1991; Beetham and Lord 1998).

We think that this entire process is not possible with inductive or machine learning approaches because our technique utilizes deductive codes and/or abductive codes. Thus, leveraging our deductive and abductive codes, we were able to engage in a second process of abduction that allowed us to incorporate more nuance that was relevant to our research question. This form of directed analysis emphasizes the research question of the secondary researcher rather than that of the primary researcher or probabilistic associations of words within a text (Karell and Freedman 2019, 729). This step is

analogous to Davidson et al.'s (2019) third step called "Preliminary analysis." However, rather than analyze excerpts of text that contain keywords, this step allows the researcher to perform a subsequent round of coding and preliminary analysis through abductive coding.

Step 3. In-Depth Abductive Qualitative Analysis

For our approach, abductive qualitative analysis includes two sub-steps: a) Inductive coding of reduced cases and b) (Manual) Qualitative analysis. It is important to note that through both stages of this step, the researcher(s) can apply Deterding and Waters (2018) "Stage 3: Exploring Coding Validity, Testing, and Refining Theory" (ibid: 23) abductively.

Step 3.a. Inductive coding of reduced cases. Using the verified cases from the QDA software code equation, researchers perform manual in-depth qualitative analysis. Part of this process is the generation of a subsequent inductive coding scheme specific to the codes identified by the code equation. This inductive coding scheme, which emerges from the analysis of verified cases from the code equation, allows to capture nuance in the operationalized phenomenon and to extend theory further by generating data-driven codes that capture greater detail in the phenomenon operationalized in the code equation that has not been yet documented. This sub-step may also result in a revision of the code equation and subsequently a return to step two.

Once we had isolated all of the cases that fit the final equation and eliminated all of the false positives in our data, we coded our cases inductively. Out of our abductive analysis, we found that the intersection of nationalism and neoliberalism in citizens' beliefs allowed for the emergence of a specific form of Euroscepticism in the UK from the mid-1990s until just after Brexit. Our coding reveals that this form of UK Euroscepticism was present in all our datasets – thus for more than two decades and across different social groups. More precisely, our qualitative analysis of our cases identified by the full equation captures the combination of nationalist and neoliberal beliefs that UK citizens draw from in their benchmarking of the EU with respect to the UK (Vila-Henninger et al. 2019b). We show that citizens hold a positive benchmark of national autonomy that makes the status quo of EU membership look like a bad deal (De Vries, 2018). These results were central to our analysis. It was through this second round of abductive analysis that we were able to identify anomalous cases that allowed us to understand our phenomenon of interest with a higher degree of nuance—and, therefore, build our theoretical explanation.

Step 3.b. In-depth (manual) qualitative analysis. The final step is an in-depth (manual) qualitative analysis and the writing of articles, research reports, and/or books. This step corresponds to Davidson et al.'s (2019) fourth step of "In-depth interpretive analysis" (372) and concludes Deterding and Waters (2018) third step "Exploring Validity, Testing and Refining Theory" (ibid: 23–25). We emphasize here that this is a "manual" analysis because it is conducted by the human researcher rather than by a QDA software itself or another computer program. During this step, researchers may also realize that the code equation from step 2, or the inductive codes from step 3.a, are incomplete. This would then result in returning to step 2 or 3.a to refine the analysis and provide a more complete empirical response to the research question.

Regarding our own research question, existing research has recently identified expressions of neoliberalism and nationalism in citizens' discourses (Taylor-Gooby et al. 2019; Andreouli and Nicholson 2018; Andreouli 2019). However, our approach allowed us not only to identify a hybrid form of Euroscepticism but also to illuminate how citizens navigate between nationalist and neoliberal beliefs in their benchmarking comparison of the UK's EU membership against autonomy. Below, we provide an example of the manual qualitative analysis that emerged and how we were able to mobilize our abductive coding and our in-depth qualitative analysis to contribute new theoretical insights while building on existing theories (Vila-Henninger et al. 2019a, 2019b).

Drawing on the discursive social psychology of Billig (1987), our approach proposes thereby to shift the focus from citizens' attitudes to the discursive articulation of Euroscepticism in lay citizens' discourses based on collective ideological repertoires – which we refer to as everyday Euroscepticism (Howarth & Andreouli, 2017). Specifically, this study sets out to systematically investigate whether and how UK citizens' Eurosceptic reasoning draws from nationalist and neoliberal ideology in order to contribute to the understanding of UK citizens' Euroscepticism leading up the Brexit referendum.

In this section, we illustrate the overall results presented in Table 4 by relying on quotes we consider typical of everyday UK Euroscepticism. These quotes are retrieved from our three datasets and, thereby, are illustrative of each period and the various sociodemographic and political characteristics of the participants. We discuss how nationalist (identity and sovereignty) and neoliberal beliefs (fair market competition) inform citizens' benchmarking comparison and highlight its decisive role for everyday Euroscepticism in the UK. Analytically, we focus on the commonalities between different periods of time, and different groups of citizens.

We begin with a collective interview from the CITAE data realized in Oxford in 2006. A respondent from an activist focus group, Ben – a 19-year-old conservative student – explains that “*you don’t need these sorts of crippling regulations which are sort of destroying the British economy*”. Referring to the issues surrounding the creation of “The Hazardous Waste (England and Wales) Regulations” (UK Parliament, 2005), he argues that “*regulations of certain disposals of waste [are] unnecessary*”. The respondent concludes by discussing the costs for the British economy: “*[The think tank] Civitas has calculated we’d be 40 billion pounds a year better off if we left the EU*”. Similar arguments to this were echoed by prominent proponents of Brexit – who also invoked the Civitas report mentioned by Ben. This is a typical example of the benchmarking mechanism as it shares key similarities with the attitudinal EU differential that De Vries’ (2018) found. We see a comparison of the state of the UK as a member of, versus autonomous from, the EU. But we also see that economic fairness, for Ben, is determined by the short-term interest of the British economy (“*we’d be 40 billion pounds a year better off if we left the EU*”). For him, there is an unfair imposition of EU market rules, or “*crippling regulations*”, upon the UK which are “*destroying the British economy*”. The EU’s supposed imposition of rules upon the British economy then compromised the sovereignty of the UK.

Discussion

Discussing abductive analysis and coding together is key to develop—or discover—what tactics for coding in emerging contemporary social science, in particular for secondary qualitative analysis or large-scale interview researches, could be. Not acknowledging the changing practices epitomized by (big) secondary analysis results in leaving many qualitative researchers ill-equipped to perform their analyses rigorously (Deterding and Waters 2018). We argue that abductive analysis can be further aided by careful methodological data analysis and coding. We therefore outlined how these coding tactics enrich the analysis and cultivate the conditions for generating creative and novel theoretical insights.

A key part of qualitative abduction is that there is a dialogue between each step of the process that serves to develop, extend, and refine theory. We see this both in the qualitative work on abduction (Timmermans and Tavory 2012; Tavory and Timmermans 2014) as well as the literature on Big Qual (Davidson et al. 2019). In this way, our approach seeks to engage in a dialogue between theory-driven and data-driven qualitative analysis. It is through this abductive process that researchers identify anomalous cases and incorporate them into the theoretical framework with which they are

performing abduction—thus engaging in theory building and extending abductive inference. We thus see our contribution as adding to the methodological work on abduction (Timmermans and Tavory 2012; Tavory and Timmermans 2014, 2019) that can hopefully be used to facilitate the blossoming empirical application of abduction in sociology and beyond (Carlson 2017a, 2017b, 2019; Timmermans 2017; Reilly 2018; Blee 2019; Karell and Freedman 2019; Vassenden and Jonvik 2019; Ignatow 2020; Vila-Henninger 2020).

In particular, code equations are a key tactic that we discuss. They allow researchers not only to see how different subcomponents of a phenomenon “hang together,” but also to understand how they relate to each other. Code equations then help to facilitate the abductive discovery process and are an anchor of our applied approach. Furthermore, code equations play an important role in our strategy of data reduction. We argue that code equations allow researchers to leverage codes to investigate elements of their research questions that are more complex than a single code. Machine learning and computational approaches could very well utilize theoretically-driven code equations—and we encourage those who use these tools to do so.

We also see this article as contributing to work on secondary qualitative analysis—in both the Francophone (for a review see Duchesne 2017) and Anglophone (for a review see Hughes and Tarrant 2020) literatures. We see secondary qualitative analysis emerging as an increasingly useful and necessary tool. We hope that our approach can help others to leverage their secondary analysis to build theory (Swedberg 2017). This article can also help to foster an exchange between work on qualitative secondary analysis (Duchesne 2017; Hughes and Tarrant 2020) and the abduction literature (e.g. Timmermans and Tavory 2012; Tavory and Timmermans 2014). As such, we hope that more scholars will be convinced by the opportunities offered by qualitative secondary analysis—both given the limitations of COVID-19 as well as the increased number of qualitative data repositories—and that this transition will include abduction. Finally, our approach helps to facilitate qualitative theory building through teamwork and may assist in strategies for the division of labor between researchers.

To conclude, it is important for us to note that secondary qualitative analysis should never be a replacement for primary qualitative analysis. Instead, it is a tool that can be used by researchers, when merited by the research question at hand, to build theory and to respond empirically to scientific inquiry that is impossible to answer with primary qualitative research alone. As such, primary and secondary qualitative research are complementary and can be used to extend the contribution of a given qualitative data set to further developing abductive inference.

Acknowledgements

The authors would first like to thank the primary researchers, Céline Belot, the CITAE team, Heidi Mercenier and Laurie Beaudonnet and the RESTEP team, for trusting us enough to give us access to their data. Furthermore, the authors would like to thank Sophie Duchesne, Edwards Rosalind and Lynn Jamieson and the members of the panel “New Approaches to Old Problems” at the 2020 Southern Political Science Association Conference in San Juan, Puerto Rico and the two helpful anonymous reviewers for their valuable suggestions and comments. Finally, the authors would also like to specifically thank Sophie Duchesne for her support and for inspiring our work mobilizing secondary qualitative analysis.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the project ERC Starting Grant Qualidem – Eroding Democracies. A qualitative (re-) appraisal of how policies shape democratic linkages in Western Democracies. The Qualidem project is supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement 716208).

ORCID iD

Vila-Henninger Luis <https://orcid.org/0000-0001-8871-3141>

Supplemental material

Supplemental material for this article is available online.

Notes

1. Note that the broader project includes six datasets composed of individual and collective interviews collected between 1995 and 2019 in 4 countries (Belgium, France, the UK, and the US) (Beaudonnet et al. 2022; Belot 2000; Duchesne et al. 2013; Mercenier 2019; Vila-Henninger 2020).
2. This list is in no way exhaustive. For a review of the broad range of archives and data repositories available for qualitative secondary analysis see Gaspard (2017), Edwards et al. (2021), and Haaker (2020).

3. Abduction has key precursors in qualitative sociological methods (Stryker 1996) and is related to a hand full of other qualitative approaches (Goldberg 2015; Mohr et al. 2015; Johnston 2017). These developments seem to respond to Lamont and Swidler's (2014) emphasis on the need in qualitative research for analytical approaches that apply broad theoretical toolkits.
4. According to Timmermans and Tavory (2018), Pierce distinguished two kinds of surprise: "novelty—or a new experience—and anomaly—or an unexpected experience." (171) In this article, we refer indifferently to surprising, anomalous, or novel results.
5. Davidson et al.'s (2019) qualitative "breadth-and-depth method" provides a very clear and stimulating approach to data reduction—which uses automatic coding to perform an iterative investigation of the data that combines computer-assisted analysis with in-depth interpretive analysis. Specifically, they use a technique called "recursive surface thematic mapping" that reduces data through auto-coding keywords that match a theme from theory or the literature. The approach is essentially to reduce data through auto-coding keywords that match a theme from theory or the literature—and is initially deductive in this sense. In this way, Davidson et al.'s (2019) method allows for deductive inferences about social processes.
6. Davidson et al. (2019) outline the first step in their approach as an "Overview survey of archived qualitative data and construction of a corpus" (369). We do not discuss in the article how the corpus is constructed. If applied in conjunction with Davidson et al.'s (2019) procedures, our step 1.a should be applied after this overview.

References

- Andreouli, Eleni. 2019. "Constructions of Europe in the run-up to the EU Referendum in the UK." *Identities: Global Studies in Power and Culture* 26(2):241-59.
- Andreouli, Eleni and Cathy Nicholson. 2018. "Brexit and Everyday Politics: An Analysis of Focus-Group Data on the EU Referendum." *POPS Political Psychology* 39(6):1323-38.
- Beaudonnet, Laurie, Céline Belot, Hélène Caune, Claire Dupuy, Anne-Marie Houde, Morgan Le Corre Juratic, Damien Pennetreau, Tiago Silva, and Virginie Van Ingelgom. 2022. "Studying (De-)Politicization of the EU from a Citizens Point of View: A New Comparative Focus Group Study". *Politique européenne* [accepted for publication].
- Belot, Céline. 2000. *L'Europe en citoyenneté. Jeunes Français et Britanniques dans le processus de légitimation de l'Union européenne*. Thèse doctorale de science politique, University of Grenoble.
- Biernacki, Richard. 2015. "Erratum: How to Do Things with Historical Texts." *American Journal of Cultural Sociology* 3(3):311-52.

- Beck, Cheryl Tatano (Ed.). 2018. *Secondary Qualitative Data Analysis in the Health and Social Sciences*. 1st Ed. London: Routledge.
- Beetham, David. 1991. *The Legitimation of Power*. Basingstoke: Palgrave Macmillan Ltd.
- Beetham, David and Christopher Lord. 1998. *Legitimacy and the European Union*. London/New York: Longman.
- Billig, Michael. 1987. *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge: Cambridge University Press.
- Bishop, Libby and Bren Neale. 2011. "Sharing Qualitative and Qualitative Longitudinal Data in the UK: Archiving Strategies and Development." *IASSIST Quarterly* 34(3-4):23–29.
- Blee, Kathleen. 2019. "How Field Relationships Shape Theorizing." *Sociological Methods & Research* 48(4):739-62.
- Brandt, Philipp and Stefan Timmermans. 2021. "Abductive Logic of Inquiry for Quantitative Research in the Digital Age." *Sociological Science* 8:191-210.
- Burawoy, Michael. 1998. "The Extended Case Method." *Sociological Theory* 16(1):4-33.
- Burawoy, Michael. 2009. *The Extended Case Method: Four Countries, Four Decades, Four Great Transformations, and One Theoretical Tradition*. Berkeley: California University Press.
- Bussi, Margherita, Claire Dupuy, and Virginie Van Ingelgom. 2022. "Do Social Policy Change Impact on Politics? Review of Policy Feedbacks on Citizens' Political Participation and Attitudes Towards Politics." *Journal of European Social Policy* [Accepted for publication].
- Campbell, Andrea Louise. 2003. *How Policies Make Citizens. Senior Political Activism and the American Welfare State*. Princeton: Princeton University Press.
- Carleheden, Mikael. 2016. "What Conception of the Theoretical Does Theorizing Presuppose? Comment on Richard Swedberg's Before Theory Comes Theorizing or How to Make Social Science More Interesting What Conception of the Theoretical Does Theorizing Presuppose?" *The British Journal of Sociology* 67(1):36-42.
- Carlson, Jennifer. 2017. "The Hidden Arm of the Law: Examining Administrative Justice in Gun Carry Licensing." *Law & Society Review* 51(2):346-78.
- Carlson, Jennifer. 2018. "Legally Armed but Presumed Dangerous: An Intersectional Analysis of Gun Carry Licensing as a Racial/Gender Degradation Ceremony." *Gender & Society* 32(2):204-27.
- Carlson, Jennifer. 2019. "Revisiting the Weberian Presumption: Gun Militarism, Gun Populism, and the Racial Politics of Legitimate Violence in Policing." *American Journal of Sociology* 125(3):633-82.

- Charmaz, Kathy. 2000 "Grounded Theory: Objectivist and Constructivist Methods." Pp. 509-35 in *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 2nd ed. Thousand Oaks, CA: Sage.
- Coffey, Amanda Jane and Paul A. Atkinson. 1996. *Making Sense of Qualitative Data: Complementary Research Strategies*. Thousand Oaks/New Delhi/London: Sage Publications, Inc.
- Corti, Louise. 2006. "Qualitative Archiving and Data Sharing: Extending the Reach and Impact of Qualitative Data." *IASSIST Quarterly* 29(3):8–13.
- Corti, Louise. 2012. "Recent developments in archiving social research." *International Journal of Social Research Methodology* 15(4):281–290.
- Corti, Louise, Andreas Witzel, and Louise Bishop. 2005. "On the Potentials and Problems of Secondary Analysis. An Introduction to the FQS Special Issue on Secondary Analysis of Qualitative Data." *FORUM: Qualitative Social Research* 6(1).
- Dalton, Russell J. 2007. *Democratic Challenges, Democratic Choices: The Erosion of Political Support in Advanced Industrial Democracies*. Oxford: Oxford Univ. Press.
- Davidson, Emma, Rosalind Edwards, Lynn Jamieson and Weller and Mary C. Waters. 2018. "Flexible Coding of In-Depth Interviews: A Twenty-First-Century Approach." *Sociological Methods & Research* 42(3):294-320.
- De Vries, Catherine E. 2018. *Euroscepticism and the Future of European Integration*. Oxford: Oxford University Press.
- DiMaggio, Paul, Manish Nag and David Blei. 2013. "Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41(6):570-606.
- Duchesne, Sophie, Elizabeth Frazer, Florence Haegel, and Virginie Van Ingelgom (2013). *Citizens' Reactions to European Integration Compared. Overlooking Europe*. Basingstone: Palgrave MacMillan.
- Duchesne, Sophie 2017. "Introduction. De L'analyse Secondaire à la Réanalyse. Une Innovation Méthodologique en Débats. *Recherches Qualitatives*, ARQ Association Pour la Recherche Qualitative, 2017, La Réanalyse à L'épreuve de L'expérimentation. Textes en Hommage à Annie-Claude Salomon." *Hors-série* 21:7-28.
- Easton, David. 1965. *A Systems Analysis of Political Life*. New York, NY: Wiley.
- Edwards, Rosalind, Emma Davidson, Lynn Jamieson and Susie Weller. 2021. "Theory and the breadth-and-depth method of analysing large amounts of qualitative data: a research note." *Quality & Quantity* 55:1275-80. <https://doi.org/10.1007/s11135-020-01054-x>.
- Fiedling, Nigel. 2004. "Getting the Most From Archived Qualitative Data: Epistemological, Practical and Professional Obstacles." *International Journal of Social Research Methodology* 7(1):97–104.

- Fine, Gary Alan and Iddo Tavory. 2019. "Interactionism in the Twenty-First Century: A Letter on Being-in-a-Meaningful-World." *Symbolic Interaction* 42(3):457-67.
- Glaser, Barney G. 1992. *Emergence vs. Forcing: Basics of Grounded Theory Analysis*. Mill Valley, CA: Sociology Press.
- Glaser, Barney G. and Anselm L. Strauss. 2005. *Awareness of Dying*. New Brunswick, NJ: Aldine Transaction.
- Goldberg, Amir. 2015. "In Defense of Forensic Social Science." *Big Data & Society* 2(2). <https://doi.org/10.1177/2053951715601145>
- Grimmer Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts". *Political Analysis* 21 (3): 267–297.
- Halford, Susan and Mike Savage. 2017. "Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research." *Sociology* 51(6):1132-48.
- Hammersley, Martyn. 2010. *Methodology*. London: Sage Publications. <http://public.e-liblib.com/choice/publicfullrecord.aspx?p=689522>.
- Hammersley, Martyn. (1997). "Qualitative Data Archiving: Some Reflections on its Prospects and Problems." *Sociology* 31(1):131–142.
- Harmes, Adam. 2012. "The Rise of Neoliberal Nationalism." *Review of International Political Economy* 19(1):59-86. <https://doi.org/10.1080/09692290.2010.507132>
- Hay, Colin. 2004. "The Normalizing Role of Rationalist Assumptions in the Institutional Embedding of Neoliberalism." *Economy and Society* 33(4):500-27. <https://doi.org/10.1080/0308514042000285260>
- Hay, Colin. 2007. *Why We Hate Politics*. Cambridge: Polity.
- Heaton, Janet. 2004. *Reworking Qualitative Data*. London: Sage.
- Howarth, Caroline and Eleni Andreouli (Ed.). 2017. *The Social Psychology of Everyday Politics*. London: Routledge.
- Hughes, Kahryn and Anna Tarrant. 2020. *Qualitative Secondary Analysis*. London: Sage.
- Ignatow, Gabe. 2020. *Sociological Theory in the Digital Age*.
- Irwin, Sarah. 2013. "Qualitative Secondary Data Analysis: Ethics, Epistemology and Context." *Progress in Development Studies* 13(4):295-306.
- Jerolmack, Colin and Shamus Khan. 2017. "The Analytic Lenses of Ethnography." *Socius* 3 <https://doi.org/10.1177/2378023117735256>.
- Johnston, Erin F. 2017. "Failing to Learn, or Learning to Fail? Accounting for Persistence in the Acquisition of Spiritual Disciplines." *Qualitative Sociology* 40(3):353-72.
- Kaltwasser, Cristóbal Rovira. 2017. *The Oxford Handbook of Populism*. New York, NY: Oxford University Press.
- Karell, Daniel and Michael Freedman. 2019. "Rhetorics of Radicalism." *American Sociological Review* 84(4):726-53.

- Kelle, Udo. 2007. "'Emergence' vs. 'Forcing' of Empirical Data? A Crucial Problem of 'Grounded Theory' Reconsidered." *Historical Social Research*. Supplement 19:133-56.
- Kitschelt, Herbert and Steven I. Wilkinson, (eds.). 2007. *Patrons, Clients and Policies. Patterns of Democratic Accountability and Political Competition*. Cambridge: Cambridge University Press.
- Lamont, Michèle. and Ann Swidler. 2014. "Methodological Pluralism and the Possibilities and Limits of Interviewing." *Qualitative Sociology* 37(2):153-71.
- Linde, Jonas and Joakim Ekman. 2003. "Satisfaction with Democracy: A Note on a Frequently Used Indicator in Comparative Politics." *European Journal of Political Research* 42(3):391-408. <https://doi.org/10.1111/1475-6765.00089>
- Mauthner, Natasha S. and Odette Parry. 2009. "Qualitative Data Preservation and Sharing in the Social Sciences: On Whose Philosophical Terms?" *Australian Journal of Social Issue* 44(3):291-307.
- Mauthner, Natacha S. and Odette Parry. 2013. "Open Access Digital Data Sharing: Principles, Policies and Practices." *Social Epistemology* 27(1):47. -67.
- Mauthner, Natasha S., Odette Parry and Kathryn Backett-Milburn. 1998. "The Data Are out There, or Are They? Implications for Archiving and Revisiting Qualitative Data." *Sociology* 32(4):733-45.
- Mercenier, Heidi. 2019. *C'est compliqué !" : L'Union européenne vue par les jeunes Bruxellois*. Thèse doctorale de science politique, Université Saint-Louis Bruxelles.
- Mettler, Susan. 2002. "Bringing the State Back In to Civic Engagement: Policy Feedback Effects of the G.I. Bill for World War II Veterans." *American Political Science Review* 96(2): 351-365.
- Mohr, John W., Robin Wagner-Pacifici, and Ronald L. Breiger. 2015. "Toward a Computational Hermeneutics." *Big Data & Society*, 2(2). <https://doi.org/10.1177/2053951715613809>
- Moore, Niahm. 2007. "(Re)-using Qualitative data?" *Sociological Research Online* 12(1):1-13.
- Nelson, Laura K. 2017. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods and Research*. 49(1):3-42.
- Norris, Pipa. 2011. *Democratic Deficit: Critical Citizens Revisited*. New York: Cambridge University Press.
- Parry, Odette and Natasha S. Mauthner. 2005. "Back to Basics: Who Re-uses Qualitative Data and Why?" *Sociology* 39(2):337-342.
- Peirce, Charles. 1934. *Collected Papers of Charles Sanders Peirce. Vol. 5, Pragmatism and Pragmaticism*, edited by Charles Hartshorne and Paul. Weiss. Cambridge, MA: Harvard University Press.

- Pierson, Paul. 1993. "When Effect Becomes Cause. Policy Feedback and Political Change." *World Politics* 45(4):595–628.
- Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs". *American Journal of Political Science* 54(1): 209–228.
- Reichertz, Jo. 2007 "Abduction: The Logic of Discovery of Grounded Theory." Pp. 214–28 in *The SAGE Handbook of Grounded Theory*, edited by A. Bryant and K. C. Charmaz. London: Sage.
- Reilly, Patrick. 2018. "No Laughter among Thieves: Authenticity and the Enforcement of Community Norms in Stand-Up Comedy." *American Sociological Review* 83(5):933-58.
- Rinehart, Kerry. 2021. "Abductive Analysis in Qualitative Inquiry." *Qualitative Inquiry* 27(2):303–311.
- Salganik, Matthew. 2018. *Bit by Bit. Social Research at the Digital Age*. Princeton: Princeton University Press.
- Savage, Mike. 2005. "Revisiting Classic Qualitative Studies." *Historical Social Research* 6(1): 118-39.
- Scharpf, Fritz W. 1997. *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*. Boulder, CO: Westview.
- Scharpf, Fritz W. 1999. *Governing in Europe: Effective and Democratic?* Oxford: Oxford University Press.
- Stoker, Gerry. and Colin. Hay. 2017. "Understanding and Challenging Populist Negativity Towards Politics: The Perspectives of British Citizens." *Political Studies* 65(1):4-23. <https://doi.org/10.1177/0032321715607511>
- Strauss, Anselm L. 1987. *Qualitative Analysis for Social Scientists*. New York: Cambridge University Press.
- Strübing, Jörg. 2007 "'Research as Pragmatic Problem-Solving: The Pragmatist Roots of Empirically-Grounded Theorizing.'" Pp. 580-602 in *The Sage Handbook of Grounded Theory*, edited by Antony Bryant and Kathy Charmaz. London: Sage.
- Stryker, Robin. 1996. "Beyond History Versus Theory: Strategic Narrative and Sociological Explanation." *Sociological Methods & Research* 24(3):304-52.
- Swedberg, Richard. 2017. "Theorizing in Sociological Research: A New Perspective, a New Departure?" *Annual Review of Sociology* 43(1):189-206.
- Tavory, Iddo. 2016. "The Pragmatist Wave of Theory Construction." *The British Journal of Sociology* 67(1):50-6.
- Tavory, Iddo and Stefan Timmermans. 2014. *Abductive Analysis: Theorizing Qualitative Research*. Chicago: University of Chicago Press.
- Tavory, Iddo and Stefan Timmermans. 2019. "Abductive Analysis and Grounded Theory." Pp. 532–46 in *The SAGE Handbook of Current*

- Developments in Grounded Theory*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd.
- Taylor-Gooby, Peter, Bjørn Hvinden, Steffen Mau, Benjamin Leruth, Mi Ah Schoyen and Adrienn Gyory. 2019. "Moral Economies of the Welfare State: A Qualitative Comparative Study." *Acta Sociologica* 62(2):119-34.
- Thorne, Sally. 2013. "Secondary Qualitative Data Analysis", in *Routledge International Handbook of Qualitative Nursing Research* (1st ed.), edited by Cheryl Tatano Beck. London: Routledge.
- Timmermans, Stefan. 2017. "Matching Genotype and Phenotype: A Pragmatist Semiotic Analysis of Clinical Exome Sequencing." *American Journal of Sociology* 123(1):136-77.
- Timmermans, Stefan and Iddo Tavory. 2012. "Theory Construction in Qualitative Research From Grounded Theory to Abductive Analysis." *Sociological Theory* 30(3):167-86.
- Torcal Mariano and Paolo Moncagatta. 2011. "Public Support". Pp. 2563-2566. in *International Encyclopedia of Political Science*. Edited by Bertrand Badie, Dirk Berg-Schlosser, and Leonardo Morlino, Thousand Oaks, California: Sage Publications.
- Mariano, Torcal and Paolo Moncagatta. 2011. "Public Support". Pp. 2563–2566. in *International Encyclopedia of Political Science*. Edited by Bertrand Badie, Dirk Berg-Schlosser, and Leonardo Morlino, Thousand Oaks, California: Sage Publications.
- Vassenden, Anders and Merete Jonvik. 2019. "Cultural Capital as a Hidden Asset: Culture, Egalitarianism and Inter-Class Social Encounters in Stavanger, Norway." *Cultural Sociology* 13(1):37-56.
- Vila-Henninger, Luis. 2020. *Social Justification and Political Legitimacy. How Voters Rationalize Direct Democratic Economic Policy in America*. Basingstoke: Palgrave MacMillan.
- Vila-Henninger, Luis, Mauro Caprioli, and Virginie Van Ingelgom. 2019. "Moral Economies of Euroscepticism: A Qualitative Diachronic Study". Paper Presented at the 115th American Political Science Association's Annual Meeting: "Populism and Privilege." August 29 - September 1, 2019, Washington.
- Vila-Henninger, Luis, Mauro Caprioli, and Virginie Van Ingelgom. 2019b. "Euroscepticism as Asymmetrical Nationalist Neoliberal Legitimation: A Case Study of British Citizens' Beliefs in EU Legitimacy". Paper presented at the 114th Annual Conference of the American Sociological Association, August 10-13, 2019, New York.
- Wagner-Pacifici, Robin, John W Mohr, and Ronald L Breiger. 2015. "Ontologies, Methodologies, and New Uses of Big Data in the Social and Cultural Sciences". *Big Data & Society* 2(2): 205395171561381.

Author Biographies

Luis Vila-Henninger is currently a postdoctoral fellow for the REACTOR grant in the Political Science Department at Aarhus University and was previously a postdoctoral fellow at UCLouvain in the framework of the ERC Starting Grant Qualidem. His dissertation and subsequent research use semi-structured interviews with respondents about economic ballot measures to investigate voter reasoning and legitimation. Luis' research areas include economic sociology, political sociology, sociology of culture, criminology, and sociological theory. Luis' work has appeared in *The Journal for the Theory of Social Behaviour*, *Sociology Compass*, *Sociological Perspectives*, and *The Sociological Quarterly*. Luis earned his PhD in sociology from the University of Arizona in August 2017.

Claire Dupuy is Professor of Comparative Politics at ISPOLE (UCLouvain). She specializes in comparative public policy with a focus on state transformations and regionalization processes in Western Europe. She is also interested in policy feedbacks and the ways in which policy changes impact on citizens' (dis)affection toward politics. Empirically, she conducts research on education and social policy in France, Germany, Belgium and the United Kingdom.

Virginie Van Ingelgom is a Research Associate Professor F.R.S. – FNRS at ISPOLE (UCLouvain). Her research interests include the issue of legitimacy at both the national and the European levels, indifference and ambivalence in citizens' attitudes towards the EU, policy feedbacks and methodological issues. Her previous work has been awarded with the Jean Blondel Ph.D. Prize by the European Consortium for Political Research (2012) Awarded with an ERC Starting Grant (2017-2023), Virginie Van Ingelgom recently developed with C. Dupuy a new research program that offers a qualitative (re)appraisal of citizens' (dis-)affection towards politics relying on the core argument of the policy feedback literature: attitudes and behaviors are outcomes of past policy.

Mauro Caprioli is currently working as a PhD student. His research is concerned with how European citizens perceive, experience, and define their relationship with politics in the current age of neo-liberalism, globalisation, and populism. His main research interests are political sociology, qualitative methods and methodologies, citizens' perceptions of politics, and right-wing populism.

Ferdinand Teuber is a postdoctoral fellow for the ERC QUALIDEM grant and research coordinator at ISPOLE (UCLouvain). His research interests center around the behavior of political elites inside and outside representative institutions. In his doctoral research he has examined the legislative agenda-setting behavior of Belgian legislators with the help of quantitative content analysis methods. Ferdinand holds a PhD degree from UCLouvain.

Damien Pennetreau received a FRESH grant from the Fonds national de la recherche scientifique (FNRS) and is currently working as a PhD research fellow in the framework of the Qualidem project (ERC Starting Grant) at UCLouvain. He is interested in research in the fields of communication, sociology, and political science. Combining these three strands (media studies, citizen opinion and discourse, and social policy), his doctoral research examines whether social policies influence citizens' relationship to politics. More precisely, it investigates how citizens discursively formulate their opinions about social policies and whether they are influenced by media coverage.

Margherita Bussi is currently an Assistant Professor at IACCHOS (UCLouvain) and was previously a postdoctoral fellow at UCLouvain in the framework of the ERC Starting Grant Qualidem. She is associate researcher at the European Trade Union Institute (Belgium). She was previously research fellow at the University of Brighton and earned her PhD in 2016 from the University of Geneva. Her interests include school-to-work transitions, activation policies and social policy implementation in Western Europe.

Cal Le Gall is currently a postdoctoral researcher in the DeVOTE project funded by the ERC at the University of Vienna. He was previously a postdoctoral researcher in the Qualidem project funded by the ERC at UCLouvain. He is associate researcher at Sciences Po Grenoble (Pacte). He previously earned his PhD in 2018 from the University of Salzburg and Sciences Po Grenoble. His research interests focus on the impact of economic globalization and European integration on voting behaviour and political attitudes.