Theory Building in Qualitative Research: Reconsidering the Problem of Induction

Pedro F. Bendassolli

Abstract: The problem of induction refers to the difficulties involved in the process of justifying experience-based scientific conclusions. More specifically, inductive reasoning assumes a leap from singular observational statements to general theoretical statements. It calls into question the role of empirical evidence in the theory-building process. In the philosophy of science, the validity of inductive reasoning has been severely questioned since at least the writings of David HUME. At the same time, induction has been lauded as one of the main pillars of qualitative research methods, and its identity as such has consolidated to the detriment of hypothetical-deductive methods. This article proposes reviving discussion on the problem of induction in qualitative research. It is argued that qualitative methods inherit many of the tensions intrinsic to inductive reasoning, such as those between the demands of empiricism and of formal scientific explanation, suggesting the need to reconsider the role of theory in qualitative research.

Table of Contents

1. Introduction
2. The Problem of Induction
3. Relationship Between Theory and Empirical Data
4. Induction and Theory in Qualitative Research
   4.1 The generic analytic cycle
   4.2 Situating the problem of induction in the current debate: Some unsolved questions
5. Suggestions for Reconsidering the Problem of Induction in Qualitative Research
6. Final Considerations
   6.1 General overview and limitations
   6.2 Contributions to scholarship: Revisiting theory building in qualitative research

References
Author
Citation

1. Introduction

One of the major claims made regarding qualitative methods is that they diverge from scientific explanation models in terms of the need for hypothesis testing. A scientific hypothesis is based on a background theory, typically assuming the form of a proposition whose validity depends on empirical confirmation. Otherwise, a hypothesis is nothing but an imaginative conjecture. Moreover, when researchers do not obtain empirical confirmation for their hypothesis, the theory in question (or part of it) may not be able to predict relevant aspects of the phenomenon under investigation. [1]
By contrast, qualitative researchers contend that their work does not consist of proposing and testing hypotheses. Their primary interest is to achieve understanding (Verstehen) of a particular situation, or individuals, or groups of individual, or (sub)cultures, etc., rather than to explain and predict future behaviors as in the so-called hard sciences, with their arsenal of laws, theories, and hypotheses employed or rejected on the basis of their predictive value. In summary, qualitative methods are primarily inductive, in contrast to the deductive methods of experimental science. [2]

The question of induction is one of the most serious issues in the philosophy of science, one that dates back to the ancient Greek philosophers, particularly ARISTOTLE (LOSEE, 2001). The debate centers around how we justify that what we know is valid. More specifically, induction is the form of reasoning based on empirical observation in the process of developing scientific laws and theories. Thus, induction negotiates the relationship between empirical reality and its theorization, in addition to the production and validation of knowledge. [3]

Induction has also had repercussions in various qualitative method domains. For example, qualitative methods have been accused of reflecting the problems pointed out by philosophers of science (e.g., POPPER, 1959), in particular that of hyper-valuing observational statements compared to their theoretical counterparts. In other words, qualitative researchers tend to prioritize logic emerging from experience, preferring to expand their knowledge from it as opposed to using a priori, deductive, concepts. Qualitative researchers have for decades reacted to this distorted view of the field (e.g., STRAUSS, 1987). [4]

The problem of induction, therefore, is nothing new to qualitative researchers, who have developed a range of strategies to overcome or at least address it. Of the many examples that could be cited, I highlight grounded theory methodology (GTM). There are differences among researchers using this approach (e.g., GLASER, 1978, 1992; STRAUSS, 1970, 1987); however, in general GTM is a hybrid method, combining induction and deduction in the theory-building process. GTM rests in a state of permanent tension between 1. the risk of "forcing" data into previous conceptual categories, that is, not being inductive enough; and 2, producing such a large volume of codes for empirical material that it hinders the categorization and theoretical development process, that is, not being deductive enough (BRYANT & CHARMAZ, 2007; KELLE, 2005). [5]

Despite attempts to address the problem of induction, as in GTM, qualitative researchers continue to be questioned about the relationship between observational and theoretical statements. What is the role of theory in qualitative research? Alternatively, what function do empirical data play in the theorizing process? Answering these questions is important for the continuing advancement of qualitative methods as well as the inclusion of this field in the discussions of similar issues that have been witnessed in the philosophy of science. [6]

In this article, my proposal is to consider the relationship between theory and empirical data based on a dialogue between the philosophy of science and
qualitative research. As a starting point, I recapitulate the main characteristics of the so-called problem of induction, arguing that it raises important questions regarding the value of theory in science. Next, I review ways of describing the theory-empirical data relationship that have been proposed in order to address the problem of induction in the realm of the philosophy of science. Against this backdrop, I discuss how qualitative researchers have dealt with the question of induction, using a "generic analytic cycle" common to qualitative methods as an illustration. In the last sections, I propose reconsidering the role of theory in qualitative research. I argue for the need to recover a substantial definition of theory in these studies. [7]

2. The Problem of Induction

The problem of induction, also known as "Hume's problem" (KANT, 2004 [1783], §§27-30), refers to the process of justifying knowledge. According to HUME (1974 [1748]), there are two primary ways to validate knowledge: by logic, as in the relation of ideas (for example, in mathematics), and by experience, in the case of matters of fact. Knowing facts is equivalent to identifying their causes and effects. However, observing facts, describing them in their manifestation, does not amount to science. There must be a leap from the visible to the invisible, and herein lies induction: knowledge building evolves from single facts to a general belief regarding their causes. The inductive leap allows us, based on singular facts, to create statements about sets of facts and their future behavior. [8]

But what sustains the argument about induction? What permits us to go from a singular fact to a statement about facts in general or future facts? According to HUME (1974 [1748]), induction does not involve a logical base. The "statement about all" is not contained in the "statement about some." The problem of induction, in this sense, is that there is no logical connection between statements, but rather an empirical connection based on repetition of experience. HUME claims that it is merely habit that causes us to think that if the sun rose today, it will do so once again tomorrow. There is therefore a psychological component in this knowledge-building process. In other words, HUME demonstrated that passing from some to all is an emotionally and imaginatively based process, and that the root of any knowledge is sensory experience. [9]

Inductive thinking is problematic because we can never be certain that a recurring (known) event will continue to occur. The past may not be the best guarantee for current knowledge; otherwise, how can we explain unpredictable events? In the well-known analogy cited by POPPER (1959), the fact that we observe innumerable white swans does not allow us to assume that there will never be a black one. Another relevant question is distinguishing between empirical generalizations, based on the observation of a recurring number of singular cases, and universal generalizations, in the form of laws. Without resorting to metaphysics, how do we attest to the truth of universal laws, which establish necessary (non-accidental) connections between events, based on observations of singular cases only (QUINE, 1975, p. 317, calls them "pegged observational sentences")? According to the skeptic HUME, all what we can do is create

© 2013 FQS http://www.qualitative-research.net/
hypotheses about how things (should) occur, drawing from our own empirical experiences or habits; we can never determine the ultimate fundamentals of the phenomena. [10]

HUME's position generated intense debate in the philosophy of science. One of these arguments is put forth by POPPER (1959). Like HUME, POPPER denies the possibility of logically justifying induction, since we have no way of guaranteeing statements based on our past or unknown experience. However, POPPER does not endorse HUME's irrationality. This irrationality is based in HUME's opinion that our beliefs have more weight than rationality does in making up our understanding. POPPER provides us with the tools for rational criticism of naïve inductivism. Naïve inductivists, according to CHALMERS (1999), believe in the origin of knowledge based on theoretically free empirical observation. They argue that a large number of observations, obtained experimentally over a wide range of circumstances, allow inference from the empirical (particular) to the theoretical (universal). Knowledge, they assert, can be constructed on the basis of repeated observations, to the point where no observational statements conflict with the law or theory thereby derived, or up to an established saturation point. [11]

POPPER (1959) diverges from naive inductivism, proposing a redefinition of the role of theory in science. He purports that if there is no logical support to infer a universal law from singular experience, there must be support for the opposite. That is, we can legitimately allege that a theory is true or false based on singular observational statements. Thus, the order is inverted: the passive "emergentist" position is replaced by an active one, in which theory enables us to conjecture about how things should function. There is no observation without theory, since perception itself is influenced by expectations, previous experiences, and accumulated knowledge. At the same time, theoretical assertions without empirical content do not tell us much about the world. Theory must be confirmed or falsified by experience. From this emerges the well-known hypothetical-deductive method. POPPER proposes a jump directly to conclusions, instead of focusing on the development of premises. The empirical world is supposed to determine if such a conclusion is confirmed (true) or pure speculation. [12]

POPPER's position has also been criticized. For example, LAKATOS (1970, 1978) states that a theory consists of a complex of universal statements (embedded in particular research programs), rather than a single statement, like a hypothesis, that can be tested straightforwardly. This calls into question the value of the falsifiability of discrete hypotheses. Moreover, QUINE (1951, 1975, 1978, 1998) proposes that we conceive theories holistically, as a web of interlocked statements, such that concepts can only be defined in terms of other concepts that make up the network and confer meaning on them, as well as relate them to experience. As a result of these criticisms, it is concluded that the value of theories is not restricted to allowing the elaboration of hypotheses to be individually tested; they are essential to explain the phenomena to be investigated. So, the primary focus of researchers should not be on data, but rather on the phenomenon, which is embedded into a given theoretical web. [13]
In the next section, I present a number of philosophical perspectives on the relationship between theory and empirical data in order to widen the discussion regarding ways of addressing the problem of induction in science in general and qualitative research in particular. [14]

3. Relationship Between Theory and Empirical Data

One of the most widely prevalent ways of thinking about the theory-data relationship is that the latter verify the former. This viewpoint is associated with the philosophy of logical positivism, which introduces a distinction between direct observation (which is not theory-laden), and theory, whose value depends on the justification allowed by empirical data. Thus, theoretical statements should have empirical content, if they are to be trusted as claims about the world. The truth about a theoretical statement depends on a "correspondence theory" of truth: referents for these statements are found in objective facts available in the world. Positivists vehemently reject any pretense of metaphysical justification for scientific activity, arguing for the impossibility of synthetic propositions, that is, non-contingent statements. Only analytic propositions (for example, logical and mathematical statements) can be aprioristically true, since they have no empirical content and therefore say nothing about what really takes place in the world. [15]

In their essence, logical positivists were empiricists. However, a difference between them and the classical empiricists of the sixteenth to eighteenth centuries, including HUME, is that the positivists gave a linguistic and logical formulation to their theory of knowledge. They focused on clarifying how a sentence could be stated in a meaningful way (ROSENBERG, 2000). A sentence with meaningfulness is a true sentence, corroborated (verified) by experience. In its strong version (SCHLICK, 1979), the criterion of verifiability assumes the existence of basic propositions that are capable of serving as the basis for the process of empirical observation. Thus, a statement is only significant (true) when we can, at least initially, verify it using basic propositions that indicate its meaning —for example, a statement which is caused, as immediately as possible, by perceptive experiences (AYER, 1952). In its weak version (REICHENBACH, 1938), the concept of probabilistic confirmation has been a field of investigation by the logical positivists, who sought to develop a system of inductive logic capable of determining the probability of a hypothesis being true as a function of a set of available data. [16]

POPPER was a critic of logical positivism, and introduced (1959) a second way of thinking about the theory-empirical data relationship. From the perspective of the previously mentioned hypothetical-deductive model, it is up to empirical data to falsify a hypotheses developed aprioristically by researchers. But what does it mean for a hypothesis to be falsifiable? It means that the hypothesis cannot in principle be true in and of itself. A hypothesis results from an exercise of intellect, creative capacity, and consideration of context, since available knowledge offers us concepts, ideas, relationships, etc., such that we almost never start from zero or from a tabula rasa. Thus, in principle, as a product of human intellect, any hypothesis can be true, even though it apparently makes no sense. Ultimately,
the data tell us if our hypotheses are consistent. If confirmed, they contribute to human progress; if falsified, they should be substituted for by others. This shows that a theory must be always subject to revision, reconsideration, and improvement. \[17\]

As mentioned in the previous section, the hypothetic-deductive model was not immune to criticism. In addition to those concerns already cited, another exists, related to the extent of falsification. Considering science from a historical and sociological perspective, several theories that initially seemed to have been falsified, which would indicate that they should be discarded, later proved to be true. Furthermore, when a hypothesis is falsified, it does not necessarily mean that the entire theory from which it was deduced should be discarded. This seems to show there is something more involved in the relationship between theory and empirical data—for realists, for example, this "something more" is the structure of the world itself (WORRALL, 1989), which is represented by the theory, if the latter is to be true. \[18\]

A third way of portraying the theory-data relationship was proposed by HEMPEL (1965), who developed the deductive-nomological model of scientific explanation, by which it is possible to logically deduce a statement that describes a phenomenon based on laws and on the consideration of background conditions. In other words: that which is explained (the *explanandum*) must be deduced from that which explains (the *explanans*), considering the circumstances, and that which explains is a law—a universal statement encompassing a necessary connection between antecedents and consequents, causes and effects. HEMPEL reminds us of an important characteristic of theories: that they unify the fragments of reality, considering what lies beyond, behind and underneath these fragments as well as empirical regularities or irregularities (NAGEL, 1979; ROSENBERG, 2000). When associated with statistical models, for example based on frequency distribution, theories identify or represent repetition and patterns in a particular class of events. They seek order in the world. \[19\]

The three ways of thinking about the relationship between theory and empirical data presented above illustrate a central question in the philosophy of science: how to reconcile the demands of empiricism—which says that for theories to be true they must have empirical content, derived from observation—with those of scientific explanation, by which the explicative power of the theory requires its theoretical terms not to be mere abbreviations for observational terms, but rather to say something more profound about how things work (GODFREY-SMITH, 2003; HEMPEL, 1965; HITCHCOCK, 2004; ROSENBERG, 2000; SCHEIBE, 2001). \[20\]

Contemporarily, a sound perspective on this issue can be found in the work of authors linked to scientific realism and antirealism. From a realist perspective, theories must be interpreted literally: not as a set of statements, propositions or sentences connected to observations, but as truths, in that they tell us about things and their properties (ACHINSTEIN, 2010; CHURCHLAND & HOOKER, 1985; FRENCH, 2007; KHLENTZOS, 2004; MAXWELL, 2011). There is a reality
independent from us, and in order for theories to be scientific, they must tell us the true nature of this reality. This poses several problems for realists. One, which is of interest here, is the problem of how to explain the existence of two or more empirically successful theories explaining the same phenomenon. This problem has become known as the undetermination of theory by evidence (LAUDAN & LEPLIN, 1991; QUINE, 1975). It indicates that there is no way to guarantee an essential, definitive connection between theory and any particular facts and properties of the world. The same phenomenon can be legitimately explained in different ways, using distinct theories and theoretical models. [21]

One solution to the problem of undetermination is to assume that theories have a pragmatic value (BALASHOV & ROSENBERG, 2004; FRENCH, 2007; LEPLIN, 1984; VAN FRAASSEN, 1979). In this sense, the choice of a theory may have nothing to do with the truth or the theory's approximation to the essential facts, but rather with its capacity to help us solve problems of practical interest. Therefore, the aim of a theory would not be "pegged" to the world, but would be designed to help us represent the world in aspects relevant to a proposed transformation of part of it. According to this pragmatic or antirealist perspective, phenomena are not discovered by science, but constructed by it. This argument depends on the premise that we can never come to know the true nature of the world due to the existence of unobservable entities. Phenomena themselves can be examples of the unobservable, since their postulation depends on their incorporation into a theoretical web. This reorders the relationship among a number of key concepts: it is phenomena that are immediately connected to theory, and not empirical data. Data are evidence of phenomena, not of theory (BOGAN & WOODWARD, 1988; HACKING, 1983, 2002; WOODWARD, 1989). [22]

In summary, theories are devices that systematize or organize experience. They are not only instruments for deducing hypotheses and predictions, but also resources of semiotic mediation; they do not only reflect the world in the mind's eye (RORTY, 1979), but (re) construct it according to our pragmatic interests. However, a strong empiricist culture likely persists in our research activities, sustaining a certain "theoretical allergy" and conceptualizing theory and theories in an excessively restrictive sense. Does this also apply to qualitative research? To answer this question, I will now discuss the problem of induction and the role of theory in qualitative research. [23]

4. Induction and Theory in Qualitative Research

4.1 The generic analytic cycle

The field of qualitative methods has grown significantly in recent decades, judging from the profusion of journal papers and textbooks on the subject. As a result of this growth, we have today a complex, diversified field influenced by a large number of schools, authors, and epistemological perspectives. It therefore seems risky to make assertions regarding qualitative methods (which are best given in the plural). Nevertheless, I will attempt to do so in this section. Specifically, I will illustrate what seems to me to be the analytic core of many
qualitative data analysis methods: the cycle composed of data coding, categorizing, and conceptualizing processes. I argue that this analytic cycle exposes the tensions inherent in the process of developing inductive theory from empirical data. [24]

In operational terms, I will denominate the coding and data categorizing process in qualitative research as the "generic analytic cycle." "Generic" is to be understood here as indicating a set of central procedures whose description can vary in textbooks but without altering its fundamentals. I hold that this allows me to broadly discuss the problem of induction and the role of theory in the qualitative research process—which would be technically more difficult if I had to consider the characteristic analysis cycle of each qualitative research tradition separately. Next, I will comment on the three large processes of a generic analytic cycle.

1. The process of analyzing qualitative data begins with researchers establishing initial contact with the material in their set by means of a general reading, followed by careful reading (and thick description; GEERTZ, 1973) of each piece of information—an interview, an image, excerpts from documents. In this process, researchers can (and in some cases must) take notes, in the form of memos (STRAUSS & CORBIN, 1998), to record their impressions and insights, which can help them in later stages of the analysis. Some researchers refer to these records as "audit trials" (LINCOLN & GUBA, 1985).

2. As a result of the previous procedure, it is expected that certain themes and patterns will start to emerge from the data; that is, that they will inductively reveal themselves to the researchers in the data's interaction with the empirical tools as given above. Another alternative in attempting to discover themes would be to analyze data according to an existing framework, that is, deductively. Thus, when creating codebooks for qualitative analyses, in content analysis for example, researchers can be both inductive (allowing themes, patterns, and categories to emerge from the data) and deductive (relying on previous analytical categories, obtained from a theory of reference or even an interview guide), or both at the same time (especially in mixed research designs; CRESWELL, 2008). The coding procedure develops as researchers identify themes and patterns in their data.

3. The coding procedure is complemented by categorization and conceptualization. At this point, the purpose of analysis is to reduce the material even further, at the same time raising its level of abstraction. Classifying or clustering themes or codes into categories allows researchers to organize them and develop conceptualizations about them—that is, explain them. To achieve this, researchers can contextualize their findings (thick description), encompassing a wider picture in which they make sense; compare them to theories and other findings discussed in the relevant and extant literature; compare subgroups, observing whether explanations differ depending on the individuals involved; link and relate categories among themselves (in general, following the criterion of grouping them according to similar characteristics); and use typologies, conceptual models and data...
matrices. Researchers can also try to explain outliers, that is, units of empirical material that do not fit into the theory under construction. [25]

A fundamental question related to the second large procedure as described above, and one which has a direct impact on the relation between theory and empirical data, is what researchers understand by "theme," "pattern," and "category." In general, one can say that themes are related to central meanings that organize experiences. Qualitative researchers often observe that themes can be identified in repeated ideas, sentences, concepts, words, images and sounds; in similarities among units that make up the analysis material (for example, among different interviewees; BERNARD & RYAN, 2010); in indigenous concepts used by individuals to describe their life experiences (PATTON, 2002); in the in vivo codes (STRAUSS, 1987; STRAUSS & CORBIN, 1998), or sensitizing concepts incorporated into the data (BLUMER, 1978); in the frequency and intensity of repetition in the material under analysis; in the location of the themes in discourse and in its centrality as a cognitive element and effective organizer of experience (MANDLER, 1984). In summary, themes can assume both categorical (an instance of the experience, a unit of meaning), and frequential (repetition of themes or their location in networks or schemes) forms. [26]

Identifying themes is the first transposition from the empirical to the theoretical, an initial inductive leap. This does not occur abruptly, but rather as a process of growing abstraction. Indeed, themes can be, at the onset of analysis, simply codes (labels) assigned to certain portions of empirical material—for example, to particular parts of an interview, or even to a single sentence, word, or image. Progressively, codes will merge with others, rearrange themselves, and then reflect a more abstract concept or topic, reducing raw data dispersion. [27]

Regardless of the strategy used, the last procedure of qualitative analysis (third item in the above list) should allow researchers to develop a theory that is not a simple synthesis of observational statements—that is, a description in a broad sense. Researchers must go beyond induction, and it is at this point thatconciliation problems emerge between empiricism and the criteria demanded of a formal scientific explanation. How have qualitative researchers dealt with this problem? [28]

4.2 Situating the problem of induction in the current debate: Some unsolved questions

In general, the solutions proposed are not different from those employed in the inductivist tradition in the philosophy of science (CHALMERS, 1999; LOSEE, 2001). The theory-building process is conducted against a growing backdrop of observational data. Initially, via induction, researchers start from observational data, acquired by either experimental or natural designs, making inferences from the latter by an enumerative induction process. So, theories (or general-universal statements) are proposed. Secondly, via deduction, these theories are used to explain the phenomena investigated (HENNINK, HUTTER & BAILEY, 2011). [29]
In qualitative research, as I have pointed out in the beginning of this article, this is very well illustrated by the GTM, which proposes an analytic spiral stemming from data and progressing to the explanation, combining two large vectors: one ascending, aimed at developing the theory, and the other descending, seeking to ground concepts in the data. It is, therefore, a two-handed process from description to explanation, always comparing cases and organizing them into increasingly central and abstract thematic categories. Interplay between the theory being developed (grounded) and the available deductive theories is promoted at all times (Hennink et al., 2011). Without this interplay, it would seem difficult to justify the scientific relevance of the qualitative procedure, which would be no more than just another way of cataloging and describing empirical facts without any connection with broader phenomena and theories. [30]

However, perhaps not even the interplay between small- or midrange theories that are generated inductively from a set of available empirical data and large-range (deductive) ones is able to rid qualitative methods of the induction problems discussed in this article. In the first place, as I have already mentioned, nothing guarantees that discrete empirical data, even when collected in large amounts and under widely varying conditions, can support large-range theories on their own. They may sustain parts of these theories, hypotheses, and questions, but not the theories as a whole, whose development depends on other factors (e.g., research programs agenda) and not only on the stock of discrete observational statements. [31]

Thus, on what basis can it be said that categorizing data from interviews with a determined set of individuals allows researchers to make non-observational (therefore, theoretical) statements about a phenomenon that is (say) psychological? Qualitative researchers can counterargue by stating that the purpose of their work is not to produce generalizations (in terms of law-like statements) but rather to understand the phenomenon. However, by acting this way, the research in question runs the risk of being purely descriptive and its explanation just an abbreviation for situated empirical observations (Rosenberg, 2000). This is not about the number of subjects, which is a sampling problem; it refers to the degree to which empirical data, irrespective of the amount, can support non-observational (theoretical) statements. [32]

In the second place, when a theory is inductively constructed, one can assume that empirical data are able, in and of themselves, to frame or postulate the phenomenon investigated. As a consequence, the theory-building process can advance "in the dark," since the phenomenon takes shape as the empirical data accumulate. Phenomena are directly determined by theory, and only indirectly confirmed by empirical evidence (Apel, 2011; Bogan & Woodward, 1988; Hacking, 1983; Woodward, 1989). They are not independent from the way in which I posit and interpret them; that is to say, they are theory-laden (French, 2007; Schindler, 2007). Next, my perception of reality depends on my previous experience and above all my prior knowledge. Therefore, the choice of which facets, properties, or qualities of a phenomenon will be considered
depends on its integration into a theoretical web, in the holistic sense advocated by LAKATOS (1978), and especially by QUINE (1978, 1998). [33]

"Generic" qualitative methods are not necessarily confined to a single theoretical web or research program, even though most of their assumptions are derived from such a theoretical basis—for example, symbolic interactionism and pragmatism, in the case of GTM (KELLE, 2005). Considering this point, I ask the following question: how are researchers who do not align themselves with the central principles of symbolic interactionism, or other microsociological theories that form the basis of many qualitative methods, supposed to justify the use of the generic analytic cycle illustrated here to conduct their research and analyze their data? The use of generic analysis methods can be an *ad hoc* resource. When research programs (LAKATOS, 1978) define their phenomena of interest, they create their own methodological criteria. Methods, in the sense of techniques, must be understood in the theoretical context of a research methodology (CROTTY, 1998; VALSINER, 2000). This last proposition is certainly not alien to qualitative researchers. However, I believe there is a need to re-emphasize this point, which seems to be critical if we are to address the problem of induction in qualitative research. This assertion will be discussed in the following section. [34]

5. Suggestions for Reconsidering the Problem of Induction in Qualitative Research

I propose three brief suggestions for addressing the problems outlined in the preceding section regarding how research using what we call "generic" methods in this paper can deal with the problem of induction and theory building. [35]

The first suggestion, already alluded to in previous sections, is that qualitative researchers rehabilitate concepts that depend more substantially on a theoretical web, in the sense used by QUINE (1978, 1998). This is based on the assumption that concepts acquire meaning in the theoretical context to which they belong. However, rehabilitating concepts involves reflecting more vigorously on the meaning of the theory used over the entire course of the research, and not only when analyzing empirical data. Obviously, this is not an unfamiliar point to qualitative researchers. Indeed, they have been discussing this issue since the late 1960s (BRYANT & CHARMAZ, 2007; CHARMAZ, 2006; FLINDERS & MILLS, 1993; GLASER, 1978; GLASER & STRAUSS, 1967; LAYDER, 1993; SANDELOWSKI, 1993; STRAUSS, 1987; STRAUSS & CORBIN, 1998). [36]

Nevertheless, I believe that the debate is far from over. If we take a look in recent textbooks covering qualitative research, we will notice that the authors' focus seems still to fall on the distinction between and combination of induction and deduction in the coding and classification process (e.g., BERG, 2001; BERNARD & RYAN, 2010; COFFEY & ATKINSON, 1996; CHARMAZ, 2006; FLINDERS & MILLS, 1993; GLASER, 1978; GLASER & STRAUSS, 1967; LAYDER, 1993; SANDELOWSKI, 1993; STRAUSS, 1987; STRAUSS & CORBIN, 1998). [36]
BOGDAN, 1998). It seems less common to find a metatheoretical reflection that questions this traditional conception of the knowledge-producing cycle, or attempts to apply qualitative literature to current debates in the philosophy of science. For example, in a historical study aimed at clarifying the concept of theoretical sensitivity and its role in the categorization and theory-building process, GLASER (1978) proposes a distinction between two types of codes: substantive codes, developed during the open coding stage; and theoretical codes, which refer to the formal categories of the social sciences and for that reason, bear the mark of their background theories. However, as suggested by KELLE (2005), GLASER was unable to effectively show how formal terms are related to substantive or observational ones. Instead, he seems to endorse the distinction between observational statements, on the one hand, and theoretical ones on the other. I believe that the same problem occurs in other generic qualitative methods. [37]

One reason for this can be connected to the implicit concept of theory held by these methods. In some cases, qualitative researchers seem overlap theory with "categories" (STRAUSS & CORBIN, 1998). In this case, theory is thought of as the conceptual component that links empirically grounded thematic categories. Thus, its role seems to be to sustain bonds or mediate between empirical categories and wider theoretical concepts. [38]

In other cases, qualitative researchers seem to understand theory in a paradoxically similar way as do logical positivists: as a set of statements that depend on empirical content for their validity. Depending on its objectives with respect to empirical verification, qualitative research can be confirmatory or exploratory (GUEST et al., 2012). Both analytical induction (ZNANIECKI, 1934) and classical content analysis (Krippendorff, 2003; Schreier, 2012) are examples of this. Thus, qualitative research may aim to refine existing theories; confirm or falsify hypotheses (derived from current theories); develop new inductive theories; present counterfactual inferences (that is, cases that do not confirm one current theory); and even make inferences, in the sense of prospective causal explanations. The work of KING, KEOHANE and VERBA (1994) is a good example of this last position. [39]

The second suggestion has to do with the insistence that qualitative researchers, especially novices, consider their research within wider theoretical traditions (or theoretical webs), avoiding, as much as possible, general and standard methods as well as a "technist" approach to research. To that end, they must have at least minimal knowledge of their basic theoretical assumptions. Some common theoretical traditions present in the qualitative research literature are phenomenological, hermeneutical (including narrative research), discursive, ethnographic, and also grounded theory. Researchers espousing other theoretical traditions can and do equally benefit from the qualitative perspective, provided they manage to justify its use vis-à-vis the fundamental assumptions of their respective theoretical orientations. [40]
My third suggestion is that qualitative researchers rethink the role of "emergence" or unexpected facts in qualitative research, as well as the relationship of these facts with the theorizing process (e.g., BEDAU & HUMPHREYS, 2008). Throughout this article, I have insisted that investigation of a scientific phenomenon depends on its incorporation into a particular theoretical web. Moreover, this web is not merely a set of hypotheses from which predictions can be made. If this were so, I would simply be recapitulating the hypothetical-deductive approach in the domain of qualitative methods, saying that theory comes "before" data. Instead, I suggest, based on SCHEIBE (2001), that the dynamic between theory and empirical data involves a reconstruction process, and that the theoretical web is actually a background that guides us, sometimes tacitly (POLANYI, 1966), in relation to a phenomenon, its relevant dimensions, and ways to better access it. The "meeting" between theory and phenomenon can often occur in a casual, unpredictable, and unexpected manner, although always within a scientific and theoretical context. In this sense, to explain the situation in which the theory-building process results from unexpected events or phenomena, qualitative researchers (e.g., KELLE, 2005; REICHERTZ, 2009; RICHARDSON & KRAMER, 2006) have proposed using PEIRCE’s (1955) concept of abductive reasoning, which, roughly speaking, stimulates the researcher to overcome the initial surprise provoked by an unexpected fact, leading to the creation of new rules (theories) for its explanation. [41]

A last comment: when I refer to a fact, occurrence or event as unexpected, I may think this occurs because, even though I depend on semiotic systems (for example, theories) to deal with the world, the latter hardly can be "totalized" by the former. In other words, my comprehensive systems are unable to capture reality in all its complexity. At the same time, this may mean that there is "something more" beyond my symbolic systems, causing them to be continuously subject to revision. This is the realist position in a broad sense. Currently, a specific version of this position, called critical realism, advocates the existence of an objective reality formed by events and their underlying causes, about the latter of which one can never acquire definitive knowledge. In qualitative research, we observe recent efforts to move closer to this form of realism (e.g., BHASKAR, 2008; MANICAS, 2006; MAXWELL, 2011). This perspective seeks to position itself in a field challenged by forces such as those represented by empiricism, materialism, idealism, relativism, constructionism, and the like. It also advocates the use of abductive reasoning (CLARK, 2008), and defends the importance of theoretical models in science (JACCARD & JACOBY, 2010). Due to the recentness of the embrace of critical realism by qualitative researchers, it is still difficult to predict its impact on the theory-building process, although it is apparently a positive development for the field to incorporate new philosophical perspectives in order to evaluate its own practices. [42]
6. Final Considerations

6.1 General overview and limitations

The purpose of this article was to reflect on the ramifications that the problem of induction poses for qualitative research. My expectation, by bringing discussion of the philosophy of science closer to the context that emerges out of consideration of qualitative methods, was to show that the latter inherit many of the problems inherent to any criteria for justifying knowledge-claims and scientific demarcation. Decades invested in the attempt to establish the exact nature of qualitative methods and demonstrate their relation to induction have obscured the fact that they are beset by tensions similar to those already identified by philosophers of science in other areas of knowledge: tensions between distinct conceptions of theory; around the role of empirical data; and between explaining and understanding, causes and reasons, \textit{a priori} and \textit{in vivo} categories, theoretical categories and indigenous concepts, and framing and emergence. \[43\]

There are obvious limitations to the present paper. Among these is the fact that I worked with a standard or generic version of the qualitative analytical proposal, based on the processes of coding, categorization, and conceptualization. Although this decision operationally facilitates analysis, it also limits my ability to appreciate subtleties, exceptions, and counterexamples. Perhaps the discussion of the problems of induction and the theory-building process should be held in the context of the specific traditions of qualitative methods. Another limitation that can affect the scope of my arguments, related to the previous one, is the fact that I focused mainly on the data-analysis cycle. It would certainly be enriching if we could consider the qualitative research cycle as a whole, since the processes involved in defining the theme and object of study and the way it is approached operationally (for example, data collection) may reveal equally valuable information about the role of theory in the qualitative approach. \[44\]

6.2 Contributions to scholarship: Revisiting theory building in qualitative research

As noted throughout this article, the ideas discussed here are certainly not unfamiliar to qualitative researchers, particularly the problem of induction and the impossibility of conducting research without substantive theoretical assumptions (e.g., FLINDERS & MILLS, 1993). Nevertheless, I would like to revisit and reiterate three points introduced over the course of the text since I believe they may contribute possible new insights or encourage other researchers to revisit the current debate on theory building in qualitative research. The presentation of these points concludes this article. \[45\]

First, when I propose considering the "generic analytical cycle" (Section 4.1), I assume the (traditional) idea that qualitative research is a cycle of induction and deduction, and that coding (finding a concept or category that fits certain incidents in the data) is the inductive part of that cycle. It may be that the entire conceptual confusion surrounding theory building in qualitative research is rooted
in this idea. However, one can also argue that finding a category is not merely induction, but also the process of abductive inference described by Peirce. Abduction is known to be a logical, as well as rational and scientific, inference that enables the creation of new forms of knowledge (Reichertz, 2009). As we have seen, in the generic analytical cycle proposed in Section 4.1, researchers deductively draw upon concepts from an extant theory in order to explain, accommodate or embed their emergent substantive theory (the theory they were able to ground in their data). However, an alternative to this traditional use of deduction is to create new forms of explanation or rules capable of "fitting" the surprise and shock caused by their data and of going beyond explanations available in the extant theory. Abduction is precisely this process of creating a novel type of combination between features present in data as well as in extant theory (Kelle, 2005). It depends on the creativity of the researcher, on an intellectual act, a "mental leap" (Reichertz, 2009, p.7), through which previously unassociated things now become associated. [46]

Second, I feel it is important to reiterate that discussion regarding the "generic analytical cycle" seeks to elucidate an apparently serious, and in my opinion underestimated, problem in qualitative research: the use of qualitative analytical methods as *ad hoc* devices. My hypothesis is that this contributes to theoretical assumptions and concepts entering the process of inductive generation of theories unnoticed. As such, the degree to which the method is not theory-free is underestimated. Moreover, since it is considered an *ad hoc* resource, the generic analytical cycle can assume the role of a driving force behind the investigation, linking all areas of the study, including the theoretical review. Some authors even suggest that researchers should conduct the literature review only after their data analysis is set out. For example, Heat (2006) recommends that researchers develop an "inductive sensitivity," through which they can arrive at an "insightful identification of relevant literature" (p.522). This concern about *when* literature should be consulted and about its *place* in the entire research endeavor has been at the heart of the "forced vs. emergent" debate in qualitative research (e.g., Duhscher & Morgan, 2004; Dunne, 2011; Kelle, 2005), specifically in the GTM tradition. In this respect, Glaser (1992) states clearly, "There is a need to review any of the literature in the substantive area under study" (p.31). I think this position regarding the timing of the literature review and its role in research can lead researchers to overestimate issues of method and, consequently, create an imbalance regarding theoretical issues. Consequently, I think discussions about the relationship between method, theory, and, as I describe below, phenomena may represent a sensitive area in the context of qualitative research. [47]

Third, when proposing a criticism on inductive thinking present in the generic analytical cycle, I mentioned that data itself may not be sufficient to sustain a theory and, based on a tripartite theory-phenomena-data model discussed in the current philosophy of science (e.g., Apel, 2011; Bogan & Woodward, 1988), I pointed out that phenomena should be directly explained by theory, and only indirectly supported by the data. Thus, if the tension between theoretical statements and empirical statements is not exactly a novelty for qualitative
researchers today, debate concerning theory, phenomena, and data may very well be. Therefore, I believe that a possible novelty in this discussion is that of problematizing the meaning of phenomena in qualitative research. Indeed, it seems that this intuition is supported by recent literature. For example, TOOMELA (2011) directs a loud criticism at qualitative methods, arguing that due to a number of fallacies, these methods do not answer the fundamental question about what phenomena are. According to TOOMELA, "the aim of modern qualitative investigations is the study and development of concepts, and not the phenomenon itself" (p.37). The author notes that this occurs partly because what qualitative researchers study "is not the external to the research world with the help of symbols as tools but rather the tool itself—the world of symbols (...)" (p.34). To sum his critical tableau against qualitative methods, TOOMELA observes that the problem of induction occurs because qualitative study is not always guided by an explicit a priori research question, on the assumption that beginning with such a question may compromise the emergence of a (substantive) theory, the kind of theory that is relevant to participants—and not only to the researcher. [48]

Although it is beyond the scope of this article to discuss TOOMELA's arguments against qualitative research, it might be worth noting that recent literature on qualitative methods could bring forth deeper reflections on the ontological status of the phenomena they study. Furthermore, there is an inherent paradox in qualitative research (maybe in quantitative research as well), and, on this point, the discussion surrounding the problem of induction, articulated with a discussion about the substantive relationship between theory and phenomena, may be instructive. The paradox is that in order to "access" a phenomenon, theory is required, but to innovate and create new possibilities of empirically reconstructing phenomena, it is also necessary to go beyond current theoretical frames or, as stated by researchers using abductive logic, go beyond the current rules of established knowledge (e.g., REICHERTZ, 2009). The paradox may reside in the ambiguity with which a phenomenon is often defined. On the one hand we encounter definitions of phenomena as natural kinds (BROWN, 1994), that is, something already "given" in nature that must be discovered by means of a scientific method (from the Greek phainomenon, "thing appearing to view"). On the other hand, particularly according to some "radical" qualitative viewpoints, phenomena are considered just "linguistic constructions." In such cases, reality is equated with the description we give to it (TOOMELA, 2011). In this article, I have suggested another possibility: that phenomena are posited by theory in an empirical reconstruction process (SCHEIBE, 2001), meaning that empirical observation should not be disregarded but rather repositioned in existing theoretical networks or in those still to be created. This polysemy of the concept of phenomenon involves theoretical as well as rhetorical implications. In my opinion, it reveals a weakness in the manner of conceiving the substantive role of theory in research activity. [49]

In the case of qualitative methods in particular, my hypothesis is that the dearth of more detailed debate on this issue represents not only a gap in the current scenario, dominated primarily by methodological discussions, but also a
significant intellectual challenge to be overcome. It is perhaps an opportune moment, following the consolidation phase of the identity of the qualitative methods, for researchers to focus more systematically on the rational justification of their practices in order to establish deeper dialogue with the philosophy of science and other disciplines with relevant perspectives on the nature of scientific knowledge. [50]

References


Lakatos, Imre (1970). Falsification and the methodology of scientific research programmes. In Imre Lakatos & Alina Musgrave (Eds.), *Criticism and the growth of knowledge* (pp.91-196). Cambridge University Press.


Author

Pedro F. BENDASSOLLI teaches organizational and work psychology at UFRN, Brazil. He authored or edited seven books and published several papers in the fields of work psychology and epistemology. Currently, his research interests include philosophical grounds of qualitative research, and applied themes in organizational and work psychology. Professor BENDASSOLLI holds a doctoral degree in Psychology from the University of São Paulo.

Contact:
Pedro F. Bendassolli
Department of Psychology
Federal University of Rio Grande do Norte
Av. Salgado Filho, s/n, Cidade Universitária
59072-970 – Natal, Brazil
Tel.: +55 84 3215 3590
E-mail: pbendassolli@gmail.com
URL: http://www.pedrobendassolli.com/

Citation