DE GRUYTER OLDENBOURG Analyse & Kritik 01/2016; 257–285

Jörg Stolz

Opening the Black Box. How the Study of Social Mechanisms Can Benefit from the Use of Explanatory Mixed Methods*

"Never mix, never worry",
Martha in *Who's Afraid of Virginia Woolf*

Abstract: This article argues that analytical sociology—an approach that attempts to study social mechanisms 'without black boxes'—can benefit from the use of explanatory mixed methods. Analytical sociologists mainly relate their theoretical and agent-based models to representative surveys and experiments. While their central claim is to find and test the actual mechanisms that have produced the explanandum, the mechanisms they postulate often remain speculative. Neither agent-based models, nor experiments or mainstream quantitative methods, give access to some of the central elements of the causal mechanisms and the relevant subjective and objective contextual parameters. One of the most important reasons for this lies in the fact that social reality is changing fast, characterized by strong diversity and complexified by the phenomenon of cultural meanings. I argue that by creating and testing the models of analytical sociology with explanatory mixed methods, researchers have the possibility of getting closer to their object of research and therefore of having the chance to create more valid explanations.

1. Introduction

Ever since the publication of the seminal book on social mechanisms (Hedström/Swedberg, 1998), analytical sociology has become a major trend in current sociological theorizing and research. Drawing on earlier work especially by Elster (1989), Boudon (1974) and Schelling (2006[1978]), the central idea of the approach is that good sociology should be able to describe and explain social phenomena by realistically identifying the mechanisms at work. According to analytical sociologists, good explanations cannot be satisfied with reading off regression coefficients and interpreting them as causal influences. Instead, a good explanation has to "open the black box" and identify the "nuts and bolts, cogs and wheels of the internal machinery" (Elster, 1989) that have created the phenomenon to be explained.

* I thank Anaïd Lindemann and two anonymous reviewers for very helpful comments.

The claim of this article is that—due to methodological limitations—analytical sociologists have often not been able to reach their goal of really 'opening the black box'. Neither agent-based models nor experiments or mainstream quantitative methods—the methods usually applied by analytical sociologists—give access to some of the central elements of the actual mechanisms in the social world. One of the reasons for this lies in the fact that social mechanisms in the social world are very unstable and are only useful for analysis if central contextual parameters are known—parameters that cannot be investigated with the methods mentioned above.

I argue that by creating and testing the models of analytical sociology with explanatory mixed methods research, more valid explanations may be derived. Explanatory mixed methods is a special type of mixed methods research giving the analytical sociologist tools that allow for a closer inspection of both subjective (beliefs, preferences, emotions, heuristics) and objective (opportunities, resources, exogenous events) parts of the mechanisms and contextual factors involved.

The plan of the article is as follows. In *section 2* I argue that analytical sociology often does not live up to its central aim—namely to 'open the black box'. *Section 3* shows the general rationale for mixed methods and the specific added value of explanatory mixed methods for the study of social mechanisms. *Section 4*, the heart of the article, presents explanatory mixed methods in the form of five rules. These rules concern realist philosophical assumptions and the 'one logic of explanatory inference', the formulation of the explanatory research question, validity issues in the research design phase, data collection on mechanisms and contexts, and the reconstruction of mechanisms and contexts using abductive/detective triangulation. *Section 5* presents three examples showing how explanatory mixed methods have been used to study social mechanisms. *Section 6* concludes with a discussion of possible limits of the approach and next steps to take.

2. Analytical Sociology and Its Methodological Challenges

2.1 Analytical Sociology and Social Mechanisms[1]

Analytical sociology is a research program that explains social phenomena with social mechanisms.[2] While there is no generally accepted definition of 'mechanism', there seems to be convergence concerning the overall idea.[3] One of the most often cited versions stems from Machamer, Darden, and Craver (2000, 3),

[1] It is debatable if analytical sociology is the same as or different from explanatory sociology. For the former view, see Kalter/Kroneberg 2014, for the latter Manzo 2010. In the view of this writer, the two research programs are very similar. This is why I cite freely from explanatory sociologists when I present what is here called the 'analytical position'.

[2] For a history of the approach see Manzo 2010.

[3] See for a list of definitions of 'mechanism': Hedström/Bearman 2009b, 6. For a lucid comparison of analytical and relational sociology and their respective ideas about mechanisms see Koenig 2006.

who define mechanisms as "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions". According to this definition, specifying a mechanism means describing the relevant entities with their properties and their activities that jointly produce the final state. Explaining a phenomenon means—according to these authors—giving a description of the mechanism that produced it.

Analytical sociology presents a toolbox of various semi-general (or: middle-range) mechanisms that may be used in order to explain specific phenomena (Hedström/Bearman 2009b, 6). Such mechanisms are abstract in the sense that they do not specify the time, place, type of actors, etc.; yet they are specific in the sense that they only apply to a certain type of phenomenon. Thus, in the analytical sociology literature we find diffusion mechanisms, selection mechanisms, reproduction mechanisms, vacancy chain mechanisms, segregation mechanisms, mechanisms of self-fulfilling prophecies, etc.[4]

Advocates of analytical sociology see its main strength in the fact that by identifying the social mechanism it solves the black box problem. A black box problem is given when we know the inputs and the outputs of a model—but we do not know how and why the inputs are transformed into the outputs (Bunge, 2004). Thus, as Goldthorpe (1997, 57) writes, it may well be that a statistical model explains much variance of the dependent variable—but that we still do not know just "what is going on at the level of the social processes and action that underlie [...] the interplay of the variables that have been distinguished".[5] Opening the black box in this respect means pointing to a general mechanism or combination of mechanisms that exists—as a model—independently from the concrete case and giving evidence that such a mechanism is actually working in the case at hand. For example, we may notice that in France children of parents with high socio-economic status have much higher educational achievement than children of parents with lower status and we may explain educational achievement statistically through parents' socio-economic status. Opening the black box would mean showing empirically what abstract mechanism is responsible for this finding in the given French context. Is it a selection-discrimination mechanism as suggested by Bourdieu/Passeron (1985), an accumulative-decision mechanism as Boudon (1979[1973]) would have it or yet another mechanism (or a combination of various mechanisms)? Being able to solve the black box problem is—according to analytical sociologists—the central reason for the superiority of the analytical research program compared with other research programs that work with covering law explanations, purely statistical explanations or rational choice models (Hedström 2005, 13).

---

[4] See for collections of such mechanisms Boudon 1971, Mayntz 1987, Hedström/Bearman 2009a.

[5] Black box problems are not limited to quantitative research, but may appear equally in qualitative, historical, experimental or simulation research—they are given when the research is not able to get to the underlying mechanism that creates the explanandum. Goldthorpe 1997 makes this point for qualitative research in the QCA tradition—but it may be generalized. Thus, historical or qualitative research may give much detail and contextual information without clearly identifying the central causal links that account for the explanandum.

A serious problem with the analytical sociology approach is that while physical and biological mechanisms are often very stable, social mechanisms are relatively unstable. For example, mechanisms such as chemical transmission at synapses can be found in very stable ways across human brains—whereas the mechanisms of status reproduction will vary considerably across societies and different points in time. The reason for this basic instability lies in fast social change, strong social and cultural diversity, and high complexity produced through language and culture (Blalock 1984, 38ff.).

Compared with physical or biological change, social change is very rapid and its speed has risen in recent centuries (Rosa 2013). In the social world, individuals and groups may (voluntarily or non-voluntarily) innovate very quickly, leading to an immense speed of social change. Currently, social change is so rapid that as soon as we publish our results, they are often no longer true (Esser 1998). Likewise, compared with physical or biological diversity, social and cultural diversity is more pronounced (Kelle 2007, 57ff.). Because of the great plasticity of the social and cultural world and because of rapid and uneven social change, social and cultural diversity are immense. What is true of one group, milieu or society must therefore not necessarily be true for another group, milieu or society, since the relevant historical, geographic, normative, etc. contexts may be radically different. Finally, the existence of language and culture in human societies leads to an explosion of complexity that is not given in the physical and biological domain.

Less often mentioned, specific mechanism-based explanations involve not only identifying a mechanism, but also pinning down contextual parameters. Causal mechanisms always operate under specific socio-historic conditions. If, for example, we want to explain a specific case of the diffusion of an agricultural innovation, we have to specify not only the mechanisms, but also the specific starting values such as the number of actors, the starting point of the diffusion, the number of interactions between agents, the distribution of the probability of acceptance of the innovation, etc. (Hägerstrand 1965). In economics and rational choice literature, these contextual parameters are called 'initial conditions' with which the model starts out as well as the 'exogenous shocks' that influence the process during its operation. In explanatory sociology, the same issue is discussed under the heading of 'bridge hypotheses' (Esser 1998; Kelle/Lüdemann 1998; Lindenberg 1996). While explanations both in the natural and the social sciences require such contextual parameters, the importance of context is stronger in the latter case because of the basic instability of the mechanisms mentioned above.

To sum up, if sociology wants to explain by means of social mechanisms, it will have to accept that these mechanisms are much less stable than those in the physical and biological world and it will have to use powerful methods in order to find the relevant contextual parameters in the specific social situation at hand.

2.2 The Methodological Challenges of Analytical Sociology

An important challenge of analytical sociology to date is that its preferred methods—experiments, agent-based models or mainstream quantitative methods—are often not able to identify the correct mechanism and the important contextual parameters at work.[6]

Controlled experiments are strong with respect to internal validity, but weak with respect to external validity (Bernard 2000, 108; Bryman 2004, 34). Their causal claims, based on manipulation of the independent variable, randomization, and a control group, are often trustworthy and robust (internal validity), but it is often not clear whether these findings can be generalized to settings outside the experimental context (external validity). Since analytical sociologists—just as sociologists in general—are mainly interested in social processes in natural settings, this means that experiments often cannot open the black box of what is really of interest. Natural experiments (occurring naturally) and quasi-experiments (using a treatment and a control-group, but lacking randomization) can get us somewhat closer to the social realities we are studying and therefore augment external validity—but again the larger part of what interests sociologists will not be researchable by these methods (see for such experiments Bernard 2000, 127ff.; Campbell 1988[1974]).

Agent-based models are extremely attractive to analytical sociologists, since they allow building a model of the supposed mechanism by specifying the relevant entities (agents), properties of entities (attributes) and activities (procedures) (Manzo 2010, 147). The agent-based model then allows simulating the working of the supposed mechanism under a wide variety of parameters (Epstein 2006, 8). Yet, such models also have important disadvantages. In the stage of creation of the model, one often has to simplify matters very strongly, creating a sense of irrealism of the model (Manzo 2007, 59). If one tries to counteract this problem by making the model more complex and realistic, one often quickly encounters the further problem of no longer understanding just what is happening in the model and/or having to deal with the presence of too many exogenous variables.

In the stage of model specification, one often finds that the model needs a certain number of initial parameters that just have to be assumed and that cannot be justified with empirical data (Yang/Gilbert 2008, 6). In the stage of model validation, finally, it has to be noted that agent-based models give sufficient, but not necessary conditions for the emergence of macro-level phenomenon (Tubaro/Casilli 2010). Even if the model is able to perfectly reproduce the phenomenon to be explained, we cannot be sure that it has captured the right causal mechanism operating in the social reality.

Mainstream quantitative methods[7] are attractive since they permit multivariate analysis of representative samples. The results they present are often closer to the social realities than those by experimenters or agent-based modellers.

[6] See for other points of critique Opp 2005. See for elaborate responses to a whole list of critiques Manzo 2010.

[7] In the literature, various terms are used for what I call mainstream quantitative methods, e.g. variable sociology (Esser 1996), quantitative analysis of large-scale data sets (Goldthorpe 1996), etc. See for a whole list Manzo 2007, 35.

Again, though, mainstream quantitative methods have significant drawbacks. First, causal findings are often questionable (internal validity is low), since one can never be sure that all important independent variables have been controlled for (Manzo 2007, 39). Second, quantitative methods are normally unable to tell us just how a study should be designed, what data should be collected and what model should be specified (Freedman 2010, 221). Third, and central to analytical sociology's concern, quantitative mainstream methods describe the relations between variables, but they are normally not able to describe the processes that have generated these relations (Goldthorpe 2000a, 138). It is here that the black box problem is the most obvious, and it is this research that has prompted many analytical sociologists to search for other methods.

It has to be noted that leading exponents of the analytical sociology approach recommend *combining mainstream quantitative methods and agent-based modelling* (Hedström 2005, 116). Manzo (2007, 56) summarizes this line of thinking as follows: "[...] describe by means of variables—explain by means of mechanisms—formalize by means of simulations". While this seems to go in the right direction, it is obvious that in most cases important black box problems must persist since there is often no way to decide which one of a large number of possible agent-based models describes the correct mechanism that has created the correlations described by mainstream quantitative methods (Moss/Edmonds, 1128). In other words, we again face the validity issue of agent-based models.

Some authors have suggested that the way to meet these challenges is a combination of various methods, and some of the most convincing studies in analytical sociology in fact do use such methodologies (Manzo 2007, 147). On the whole, however, such combinations and especially the addition of qualitative and mixed methods in the analytical sociology framework have been rare.[8] To my knowledge there has so far been no link at all between the literature on analytical sociology and the literature on mixed methods. The idea that such a link might strongly benefit analytical sociology is the main contribution of this paper.

3. The Added Value of Mixed Methods

A growing number of researchers in the social sciences use mixed methods, that is, they combine quantitative and qualitative methods and data (M. M. E. Bergman 2008; Creswell/Clark/Gutmann/Hanson 2003; Kuckartz 2014; Tashakkori/Teddlie 1998; von der Lippe/Mey/Frommer 2011). Mixed methods methodology is by now established, with a large number of publications, a handbook (Tashakkori/Teddlie 2010), an international research association,[9] and two journals.[10] In this paper I define mixed methods as a research strategy that com-

[8] The index of the arguably three most prominent books on analytical sociology to date—Hedström 2005, Hedström/Swedberg 1998 and Hedström/Bearman 2009a—all lack the keyword 'qualitative' or 'mixed methods'.

[9] See http://mmira.wildapricot.org.

[10] The *Journal of Mixed Methods Research* and the *International Journal of Multiple Research Approaches*.

bines data collection strategies of different levels (more or less structured, variables/cases) and combines and triangulates different data types (text and numbers, nominal/ordinal/interval) in order to create better descriptive and explanatory inferences.[11] In order to get a grasp of mixed methods, it is important to understand (1) the nature of the distinction and combination of qual and quan, and (2) the rationale for mixed methods. I will treat both points in order.

3.1 The Difference between Quantitative and Qualitative Methods

Two major ways to distinguish between qualitative and quantitative methods can be identified in the literature (Bryman 1988, 5). Some researchers opt for a distinction mainly on *epistemological grounds*. According to this account, quantitative and qualitative methods are grounded in different philosophical and epistemological assumptions—in different 'paradigms'. For example, qualitative methods would imply the negation of an external reality, an espousal of constructivism, the abnegation of causality and the belief that research is necessarily value-laden, while quantitative methods would rest on the assumptions of an external reality, a post-positivist position, the search for causal relationships and a search for value-free positions. There are serious problems, both empirical and practical, with the epistemological position (M. M. Bergman 2008). Empirically, one can point to many examples that do not respect the alleged link between methods and epistemology. Practically, it seems that by insisting on two completely different worlds of research, good research opportunities that would require a combination of methods and data are excluded from the start.

Other researchers distinguish qualitative and quantitative methods on *technical grounds*. For them, "quantitative and qualitative research are simply denotations of different ways of conducting social investigations and which may be conceived of as being appropriate to different kinds of research question [...]" (Bryman 1988, 5). According to this position, some of the most important differences are that qualitative research uses relatively small Ns (mostly) text, an only nominal level of measurement, and relatively unstructured instruments, while quantitative research uses relatively large Ns (mostly) numbers, all kinds of levels of measurement (from nominal to metric), and relatively structured instruments. Often, these and other distinctions are to be seen not as implying either/or choices, but as continua. Researchers can design their instruments as more or less structured; they can vary their N, etc.

The technical position allows for an excellent compatibility with analytic sociology: it fits with the realist paradigm most often used in analytical sociology; it is compatible with the idea of 'one logic of inference' (that will be explained below); and it allows the possibility of explaining with mechanisms. It is there-

[11] This definition shows that I reject the view of two radically different worlds of quan and qual. Accordingly, already the talk about 'mixed methods' is somewhat problematic, since it suggests a strong difference that I would rather try to overcome. However, for convenience and in order to locate my work in the state of the art, I keep the label 'mixed methods'. See for a critique of the 'two worlds approach' Meulemann 2002; Bergman 2008. For an overview of different definitions see Johnson/Onwuegbuzie/Turner 2007.

fore this technical view of the difference between qual and quan that will be used in the paper.

3.2 The Rationale for Mixed Methods

In the literature many different *rationales* for mixed methods are given (Greene/Caracelli/Graham 1989). The rationale that seems most important for analytical sociologists is that mixed methods—because of non-overlapping strengths and weaknesses of the quan and qual parts—can be a way to strengthen the validity of our results, and especially our inferences about causal mechanisms and contextual parameters (Hammersley 2008; Kelle 2007, 47; Onwuegbuzie/Teddlie 2003).[12]

This leads, of course, to the question of what is meant by the terms *validity* and *validity threat*. Here, I suggest following the very broad definitions given by Maxwell (2005, 106). Validity, according to him, is the "[...] correctness or credibility of a description, conclusion, explanation, interpretation, or other sort of account."[13] Two important points have to be noted concerning this view of validity. First, validity is not a matter of correct procedures, but of the correctness of the result of procedures. As Shadish, Cook, and Campbell (2002, 34) write: "Validity is a property of inferences. It is not a property of designs or methods, for the same designs may contribute to more or less valid inferences under different circumstances. [...] No method guarantees the validity of an inference." Second, as Maxwell (2010, 19) notes, validity of inferences has to be shown not only by supporting evidence, but also by excluding alternative accounts (hypotheses, descriptions, mechanisms etc.). A validity threat is then "[...] a way you might be wrong" about the claims made by a piece of research (Maxwell 2005, 105ff.). Very often, such validity threats take the form of rival explanations.

We can now reformulate the additional benefit of mixed methods. It lies in the fact that it allows *eliminating validity threats given in just one but not the other methodological tradition*, leading to more valid inferences (Kelle 2007, 227ff.). The typical validity threats of mono-quan research are that the real mechanisms are often not observable; it is difficult to know what variables should be included in models; the attributes of the situations and their perception by actors remains unclear; the rules of the game have to be guessed; and causal

[12] Many authors see a rationale for mixed methods in that it allows researchers to present different perspectives, thus giving a more complete picture of the phenomenon. However, I would argue that there is no obvious scientific merit in adding an additional perspective or giving a more complete description. It is extremely easy to add some information on a given phenomenon—but it is not always better to do so. Furthermore, it is impossible to present a complete picture of a phenomenon in all its complexity, however thick one's description might be. Additional perspectives therefore only make sense when they allow the researchers to better answer their research questions, in other words when it leads to higher validity. See for a similar argument Bergman 2011, 274.

[13] This definition embraces the various types of validity that are normally distinguished such as measurement validity, internal validity (concerning causal claims), external validity (concerning generalization), and ecological validity (concerning relevance in everyday life). See for these concepts Bryman 2004.

complexity is often difficult to capture—in all of these areas, qualitative methods are often stronger. In mono-qual research, on the other hand, it is difficult to capture weak causalities, assess the relative strength of causal factors, assess statistical significance, control for other variables, capture central tendencies and variability, make assertions on the general distribution of types and assess representativeness in general—in all of these areas, quan methods are much stronger.

However, mixed methods is no miracle solution to all of our methodological problems and has to be used with caution. First, it has always to be remembered that, since validity cannot be bought with method, we cannot just trust our mixing to give more valid inferences—we have to judge the results of our research as to its validity. Second, as we have argued above, the distinction between quan and qual methods and data is less clear-cut than one might normally think and the validity threats attached to each method and data type have to be judged in a specific manner in their specific context in every research. Third, it does not suffice to point to different validity threats of different methods and data in a given mixed methods research project—we also have to convincingly show how the combination is able to eliminate the specific validity threats of the other method or data type. Fourth, quan and qual methods are sometimes fraught with the same or similar validity threats—and mixing will therefore not eliminate them. For example, standardized and in-depth interviews both have a validity issue with reactivity, with the fact that process is not very well captured, and with the fact that reported action is not the same as actual action (Bryman 1988, 112).

4. Investigating Social Mechanisms with Explanatory Mixed Methods

Not every mixed methods study lends itself automatically to the study of social mechanisms. In what follows, I present the research program that I call 'explanatory mixed methods' in the form of five interlinked rules. I argue that only this specific form of mixed methods will allow to 'open the black box'.

4.1 Use a Realist Philosophical Paradigm and One Logic of Explanatory Inference

Explanatory mixed methods uses what might roughly be called a *realist philosophical stance* for one's research (Maxwell 2010). I do not think that for practical purposes it is necessary to dig too deep into the philosophical discussion. Suffice it to say that most realists would assume with Brante (2001) that (1) there is both a material and a social "reality existing independently of our representations or awareness of it (ontological postulate)", that (2) "it is possible to achieve knowledge about this reality (epistemological postulate" and that (3) "all knowledge is fallible—and correctable (methodological postulate)".

Such a position—quite close to common sense—is perfectly compatible with analytical sociology, since it allows for the idea of causality through mechanisms and allows for unobserved entities. It is also perfectly compatible with certain types of both quantitative and qualitative methods. While this might be obvious for quantitative research, recall that well-respected scholars working with qualitative methods, such as Martyn Hammersley or Matthew Miles and Michael Huberman, have implicitly or explicitly leaned towards realist positions. Such a position differs from many constructivist views often taken by qualitative researchers who deny the usefulness of the concept of causality and think that explanation in the social sciences is either impossible or useless. It also differs from positivist positions often taken by quantitative researchers who would see causality in a Humean way as 'robust dependency', thus rejecting the need for mechanism in-depth explanation.

A realist stance is also well suited for an analytically oriented mixed methods because it allows using the concepts of *"one logic of inference"* (Goldthorpe 2000b; King/Keohane/Verba 1994).[14] According to this idea, all science builds on the same logic of inference, independently of what methods are used—quantitative, qualitative, historical, physical, etc. *Figure 1* shows what is meant by this. We start with the realist assumption of a 'reality out there', where real things happen. We cannot observe this reality directly and we cannot embrace it completely, which is why we necessarily have to resort to some sort of sampling and to (more or less structured) data collection, e.g. interviews, observation, document analysis, etc. On the basis of the analyzed data, we then make inferences, that is, we draw conclusions about what we think is true about the world.[15] Inference can be descriptive (point to facts) or explanatory (point to causal mechanisms). Adherents of the one logic of inference therefore believe that it is possible to create knowledge about the external, existing world; however, this knowledge is always uncertain. There could always have been biases in the sampling, errors when collecting or analyzing the data, faulty assumptions when drawing conclusions, etc. This means that we always have to address the question of the quality of (a) the data collection, (b) the resulting data, and (c) our inferences. Mixed methods are used precisely because researchers believe that they will lead to better—more reliable and valid—inferences about facts and causal mechanisms in the real world.

---

[14] It seems important to note that analytically minded scholars using mixed methods do not have to accept the entire program suggested by King/Keohane/Verba 1994 to go with "one logic of inference". In my view, King/Keohane/Verba have rightfully been criticized for applying the quantitative template too strictly to qualitative methods (Brady/Collier 2010; McKeown 1999). Thus, I think that it is generally useful to think of social research as making inferences to an underlying reality and that there is always an uncertainty—but it is not always useful to express this relationship in the language of means and variance as King et al. do.

[15] This inference to an underlying reality can and should be distinguished from the inference quantitative researchers mean when they make inferences from the sample to the population in a survey. The first inference is always given, since we can view all data as hypothetical results of an infinite number of possible 'runs'. The second inference is only possible if we take a sample from a clearly delimited population of cases (and impossible in the case of a total survey). See for this point Kelle 2007, 246.

Accepting 'one logic of inference' means rejecting the temptation of wanting to combine incompatible philosophical and methodological assumptions in one mixed methods project. Some researchers want to do the quantitative and qualitative parts of their research as much as possible according to the standard epistemological assumptions and methodological practices of qual and quan communities. They then find themselves trying to be constructivists for their qual part and postpositivists for their quan parts. Or they assume causality when analyzing their quan data, but reject the notion of causality when interpreting their interview transcripts. Or again, they argue that they can only explain a certain amount of the variance with quan, but everything with qual. Such positions are evidently untenable and cannot lead to useful research.

Combining mixed methods with the idea of one logic of inference gives us the philosophical backbone for the idea that *triangulation* of different types of data will help us eliminate validity threats linked to only one type of data.[16] When analytical sociologists use mixed methods, it is because they think that triangulation will lead to better explanatory inferences, a way of better getting at the real causal mechanisms and contextual parameters at work.[17]

---

[16] The term has been popularized by Denzin 1978. See for a discussion of the limits of the metaphor Kelle 2001. Note that I use a narrow definition of triangulation. Other authors (Denzin 1978; Flick 2002, 331) use a very broad definition that included the triangulation of data, investigators, theory and methods. However, it seems to me that such a wide use of triangulation continues to think in a kind of methodological dualism that I would like to leave behind. In contrast, my use of triangulation is linked to the idea of "one logic of inference".

[17] As Hammersley 2008 has shown, the concept of triangulation runs into serious troubles if we accept the constructivist and postmodernist assumption that different methods construct radically different realities—for why should we then put them together in a triangulation in the first place? Of course, one logic of inference does not disagree with more realist forms of constructivism. It is consistent with the idea that we see different things when applying different methods—that is why we use mixed methods. Likewise, it is consistent with the possibility that actors may have radically different subjective viewpoints, values or ideas concerning the same object. But then, these viewpoints, values or ideas are still seen to be as something 'real', 'out there', which we may more or less reliably assess with our methods. For an interesting critique of many forms of constructivism see Hacking 1999.

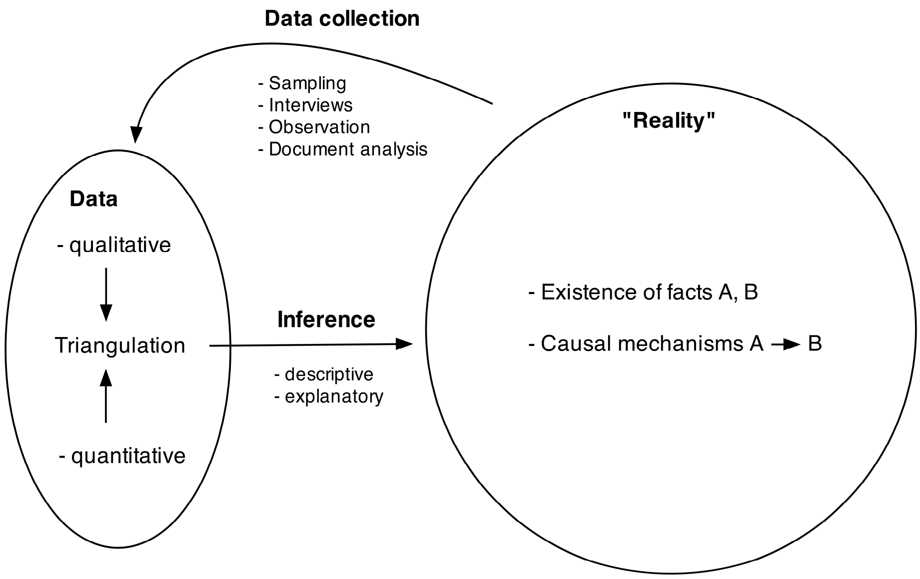

Figure 1: A model of triangulation in the framework of 'one logic of inference' (source: own figure).

4.2 Put Your Explanatory Research Question First

Just like other social scientific research designs, explanatory mixed methods research designs should be built around *one central research question.* The central research question states just what it is that the researcher wants to understand. It can normally be formulated in one or two sentences—and has the form of a question. Often a central overarching question is followed by a series of subquestions that are enclosed by the central question. When we combine analytical sociology with mixed methods, our central research question should evidently be explanatory: we want to know *why* a certain phenomenon appears or *how* the mechanisms work that have caused our explanandum.

As Maxwell (2005) shows, in a good research design, the central research question is *strongly integrated* with all other elements of the research design (the overall goals of the research, the methods, the theory, the validity aspects). As a matter of fact, most unsuccessful research projects fail because they are poorly designed—and most often the poor design is due to a lack of integration: the research question cannot be answered with this method; or the theory does not really give a clear hypothesis answering the research question, etc. A strong integration of the research design around one central question is especially important in the case of mixed methods research, because experience shows that in many mixed methods research projects the quantitative and the qualitative parts seem to answer different research questions and cannot really be integrated

(Bryman 2008, 99). The point I am making here is that a good integration at the stage of the triangulation is only possible if there is one overarching central question.

This means that we resist the temptation of formulating two different research questions for quan and qual, as is often proposed in the literature. Creswell (1998, 17) proposes that qualitative research asks "how" or "what", while quantitative research asks "why". Maxwell (2005, 74) thinks that qualitative research uses "process-questions" while quantitative research uses "variance-questions". But despite what is said in these well-known qualitative textbooks, there is no necessary link between the overarching question and the quan or qual method. Quantitative methods can very well be used to answer questions about the 'how' or the 'what' of a social object and qualitative methods can work very adequately with why-questions. This is not to say that we should use the same type of questions for quan and qual all along the research project—that would be impossible. What I do mean, however, is that on the most general level of the research, there should be only one central thing we want to know. Otherwise, we will be conducting two separate research projects and not one; triangulation will be difficult or impossible; and the central advantage of mixed methods—which is why we wanted to do it in the first place—will be lost.[18]

4.3 Address Validity Issues with Mixed Methods Already in the Research Design Phase

The central idea of explanatory mixed methods is to weed out validity threats of mechanism explanations through the combination of different methods and data types. Much thought should therefore be given already in the research design and data collection phase to just how such validity threats are going to be addressed.

Of course, the specific validity threats depend on the specific mixed methods *research design* used. Much of the mixed methods literature is concerned with typifying and describing such designs according to various criteria (temporal sequence, dominance of the components, confirmatory/exploratory etc.). Following Creswell/Plano Clark (2007, 58ff.) we can distinguish four major designs (that each have their sub-designs for different purposes: a triangulation design,[19]

[18] Another way of saying this is that the one central research question should be more important than the mixing. We should not do mixed methods as a goal in itself. We should use it if—and only if—it leads us to a better way of answering our one research question.

[19] In the triangulation design (qual – quan, concurrent), researchers will conduct the quan and qual parts of the research independently (and both during one phase of the research) in order to triangulate the different data types.

an embedded design,[20] an exploratory design,[21] and an explanatory design.[22] I will not go more deeply into these designs and subdesigns since this has been done extensively in the literature (Tashakkori/Teddlie 1998, 40ff.; Teddlie/Yu 2008). My point is that *all* of these designs (and not just the explanatory design) can be put to good explanatory use by analytical sociologists; and in all of them, mixed methods can be used in order to address validity issues. Whatever the design, however, mixed methods studies do *not automatically* create more valid results than mono-method studies. They only do so under very *specific* conditions where researchers identify *specific* validity threats that can be eliminated with a *specific* research design. In the three examples given below in *section 5*, I show how mixed methods have been successfully used to this end in very specific ways. Independently of the specific research design, three pieces of general advice can be given.

First, *sampling* in qual and quan should be designed such that triangulation becomes as straightforward as possible. Many different mixed methods sampling strategies are discussed in the literature (Kemper/Stringfield/Teddlie 2003; Teddlie/Yu 2008). Textbooks often give the incorrect impression that qualitative research necessarily uses purposeful, non-random sampling, while quantitative research should necessarily use random sampling. However, as Bergman (2011, 273) correctly notes, "it is conceivable to draw a stratified random or random cluster sample for small-scale qual research, or to draw a snowball or atypical case sample for research associated with a quan (non-inferential) methods". Nothing forbids using similar sampling in qual and quan—and it is often preferable to use sampling that is as close as possible in both parts. In fact, if very different sampling is employed for the respective qual and quan data sets (e.g. theoretical sampling or maximum contrast sampling for qual and random sampling for quan), it is often difficult to see how the data sets may be integrated in order to eliminate validity threats. It is then not clear if differences or complementarities found are due to substantial phenomena or to the different sampling procedures.

In all four research designs mentioned, it is at least in principle possible to use *nested samples*, that is, to have the qual sample be a subsample of the quan sample. This is attractive since it allows us to compare the two samples and to judge the extent to which the two samples really focus the same reality. We can therefore say something about the quality of our triangulation. For example, Stolz et al. (2014) analyzed a representative sample of members of free churches in Switzerland and then drew 32 respondents in a stratified random way from this sample in order to conduct in-depth interviews; they were able to compare and

[20] In the embedded design (i.e. Quan – qual – Quan; Qual – quan – Qual), "one data set provides a supportive, secondary role in a study based primarily on the other data type" (Creswell/Plano Clark 2007, 67). For example, a standard quantitative laboratory experiment on heuristics may be supplemented with a small number of open-ended questions on how participants perceived and solved the questions from their subjective point of view.

[21] The exploratory design (qual – quan, sequential) is a two-phased design where researchers conduct a qualitative study and then set out to test their resulting hypotheses quantitatively.

[22] The explanatory design (quan – qual, sequential) is a two-phased design that is used when quantitative studies lead to astounding results that are inexplicable with the quantitative data at hand. In a second qualitative phase, researchers therefore look for appropriate qualitative, historical, ethnographic, interview and other material in order to explain their findings.

judge their two data-sets. Likewise, Cherlin et al. (2004) drew a random sample of 2402 children and their care-givers in low-income neighborhoods in three US cities and recruited 256 additional children and families for an ethnographic study, non-randomly, but from the same neighborhood. While a properly nested random qual sample would have been even more attractive, the qual sample of Cherlin et al. can be considered a subsample of the quan sample - and the authors compare the two samples in their paper.

Nesting samples also—at least in principle—allows us to describe mechanisms in the sample with both qual and quan methods and to make population inferences, that is, argue that these mechanisms exist in the population. However, as Bergman (2011, 273) convincingly argues, we have to be very careful when doing this. If the qual sample is only exploratory and non-random, it "is not suitable for population inference because dimensionality identified in a small sample drawn nonprobabilistically may not include all relevant dimensions. Although the large-scale survey may indeed allow for population inference, it is nevertheless limited to the constraints imposed by the dimensionality identified from the sample associated with the qual component." This is a fundamental point that is almost never raised in the literature. Mixed methods designs that want to counter it, will have to make sure that, for example, we can be confident that the qual sample is a representative subsample of the quan sample, that the qual sample is relatively large, or that we have other reasons to believe that we can generalize from our qual sample to the population (e.g. because we have assembled all existing critical cases) (see for proposals for such mixed methods sampling strategies Kelle 2007, 227ff.).[23]

Second, *data collection should be designed such that triangulation becomes as straightforward as possible.* Triangulation is a comparative technique and comparison is easier if not too many dimensions vary. It is thus often a good idea to capture the same theoretical dimensions in different, but comparable data modes (as 'codes' in qual or independent and dependent 'variables' in quan). Also, it is a good idea to measure/capture the different qual and quan dimensions on the same cases (individuals, families, countries). There exist many cases of mixed methods research where the qual and quan components are so different that the respective parts cannot address each other's validity problems. The added value of such studies is small; they often only produce "thematically connected mono-method research outputs" (Bergman 2011). For example, Way et al. (1994) use an explanatory design (in Creswell/Clarks' terminology). They find in a quan study that depression and drug use is highly correlated in one (suburban) school, but not in another (urban) school in the US. They follow up their quan-substudy with a qual-substudy in order to explain this interesting finding. They draw a nested subsample, but unfortunately, they only sample the most depressive students in the two schools for in-depth interviews. While they find some interesting new results, they cannot answer their research question, because they are not sure what they would have found had they also taken an

[23] See for Mixed Methods studies that solve this problem by using relatively large qualitative samples, sampling plans and a strong integration of the qual and quan datasets Stolz et al. 2015; Cherlin et al. 2004.

in-depth look at the not depressive in both schools. Quite evidently, the authors should have used a more comparable sampling in both the quan and qual parts of their study.

Third, researchers should build as many *cross-validating opportunities* into their research design as possible. In what has been called a triangulation design above, it is possible to let the respondents of the smaller qual sample also respond to the quan questionnaire. This permits a quan comparison of both datasets. If an ethnography is used to create a typology that is then used in a quantitative study, it would be a good idea to apply the quantitative survey also to the site of the ethnography—to judge if the quantitative instrument actually captures the typology created with qualitative means (at least in this one case).[24] If a representative survey is combined with in-depth interviews, it is a good idea to let the respondents of the in-depth interviews also respond to selected items from the quan study and/or to include some open questions from the qual study in the quan survey. Often quantitizing and qualitizing techniques can be used for such additional cross-validating purposes (Tashakkori/Teddlie 1998, 125).

4.4 Collect Systematic Data on Mechanisms and Relevant Contextual Parameters

Again, the collection of data depends strongly on the research question and design chosen. In mixed methods studies it can take the form of ethnography, observation, documentary analysis, in-depth interviews, expert interviews, surveys; it can be one-shot or longitudinal, experimental or non-experimental etc. Four general pieces of advice specifically for explanatory mixed methods research can however be given.

First, *collect data that is able to represent your mechanism narrative and relevant contextual parameters* (Abbott 1992; Miles/Huberman 1994, 177). Explaining means testing a causal story that functions with one or several mechanisms, including several actors-in-situations. It is therefore crucial to collect data that may map the temporal order of the implied mechanism, i.e. stories about what has happened, variables and codes that represent initial situations, intermediate situations and the explanandum, etc. Furthermore, it is important to collect data that describe the essential elements of the situations and actions-in-situations of actors. This means, for example, to describe and measure the opportunities, desires and beliefs of actors at different points in the causal story.

Second, *collect all the data implied by your theory that you can get, irrespective of the level or data type*. King/Keohane/Verba (1994, 46) suggest that "the best scientific way to organize facts is as observable implications of some theory or hypothesis". Researchers should ask themselves just what observable implications their theory should have and should then go looking for those data—be they qualitative, quantitative, or historical. In a similar way, researchers should

[24] See Laubach 2005 for an example of a good mixed methods study that could have been improved by such a cross-validation.

ask themselves just what data would be needed in order to refute their theory (King et al. 1994, 19).[25]

Third, collect your data *systematically* both in the qual and the quan parts. Even if there are exploratory and inductive parts in your research design (often in the qual research), there has to be a point where systematic data collection sets in and where clear rules of data collection can be mentioned. Systematic data collection according to explicit rules is important since it guarantees transparency, allows judging the quality of the data and possible validity threats and permits and—at least in principle—replication (Goldthorpe 2000b). It is only with a sense of the quality of our data and possible validity threats that we can set out to try to triangulate different data sets, hoping that different validity threats will be attenuated by the combination of methods.

Fourth, *do not associate qual and quan with different objects of research*. Thus, it would be erroneous to think that one of the methods (say: quan) were better at focusing on mechanisms and the other methods were better at focusing on contexts (say: qual). Nor should one think that quan was better at specifying the explanandum and qual the explanans (mechanisms and contexts); or that quan was better at showing macro conditions while qual was better at micro behaviour; or that objective facts should be investigated with quan and subjective phenomena with qual; or that structure had to be addressed with quan and process with qual. All of these often made associations break down when we look at mixed methods research as it is actually practiced. Rather, it depends on the specific research domain as to just what method may elucidate what element of the phenomenon, and most times, at least in principle, qual and quan methods and data can focus on the same elements of the research phenomenon.

4.5 Reconstruct the Mechanisms and Contexts Using Abductive/ Detective Triangulation

Data analysis in explanatory mixed methods research takes the form of triangulation (Tashakkori/Teddlie 1998, 41). Triangulation may be defined as a kind of data analysis that uses different types of data in order to make better—more valid—inferences to an unobserved reality. We can distinguish (a) descriptive triangulation that combines different data sources in order to better describe a social fact, from (b) explanatory triangulation that combines data sources in order to make inferences to a causal mechanism or narrative.[26] Analytical sociology will normally aim for explanatory triangulation. How does explanatory triangulation work in practice? While this varies with the type of research design, there seem to be some general pieces of advice that can be given.

First, work in an *iterative, abductive and detective-like* way. Start with the assumption that the same reality has created the different data sets. Therefore, what you find in one data set can or should show up in the other one. By

---

[25] This is of course a useful way of rephrasing the Popper (1980[1959]) criteria of falsifiability.

[26] This distinction is made in analogy to descriptive and explanatory inference in King et al. 1994.

analyzing one data set you will therefore find new hypotheses, new questions to ask that you can use to analyze the other data set and vice versa. In this way you can switch iteratively from one data set to the other, from variable view to case view and from qual to quan by trying to get an ever finer-grained view of the causal mechanism. This way of analyzing data is abductive in the sense that you try to explain given data phenomena by assuming a hypothetical causal mechanism (Peirce 2006). Abduction, as Peirce conceives it, is not an alternative but a complement to both deductive and inductive reasoning. It needs "deduction of the consequences of that hypothesis; and inductive testing of those consequences to determine how likely it is that the hypothesis is true" (Haack 2006). Thus, researchers who triangulate can be compared with detectives. Just as a detective "solves a crime by looking at clues and suspects and piecing together a convincing explanation, based on fine-grained evidence that bears on potential suspects means, motives, and opportunity to have committed the crime in question" (Bennett 2010, 207), triangulating researchers attempt to answer their research question by continually looking out for traces, new elements in their data, but equally for new hypotheses that would explain the different elements in a consistent manner.[27] They would also continually look out for elements in their data that would allow them to rule out one or several of their hypotheses. For example, you might find a suggestion for a new intervening variable in your qual data and could try to operationalize and test this in your quan data. Or you might find a significant correlation in your quan data and set out to look for traces of this in your qual data. Or you might find a disturbing contradiction between what your qual and quan data suggest—and try to dig deeper in order to resolve the issue.

Second, construct *valid elements of the mechanism.* The causal mechanism or narrative that researchers use in order to explain their explanandum consists of typical actors-in-situations. The different data types that mixed methods provides will often allow them better to construct an as-simple-as-possible-as-complex-as-necessary model of who the main typical actors are, what their typical situations are, how and according to what selection rules they would typically act and what possible (intentional or unintentional) effects might typically follow. Often, researchers will create families of typical actors or cases with similar causal narratives. For example, in a large mixed methods study on members of evangelical free churches, Stolz et al. (2014) constructed models of three types of evangelical actors (classical, charismatic, conservative) and showed how these types behaved differently in different situations. On the basis of these typical actors, many of the aggregated phenomena in the evangelical milieu then become explicable.

Third, using the elements from step two, create a valid *overall mechanism or mechanism narrative* that explains the outcome. This often involves identifying central background parameters, showing how the different elements are interlinked over time and creating a chain of evidence that is as complete as possible. If the researcher tries to explain one case, this is often called "process-tracing"

[27] Bennett uses the detective metaphor in order to describe process tracing, but in my view it applies just as well to explanatory triangulation in general.

(Bennett 2010) or "causal-process-observation" (Brady 2010). But equally, when researchers face several or a large number of cases, the model has to show what typical causal processes lead typically to the explanandum. We look, as Abbott (1992, 68) writes, for "narrative generalizations across cases". The general point here is that analytic mixed methods research does not just use qualitative data and a case perspective, but all kinds of data, qual and quan, and switches continually between variable and case perspective in order to get to the final model. At one point, researchers have to settle for a final as-simple-as-possible-as-complex-as-necessary model in terms of typical actors-in-situations.

Fourth, *evaluate the model.* Two criteria seem to be important. For one thing, the model has to be internally consistent (King et al. 1994, 105) and complete. The model may not use axioms or lead to hypotheses that contradict each other. Also, it has to be so precise as to give an uninterrupted causal narrative from the causes to the effects. An excellent way of judging internal consistency and completeness is formal modelling and the technique of Agent-Based-Modelling. These tools force researchers to be very precise in their assumptions and bring inconsistencies to light that had been hidden under the imprecise structure of language. Agent-Based-Modelling also lets us simulate the mechanism and judge if a model as assumed by the theory could at least in principle bring about the phenomenon to be explained (Macy/Willer 2002; Manzo 2007, 49ff.). For another thing, the model should explain as much as possible with as little as possible. This may be judged in the quan data (how much variance explained) and in the qual data (how many different phenomena explained)? An important criterion is if the model explains new observable facts, hitherto unexplained (Lakatos 1978). Another important point in this respect is whether the model refrains from using easy fixes, that is, includes all kinds of additional assumptions that make the model fit.[28] While I have argued that some of these assumptions are inevitable (bridge hypotheses), researchers should use them sparingly.

5. Examples of Opened Black Boxes

What might an analytical sociology using explanatory mixed methods look like in practice? I give three examples that each highlights a different aspect of the benefits.

5.1 Who survived on the Titanic?

In the arguably most famous peacetime maritime disaster of all times, on April 14, 1912, the then-largest ship afloat, the Titanic, collided with an iceberg and sank. Due to a lack of lifeboats, about 2/3 of the people on board died (precisely: 1517 of 2223 passengers and crew) (Lord 2012[1956]). As is well known, in the accident women had a much higher chance of surviving than men, and first-class passengers had higher chances of survival than second-class passengers who in

[28] This is a critique addressed at qualitative methods such as "grounded theory" by scholars in the quantitative tradition (Goldthorpe 2000b).

turn survived with more likelihood than third-class passengers, as can be seen in *figure 2*.

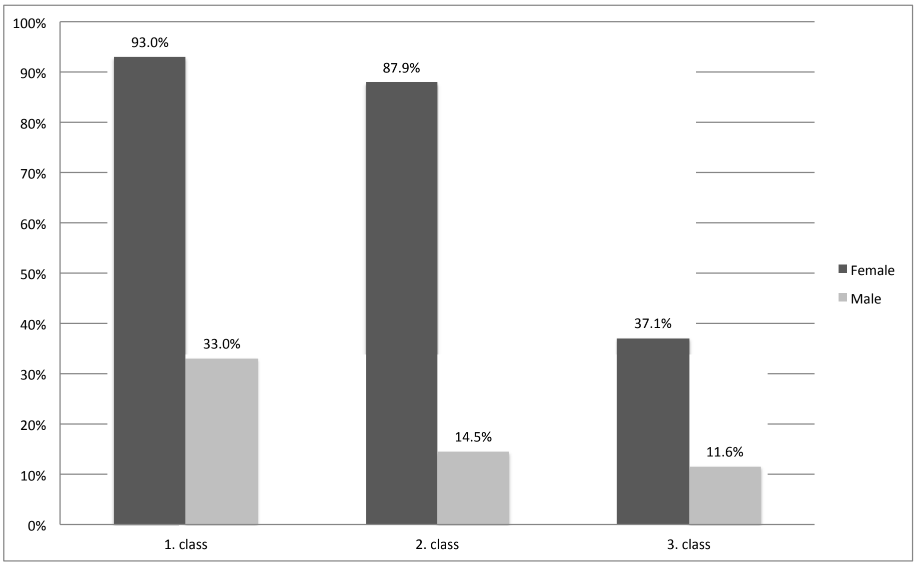

Figure 2: Probability of surviving on the Titanic (passengers only) depending on sex and class (source: own figure based on the Titanic dataset that can be downloaded at: biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3.xls. The drawback of this dataset is that crew are not included.

In a mono-method study, Frey et al. (2011) present a study—perfectly acceptable according to mainstream quantitative standards—on what individual and social dynamics led to death or survival on the Titanic. The authors use a rectangular data set, interpret the existing variables in the sense of a number of individual and social factors, and estimate various multiple regressions. They come up with the main result that social norms were a key determinant in this extreme situation of the sinking of the Titanic and set this off from a purely rational choice view that might have predicted a fight for life and death. While the quantitative analysis is sound, anyone who is fairly acquainted with the details of the Titanic case cannot help but be somewhat disappointed. Clearly, the authors do not get at the real mechanisms that were at work. While they point to some context factors in their footnotes, they do not grasp, for example, the extreme importance of policy and policy enforcement by the crew.[29]

[29] This is not an attempt to debunk Frey or his co-authors. Rather, I like this example because it shows that the results of an article that is of high quality in mainstream quantitative terms nevertheless can be considerably improved by reverting to mixed methods.

It is therefore interesting to compare the Frey et al. (2011) paper to a paper by Hall (1986)[30] who does not claim to do a mixed methods study—but in fact combines his (much less developed) statistical data analysis with an in-depth use of qualitative data including survivor accounts, the spatial arrangements on the ship, and the temporal order of happenings. With these different data, Hall is able to give us a much more convincing account of the causal mechanisms at work that led to the effect that women and higher class passengers survived more often. Compared with the Frey article, Hall shows that:

- The sex differentials were a result of *policy*. It is thus because there was an (armed) crew that supervised the filling of the boats and that used the norm 'Women and children first' that this group survived with much higher probability. According to Hall, not all first-class women were saved since some wanted to stay with their husbands and some did not believe that the ship could actually sink. Also, there seems to have been a difference in the strictness of enforcement of the rule 'Women and children first' on the two sides of the Titanic, which meant that on the one side men had a slightly higher chance of getting on the boats.
- The class differentials were due to a number of factors: "(1) the *positioning of the lifeboats* on the deck where first and second class passengers were located; (2) a *policy* of looking after the first and second class passengers first; (3) *neglect* of third-class passengers who were left to fend for themselves, and who could only find their way to the boat deck by trial and error; and (4) some *unsystematic exclusion* of third-class passengers from the boat deck by members of the crew." (Hall 1986, 9)

The Hall explanation is much more satisfying, since it tries to show, on the basis of varied pieces of evidence, both qual and quan, how different classes of typical actors have acted in typical situations and have thereby brought about the sex and class differentials. For Hall this means bringing elements into the picture that are not included in the rectangular data set that Frey et al. exclusively analyze: the policy enacted by the crew, the spatial arrangement of the ship, the beliefs of passengers that the ship was unsinkable, the decision of some first-class women to remain and die with their husbands rather than enter the lifeboats—all this is qual evidence that can be gleaned with high validity from the survivor accounts and all of this is highly relevant for a mechanism explanation, but neither can it be hypothesized on the basis of our everyday knowledge, nor can it be read off from a multiple regression analysis, as is exemplified in the Frey et al. paper.

5.2 Why Do Women in the Hairdresser Profession Stop Working Earlier Than Women Business Employees?

In a mixed methods study on the work- and family-biographies of women in five different professions, the question was asked just what professional changes oc-

[30] Frey et al. were apparently unaware of the Hall paper.

curred in these biographies and how they could be accounted for (Erzberger 1998; Kelle 2007, 258). Researchers used a standardized questionnaire that was sent to the sampled women by post (N = 220); a subsample of women was interviewed with in-depth interviews (N = 52), and at a later point a number of spouses (N = 37) of women in the subsample were again interviewed with in-depth methodology. From the quantitative analysis, it became clear that business employees worked (independently from the number of children) on average longer than the other professional groups in their first profession; dressmakers and shop assistants more often switched to other jobs; while hairdressers significantly more often left their first profession without returning to any job (Kelle 2007, 259). According to Kelle, this fact can be explained, on the one hand, by the differing situations on the labour market in the 1950s and 1960s and, on the other hand, by the differences in the compatibility between the job demands and family life. What is interesting, however, is that these significant causal mechanisms did *not* show up at all in the qualitative material. According to both women and men, the decisions depended on negotiations in the couple and the job of the male partner. The fact that the professional group of the woman gave her relative advantages or disadvantages in comparison to women in other professional groups remained latent, since respondents did not normally have the relevant information to make these comparisons. Researchers could, however, use the quantitative findings to re-analyze the qualitative material and to show that the differences in the labour market opportunities of different professional groups affected the differing labour market involvement of women *through the couple's negotiations*. In fact, only the women who could show that their work would significantly improve the household budget and that their job would not have negative effects on their work in the household succeeded in getting their paid work accepted by their husband.

This example nicely shows how *both kinds of data had their blind spots or black box problems*. The qualitative data did not make the causal importance of different professional groups apparent whose causal effect works behind the backs of the actors; the quantitative data did not show the importance of intra-couple negotiations. The combination of the data sets allowed researchers to open the black box and construct a mechanism that led from differing job opportunities of different professional groups to intra-couple negotiations (where women have different negotiation power depending on their professional group), this in turn influencing the actual labour market involvement of the women.

5.3 How Many Votes Did George Bush Lose in the 2000 Election in Florida?

In the 2000 presidential election in the USA, media networks wrongly declared Al Gore the winner in Florida after the polls had closed in some Florida counties, but before the polls had closed in a number of others (the Florida Panhandle counties that were in a different time zone). According to a quantitative analysis by John Lott (2000), this led to a loss of votes for George Bush, amounting to 10,000 votes. The analysis by Lott

> "[...] employed a 'difference-in-differences' form of regression analysis, based on data-set observations. He obtained turnout data on all sixty-seven Florida counties for four presidential elections (1988, 1992, 1996, and 2000), and he estimated a time-series cross-sectional regression with fixed county and time effects and with a 'dummy variable' for the ten Panhandle counties. In effect, Lott looked at the difference between one set of counties that got a 'treatment' in the year 2000 (the ten Panhandle counties whose polls were still open when the election was 'called') and those that did not (the remaining fifty-seven Florida counties in the eastern time zone), while controlling for differences reflected in the data from previous elections." (Brady 2010, 238)

In a critical reaction to Lott, Brady (2010)—working in an analytic fashion—comes to the conclusion that Lott's estimation of Bush's vote loss of 10,000 is dramatically overstated, that the "approximate upper bound for Bush's vote loss was 224 and that the actual vote loss was probably closer to somewhere between 28 and 56 votes".[31] Methodically, Brady gathers—much like a detective—a number of qualitative and quantitative observations, putting them into the correct causal and temporal order, which precisely describe the actual causal process taking place. He thus seeks exact information on when exactly the media called the election: he estimates how many voters might not yet have voted; how many might have heard the information; and how many might have wanted to vote for Bush—but then did not. Doing this, he finds that the media call happened only ten minutes before the polls closed in the ten Panhandle counties of Florida, and that of the estimated 4,200 people that might not yet have voted only (in a conservative estimate) 20%, that is 840, might have heard the information. Of these, only a certain percentage would have been Bush voters (Brady estimates 560) and of these, not all would have not voted just because they would have heard the call. In fact, Brady estimates that probably only 10% of voters having heard the information would have been deterred, leaving him with 56 Bush voters that might have not voted for Bush because of the media information.

The example is interesting because the combination of qualitative and quantitative information does not so much lead us better to understand the mechanism (which seems to be rather straightforward). Rather, it helps us to get at the precise contextual conditions of the process—which leads us to conclude that the Lott estimate must be completely wrong.

6. Conclusion

In this paper I have argued that analytical sociology could strongly benefit from using explanatory mixed methods. Given that social reality is changing fast, characterized by strong diversity and complexified by the phenomenon of cultural

[31] For a critical exchange concerning this example see Beck 2010 and Collier/Brady/Seawright 2010.

meanings, mixed methods are tools that allow analytical sociologists to get closer to their object of research and therefore to have the opportunity to create more valid mechanism explanations. I have given five interlinked rules that together form an explanatory mixed methods research program for analytical sociology and have pointed to three examples of successful explanations that were able to identify the actual causal mechanism through combining various data types.

Instead of summarizing the points and examples again, let us just take a brief look at *how the abstract rules given in the article are borne out in the concrete examples*. In all three examples, one single central question gives focus to the research design, the theorizing, the quan and qual data collection as well as the data analysis. Why did women and higher class passengers survive more easily on the Titanic? Why did women in the hairdresser profession stop working earlier than women business employees in Germany in the 1950s and 1960s? How many votes did George Bush lose in the 2000 election when Al Gore was falsely declared the winner in states that stopped voting earlier? These questions are treated from a realist philosophical background and with the help of one logic of inference; that is, the quan and qual data are used in order to better understand what really happened in the three cases of interest. In every one of the examples the central question is answered by pointing to a semi-general causal mechanism in conjunction with contextual parameters that produces the outcome. In the Titanic example, it is a filling mechanism with differential filling rules that works together with contextual parameters such as the number of lifeboats, the size of the gender and class groups on the boat, the spatial arrangement of the ship, the existence of a crew in charge of the operation etc. In the women hairdresser/employee example, it is a negotiation mechanism with differential negotiation power (women in professions) that works in conjunction with one central contextual parameter: the labour market situation in the 1950s/60s in Germany for the three different professions. In the 2000 Bush-Gore election, it is an information-voting mechanism that is combined with contextual parameters such as the time when the false information about the alleged victory of Gore was given, the time left to vote, the estimate of voters that had not yet voted, might have heard the information and had wanted to vote, etc. Finally, in all three examples, the explanation given is substantiated by the combined analysis—triangulation—of both qual and quan data: a rectangular data set, survivor accounts, and detailed plans of the boat (Titanic), a standardized survey and in-depth interviews (hairdressers/employees), qual and quan contextual information (Bush-Gore 2000 election). Looking at the examples, it is quite obvious that the use of only one data-set could not possibly have created an explanation as valid as the ones given—and, as we have seen, in fact in two cases such mono-method studies exist and are clearly found to be wanting. For example, the mono-method study by Frey et al. (2011) did not take survivor accounts and other qualitative contextual information into account and thus clearly misrepresents the actual causal mechanism (e.g. by not acknowledging the importance of the crew on the Titanic).

Although the advantage of using explanatory mixed methods in these examples may be uncontroversial, it is important to note that analytical sociology using explanatory mixed methods is only at the beginning. I see three domains where important *next steps are needed*. First, there is a need of exemplars of analytical explanations that use explanatory mixed methods. While I have pointed to some examples and others could be cited, we do not yet have a number of first-rate studies that could work as models for a whole research tradition. Second, there are some conceptual issues in explanatory mixed methods studies that have to be better understood and best practices have to be fixed; we have alluded to some of these issues in the point on validity (4.5). Third, there is a dire need for powerful computer programs that can better handle explanatory mixed methods data sets. Fortunately, such developments are currently underway (Kuckartz 2014).

"Never mix, never worry", says Martha in *Who's afraid of Virginia Woolf.* While this advice may indeed be commendable with respect to alcoholic beverages,[32] I have argued for a different stance in social science methodology: here, mixing methods may indeed be the way to go in order to arrive at better mechanism explanations.

Bibliography

Abbott, A. (1992), What Do Cases Do? Some Notes on Activity in Sociological Analysis, in: Ragin, C. C./H. S. Becker (eds.), *What Is a Case? Exploring the Foundations of Social Inquiry*, Cambridge, 53–82

Beck, N. (2010), Causal Process 'Observation': Oxymoron or (Fine) Old Wine, in: *Political Analysis* 18, 499–505

Bennett, A. (2010), Process Tracing and Causal Inference, in: Brady, H. E./D. Collier (eds.), *Rethinking Social Inquiry. Diverse Tools, Shared Standards. Second Edition*, Lanham, 207–219

Bergman, M. M. (2008), The Straw Men of the Qualitative-Quantitative Divide and their Influence on Mixed Methods Research, in: Bergman, M. M. (ed.), *Advances in Mixed Methods Research*, Los Angeles, 11–21

— (2011), The Good, the Bad, and the Ugly in Mixed Methods Research and Design, in: *Journal of Mixed Methods Research* 5, 271–275

Bergman, M. M. E. (2008), *Advances in Mixed Methods Research*, Los Angeles

Bernard, H. R. (2000), *Social Research Methods: Qualitative and Quantitative Approaches*, Thousand Oaks:

Blalock, H. M. (1984), *Basic Dilemmas in the Social Sciences*, Beverly Hills

Boudon, R. (1971), *Effets pervers et ordre social*, Paris

— (1974), *The Logic of Sociological Explanation*, Suffolk

— (1979[1973]), *L'inégalité des chances. La mobilité dans les sociétés industrielles*, Paris

Bourdieu, P./J.-C. Passeron (1985), *Les héritiers. Les étudiants et la culture (3. édition)*, Paris

Brady, H. E. (2010), Data-Set Observations versus Causal-Process Observations: The 2000 U.S. Presidential Election, in: Brady, H. E./D. Collier (eds.), *Rethinking Social Inquiry. Diverse Tools, Shared Standards. Second Editiion*, Lanham, 237–242

[32] And be preferable to alternative rules such as 'Liquor to beer, have no fear', 'Beer to wine, you're doing fine' or 'Beer, wine and gin, take it all in'.

—/D. Collier (2010) (eds.), *Rethinking Social Inquiry. Diverse Tools, Shared Standards. Second Editiion*, Lanham
Brante, T. (2001), Consequences of Realism for Sociological Theory-Building, in: *Journal for the Theory of Social Behaviour* 31, 167–195
Bryman, A. (1988), *Quantity and Quality in Social Research*, London
— (2004), *Social Research Methods. Second Edition*, New York
— (2008), Why Do Researchers Integrate/Combine/Mesh/Blend/Mix/Merge/Fuse Quantitative and Qualitative Research?, in: Bergman, M. M. E. (ed.), *Advances in Mixed Methods Research. Theories and Applications*, Los Angeles, 87–100
Bunge, M. (2004), How Does It Work? The Search for Explanatory Mechanisms, in: *Philosophy of the Social Sciences* 34, 182–210
Campbell, D. T. (1988[1974]), Quasi-Experimental Designs, in: *Methodology and Epistemology for Social Science. Selected Papers*, Chicago 191–221
Cherlin, A. J./L. M. Burton/T. R. Hurt/D. M. Purvin (2004), The Influence of Physical and Sexual Abuse on Marriage and Cohabitation, in: *American Sociological Review* 69, 768–789
Collier, D./H. E. Brady (2010), Outdated Views of Qualitative Methods: Time to Move On, in: *Political Analysis* 18, 506–513
Creswell, J. W. (1998), *Qualitative Inquiry and Research Design: Choosing among Five Traditions*, Thousand Oaks
—/V. L. Plano Clark (2007), *Designing and Conducting Mixed Methods Research*, Thousand Oaks
Creswell, J. W./V. L. Plano Clark/M. L. Gutmann/W. E. Hanson (2003), Advanced Mixed Methods Research Designs, in: Tashakkori, A./C. Teddlie (eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, Thousand Oaks, 209–240
Denzin, N. K. (1978), *The Research Act. A Theoretical Introduction to Sociological Methods*, New York
Elster, J. (1989), *Nuts and Bolts for the Social Sciences*, Cambridge
Epstein, J. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton
Erzberger, C. (1998). *Zahlen und Wörter – Die Verbindung quantitativer und qualitativer Daten und Methoden im Forschungsprozess*, Weinheim
Esser, H. (1996), What Is Wrong with "Variable Sociology"? In: *European Sociological Review* 12, 159–166
— (1998), Why Are Bridge Hypotheses Necessary? In: Blossfeld, H.-P./G. Prein (eds.), *Rational Choice Theory and Large-Scale Data Analysis*, Colorado, 94–111
Flick, U. (2002), *Qualitative Sozialforschung. Eine Einführung*, Reinbek
Freedman, D. A. (2010), On Types of Scientific Inquiry: The Role of Qualitative Reasoning, in: Brady, H. E./D. Collier (eds.), *Rethinking Social Inquiry. Diverse Tools, Shared Standards*, second Edition, Lanham, 221–236
Frey, B. S./D. A. Savage/B. Torgler (2011), Who Perished on the Titanic? The Importance of Social Norms, in: *Rationality & Society* 23, 35–49
Goldthorpe, J. H. (1996), The Quantitative Analysis of Large-Scale Data-sets and Rational Action Theory: For a Sociological Alliance, in: *European Sociological Review* 12, 109–126
— (1997), Current Issues in Comparative Macrosociology. A Debate on Methodological Issues, in: *Comparative Social Research* 16, 1–26
— (2000a), Causation, Statistics, and Sociology, in: idem (ed.), *On Sociology. Numbers, Narratives, and the Integration of Research and Theory*, Oxford, 137–160

— (2000b), Sociological Ethnography Today: Problems and Possibilities, in: *On Sociology. Numbers, Narratives, and the Integration of Research and Theory*, Oxford, 65–93

Greene, J. C./V. J. Caracelli/W. F. Graham (1989), Toward a Conceptual Framework for Mixed-Method Evaluation Designs, in: *Educational Evaluation and Policy Analysis* 11, 255–274

Haack, S. (2006), Introduction: Pragmatism, Old and New, in: Haack, S. (ed.), *Pragmatism, Old & New. Selected Writings*, Amherst, 15–68

Hacking, I. (1999), *The Social Construction of What?*, Cambridge

Hägerstrand, T. (1965), A Monte Carlo Approach to Diffusion, in: *Archives européennes de Sociologie* 6, 43–57

Hall, W. (1986), Social Class and Survival on the S. S. Titanic, in: *Social Science & Medicine* 22, 687–690

Hammersley, M. (2008), Troubles with Triangulation, in: Bergman, M. M. (ed.), *Advances in Mixed Methods Research*, Los Angeles, 22–36

Hedström, P. (2005), *Dissecting the Social. On the Principles of Analytical Sociology*, Cambridge

—/P. Bearman (2009a), *The Oxford Handbook of Analytical Sociology*, Oxford

—/— (2009b), What Is Analytical Sociology All about? An Introductory Essay, in: Hedström, P./P. Bearman (eds.), *The Oxford Handbook of Analytical Sociology*, Oxford, 3–24

—/R. Swedberg (1998) (eds.), *Social Mechanisms. An Analytical Approach to Social Theory*, Cambridge

Johnson, R. B./A. J. Onwuegbuzie/L. A. Turner (2007), Toward a Definition of Mixed Methods Research, in: *Journal of Mixed Methods Research* 1, 112–133

Kalter, F./C. Kroneberg (2014), Between Mechanism Talk and Mechanism Cult: New Emphases in Explanatory Sociology and Empirical Research, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Special Issue 54: Social Contexts and Social Mechanisms*, 91–115

Kelle, U. (2001), Sociological Explanations between Micro and Macro and the Integration of Qualitative and Quantitative Methods, in: *Forum: Qualitative Social Research* 2, URL: http://qualitative-research.net/fqs/fqs-eng.htm [04/10/2008]

— (2007), *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte*, Wiesbaden

—/C. Lüdemann (1998), Bridge Assumptions in Rational Choice Theory: Methodological Problems and Possible Solutions, in: Blossfeld, H.-P./G. Prein (eds.), *Rational Choice Theory and Large-Scale Data Analysis*, Colorado, 112–125

Kemper, E. A./S. Stringfield/C. Teddlie (2003), Mixed Methods Sampling Strategies in Social Science Research, in: Tashakkori, A./C. Teddlie (eds.), *Handbook of Mixed Methods in Social & Behavioral Research*, Thousand Oaks, 273–296

King, G./R. O. Keohane/S. Verba (1994), *Designing Social Inquiry. Scientific Inference in Qualitative Research*, Princeton

Koenig, M. (2006), Soziale Mechanismen und relationale Soziologie, in: Rehberg, K.-S. (ed.), *Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006. Teilbd. 1 u. 2.*, Frankfurt, 2869–2906

Kuckartz, U. (2014), *Mixed Methods. Methodologie, Forschungsdesigns und Analyseverfahren*, Wiesbaden

Lakatos, I. (1978), *The Methodology of Scientific Research Programmes. Philosophical Papers Vol. 1*, Cambridge

Laubach, M. (2005), Consent, Informal Organization and Job Rewards: A Mixed Methods Analysis, in: *Social Forces* 83, 1535–1565

Lindenberg, S. (1996), Die Relevanz theoriereicher Brückenannahmen, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 48, 126–140

Lord, W. (2012[1956]), *A Night to Remember*, London

Lott, J. R. J. (2000), Gore Might Lose a Second Round: Media Suppressed the Bush Vote, in: *Philadelphia Inquirer, Tuesday, 14 November*, 23A.

Machamer, P./L. Darden/C. F. Craver (2000), Thinking about Mechanisms, in: *Philosophy of Science* 67, 1–25

Macy, M./R. Willer (2002), From Factors to Actors: Computational Sociology and Agent-Based Modeling, in: *Annual Review of Sociology* 28, 143–166

Manzo, G. (2007), Variables, Mechanisms, and Simulations: Can the Three Methods Be Synthesized? A Critical Analysis of the Literature, in: *Revue française de sociologie* 48, 35–71

— (2010), Analytical Sociology and Its Critics, in: *European Journal of Sociology* 51, 129–170

Maxwell, J. A. (2005), *Qualitative Research Design. An Interactive Approach. Second Edition. Applied Social Research Methods Series. Volume 41*, Thousand Oaks

— (2010), Realism as a Stance for Mixed Methods Research, in: Tashakkori, A./C. Teddlie (eds.), *Sage Handbook of Mixed Methods in Social and Behavioral Research*, 2nd ed., Thousand Oaks, 145–167

Mayntz, R./B. Nedelmann (1987), Eigendynamische soziale Prozesse. Anmerkungen zu einem analytischen Paradigma, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 39, 648–668

McKeown, T. J. (1999), Case Studies and the Statistical Worldview: Review of King, Keohane, and Verb's *Designing Social Inquiry: Scientific Inference in Qualitative Research*, in: *International Organization* 53, 161–19

Meulemann, H. (2002), Die empirische Sozialforschung in der soziologischen Ausbildung und der Berufspraxis von Soziologen, in: Engel, U. (ed.), *Praxisrelevanz der Methodenausbildung. Sozialwissenschaftliche Tagungsberichte* 5, 45–67

Miles, M. B./A. M. Huberman (1994), *Qualitative Data Analysis. An Expanded Sourcebook*, Thousand Oaks

Moss, S./B. Edmonds (2005), Sociology and Simulation: Statistical and Qualitative Cross-Validation, in: *Amercan Journal of Sociology* 110, 1095–1131

Onwuegbuzie, A. J./C. Teddlie (2003), A Framework for Analyzing Data in Mixed Methods Research, in: Tashakkori, A./C. Teddlie (eds.), *Handbook of Mixed Methods in Social & Behavioural Research*, Thousand Oakes, 351–383

Opp, H.-D. (2005), Explanations by Mechanisms in the Social Sciences. Problems, Advantages and Alternatives, in: *Mind and Society* 4, 163–178

Peirce, C. S. (2006), Some Consequences of Four Incapacities, in: Haack, S. (ed.), *Pragmatism, Old & New. Selected Writings*, Amherst, 69–106

Popper, K. R. (1980[1959]), *The Logic of Scientific Discovery*, London

Rosa, H. (2013), *Social Acceleration: A New Theory of Modernity*, New York

Schelling, T. C. (2006[1978]), *Micromotives and Macrobehavior*, New York

Shadish, W. R./T. D. Cook/D. T. Campbell (2002), *Experimental and Quasi-experimental Designs for Generalized Causal Inference*, Boston

Stolz, J./O. Favre/C. Gachet/E. Buchard (2014), *Phänomen Freikirchen. Analysen eines wettbewerbsstarken Milieus*, Zürich

—/J. Könemann/M. Schneuwly Purdie/T. Englberger/M. Krüggeler] (2015), *(Un-) Believing in Modern Society. Religion, Spirituality, and Religious-secular Competition*, London

Tashakkori, A./C. Teddlie (1998), *Mixed Methodology. Combining Qualitative and Quantitative Approaches*, Thousand Oaks

—/— (2010), *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, Thousand Oaks

Teddlie, C./F. Yu (2008), Mixed Methods Sampling: A Typology with Examples, in: Plano Clark, V. L./J. W. Creswell (eds.), *The Mixed Methods Reader*, Thousand Oaks, 197–228

Tubaro, P./A. A. Casilli (2010), 'An Ethnographic Seduction': How Qualitative Research and Agent-based Models Can Benefit Each Other, in: *Bulletin of Sociological Methodology/Bulletin de Möthodologie Sociologique* 106, 59–74

von der Lippe, H./G. Mey/J. Frommer (2011), Zur Frage der Integration qualitativer und quantitativer Forschung in der Psychologie: Eine Einführung und ein Diskussionsbeitrag, in: *Zeitschrift für qualitative Forschung* 1, 3–25

Yang, L./N. Gilbert (2008), Getting away from Numbers: Using Qualitative Observation for Agent-Based Modelling, in: *Advances in Complex Systems* 11, 1–11