2007

On the Replication of Positivist Case Study Research

Rene Riedl
University of Linz, rene.riedl@jku.at

ON THE REPLICATION OF POSITIVIST CASE STUDY RESEARCH

Riedl, René, University of Linz, Altenberger Straße 69, 4040 Linz, Austria, rene.riedl@jku.at

Abstract

A major characteristic that is often said to distinguish the scientific method from other approaches to knowledge is its objectivity. Conducting replication studies allows for critically testing of objectivity and therefore significantly contributes to scientific progress. This article deals with the nature of replicating positivist case study research and outlines that the replication of case studies—if conducted well—may result in both (i) critical test of existing findings and (ii) groundbreaking discovery of new scopes of validity of existing theories. In this article, at the beginning philosophy of science considerations are used to substantiate the importance of replication research for scientific progress in the IS discipline. Afterward, well-known case study replication concepts by Yin (2003) and Lee (1989) as well as more general remarks from the behavioral, social, and administrative sciences are taken into consideration and then integrated in order to develop a set of positivist case study replication strategies.

Keywords: Replication, Case Study, Positivist Case Study Research, Philosophy of Science.

1 INTRODUCTION

No doubt, we would all agree that a person who concluded that all cats are black after seeing only one cat that happened to be black, was unintelligent. Scientists in the IS discipline have never been as unintelligent in drawing conclusions about their phenomena—have they? If one reads, for instance, the publication guidelines on "What Not to Submit to the MIS Quarterly", the answer might challenge the view that replication is something important:

> "MIS Quarterly should not publish articles that offer little that is truly new to the information systems literature. Journal space is a scarce commodity; and there is increased competition for the space that exists. Thus, even if a manuscript describes a study that is carefully executed and internally valid, it is unlikely the manuscript would be published if its primary contribution was a corroboration of earlier studies. Authors of such manuscripts are strongly encouraged to submit them elsewhere." (Zmud 1995)

But not only editorial boards of IS publication outlets often reject replication studies. Rather, also in the behavioral, social, and administrative sciences replication studies are not common. For example, a survey study performed by Neuliep and Crandall (1990) found that almost 94% of the sampled editors of social and behavioral science journals stated that replication studies are not included as examples of research encouraged for submission in the editorial policy of their journal. If one analyzes the journal articles itself, the number of replication studies clearly reflects the editorial policies. For example, Bozarth and Roberts (1972) found more than three decades ago that less than 1% of a sample of 1,334 published articles in three psychology journals reported on replication studies. Greenwald (1975) even found no replication studies in a sample of 362 *Journal of Personality and Social Psychology* articles. A more recent content analysis of 18 leading business journals covering the 22-year time period from 1970 to 1991 found that replication research constitutes less than 10% of published empirical work in the accounting, economics, and finance areas, and 5% or less in the management and marketing fields (Hubbard & Vetter 1996).

In contrast, in natural sciences replication is considered as an essential part of the business of doing research. Hendrick (1990), for instance, describes an interesting example of what is happening when a nuclear physicist announces the discovery of a new particle: Every researcher who is lucky enough to own a linear accelerator is going to crank it up and go searching for the new particle. The objective is to replicate the study in order to establish a basic existence law for the previously unknown particle. Once the physics community is certain that the particle exists, attention turns to the systematic study of cause-effect-relationships to establish mathematical laws of the particle behavior—the scientists develop a physical theory. However, arguing that every study in the IS discipline should be replicated would be ridiculous.

At this point, one might argue that the research field of social sciences differs from the one of natural sciences and therefore replication is not an equally important issue. However, this argument only holds true when one does *not* follow a positivist philosophy. The main objective of positivist research is the development of law-like generalizations independent of time or context (Orlikowski & Baroudi 1991). Hence, within the positivist paradigm replication is important in order to assess the predictive power of nomothetic statements. In contrast, research with an underlying interpretive philosophy adopts the position that our knowledge of reality is a social construction by human actors (see, for example, Walsham (1995a) for a discussion of the interpretive perspective in IS research). If one follows the interpretive paradigm, value-fee data cannot be collected, because the researcher uses her preconceptions in order to guide the research process. Additionally, the researcher often interacts with the human subjects of the investigation, changing the perceptions of both parties. Hence, interpretivism contrasts with positivism, where nomothetic statements are possible, whereas interpretive research aims at deriving idiographic statements; i.e., the objective is to explain and understand past data and not to predict future situations (Walsham 1995b, Walsham 2006). Therefore, it is clear that replication takes on a different role in the subjective paradigm. As Berthon et al. (2002) point out, the generation of valid knowledge concentrates on the process of self-reflection in the

subjective paradigm. Thus, the aim of replication studies is not to increase accuracy (i.e., the development of an ever-more accurate representation of some external reality). Rather, replication in interpretive research aims at gaining increasingly more in-depth understanding of the phenomena investigated (i.e., building richness of phenomenological experience).

In the IS discipline both the positivist (Benbasat et al. 1987, Eisenhardt 1989b, Lee 1989, Yin 2003) as well as the interpretive (Carroll & Swatman 2000, Walsham 1995b, Walsham 2006) stance are taken up on case study research. (There are also some researchers, for example, Lee (1991) or Gable (1994), who suggest a combination of positivist and interpretive research designs.) Originally, the term "case study" indicated a research method that was characterized by little methodological rigor (Dubé & Paré 2003). Fortunately, during the past two decades increasingly more researchers called for more methodological rigor in case study research—especially those from the positivist tradition (e.g. Benbasat et al. 1987, Lee 1989, Yin 2003). However, the main objective of this article is the development of a set of replication strategies that outline possible ways of how to replicate positivist case study research. Hence, this article may be considered as a detailed specification of the currently available more general remarks by the positivist case study community (e.g. Yin 2003).

The remainder of this article is organized as follows: First, philosophy of science considerations are used to stress the relevance of replication research for scientific progress in the IS discipline (section 2). Then, well-known case study replication concepts by Yin (2003) and Lee (1989) as well as more general remarks from the behavioral, social, and administrative sciences are taken into consideration (section 3) to develop a set of strategies that allow for sound replication of positivist case research (section 4). It is outlined that the replication of case studies—if conducted well—may result in critical test of existing findings (i.e., pure replication or reproducibility) as well as groundbreaking discovery of new scopes of validity of existing theories (i.e., extension or generalization). Finally, section 5 draws some overall conclusions on the topic.

2 ON THE IMPORTANCE OF REPLICATION

Considering the observations of philosophers like Kuhn (1962), Lakatos (1970), or Popper (1972), there is no doubt that the development of scientific progress and the accumulation of scientific knowledge are *not* straightforward processes. Rather, the scientific field is a political as well as an intellectual enterprise (Bornstein 1990). Furthermore, human errors of procedure, observation, recording, computation, or report are well known to make researchers wary of the unreplicated study (Rosenthal 1990). Consequently, researchers of all empirical disciplines have long been aware of the importance of performing replication studies (e.g. Campbell & Jackson 1979). If asked, whether replication studies are important or not, the majority of the scientific community would probably say "Yes". However, in our rush for new knowledge, generation rather than replication, "search" rather than "re-search" predominates (Berthon et al. 2002).

Thinking about the importance of replication research implies knowing what the term "replication" exactly means. The term "replication" has its origins in the Latin term "replicare"—and this means to "fold back". Basically, replication is the process of going back and "re-searching" an observation, investigation, or experimentation to compare findings (Berthon et al. 2002). "Re-searching" observations is of paramount importance because no single research study can simultaneously consider (at least control) all variables that might have influence on a dependent variable observed. Hence, the investigation of a particular phenomenon should not be reserved to one single researcher. Otherwise, it is likely that important independent variables are not considered, and this might lead to inaccurate conclusions. Additionally, scepticism against research findings of other scholars is of particular importance for scientific progress, too (remember, for example, Hendrick's (1990) example on what is happening when a physicist discovers a new particle.)

A widely accepted assumption in the IS discipline is that if the result of a study is significant, then theories should be based on it. As pointed out by many researchers in the social sciences (for example,

Amir & Sharon 1990), this assumption is a severe fallacy. The main criticism is that a theory should fulfil at least two prerequisites: (1) the test of reproducibility (that can be defined as the possibility to find an originally found cause-effect-relationship in a *similar* sample, situation, context etc. again) and (2) the test of generalization (that can be defined as the possibility to find an originally found cause-effect-relationship in a *different* sample, situation, context etc.).

Basically, by performing replication studies both the reproducibility test as well as the generalization test can be conducted (Amir & Sharon 1990). It has to be mentioned that strictly speaking, in the social sciences the concept of "significance" (i.e., $p < .05$) can only be applied to pure quantitative research such as experimental and survey research. However, also the results of positivist case study research can be "significant", especially, if they build and test theories (Eisenhardt 1989b, Yin 2003). But, in contrast to pure quantitative research, within the qualitative and mixed-method research paradigm, respectively, the significance of the results is not determined by statistical techniques. Rather, the significance of the results is determined by the potential to extend and generalize theories. To speak with the words of Yin (2003): "The significance of positivist case research results is defined by the potential for 'analytic generalization' and not 'statistical generalization'."

As already mentioned, many researchers do not conduct replication studies because it is difficult to find a journal that publishes this type of research (Neuliep & Crandall 1990). Even more controversially, if a researcher decides to conduct a replication study, then he is confronted with a paradox (Bornstein 1990). If a replication study is performed, there are two possible outcomes: (1) The original finding is replicated. As a consequence, manuscript reviewers, journal editors, and the majority of the scientific community are likely to consider the results as trivial, because the study is only a demonstration of something that was already known. (2) The original finding is *not* replicated. In case of pure quantitative research, the researcher has to report then a nonsignificant result (and usually there is little trust in nonsignificant results, see for example, Greenwald 1975 or Kupfersmid 1988), or in case of qualitative or mixed-method research, the scientist shows a result that is different from the original one. In both cases, the major problem is that the result of the replication study cannot be interpreted, because the scientific community cannot be sure whether the diverging result was due to the absence of a relationship between the independent and the dependent variable(s), or due to methodological issues. The missing reliability of the measurement tool as well as flaws in the design of the experimental study, for instance, may be responsible for the diverging result between a replication and the original study.

3 RELATED WORK

Dubé and Paré (2003, p. 600) state the following characteristics of case research: (1) a contemporary phenomenon is examined in a real-life context or setting, (2) one or few entities (person, group, organization, technology) are examined, (3) the complexity of the unit is studied intensively, (4) the phenomenon of interest is not isolated from its context, especially at the data analysis stage, and (5) no controlled observation that involves manipulation is involved. As a result of characteristic (3), it is clear that performing case studies implies the usage of *several* data collection methods (e.g. interviews, document analysis, and questionnaires) as well as data analysis methods (e.g. field notes and coding or statistical methods such as time series analysis).

The knowledge gained from replicated case studies is usually considered to be more validated than from one case. Hence, the overall study (that consists of multiple cases) is considered to be more meaningful than the single case (Herriott & Firestone 1983). However, as Yin (2003) points out, in some situations it is simply not possible to conduct replicated case studies. Consider, for example, the unusual, rare, or critical case. Furthermore, limited resources and time may be a challenge, if a researcher wants to replicate case studies.

However, considering the diverging notions on case study replication research, it is important to define the concept of multiple-case design. It is a prevalent mistaken analogy in science to regard multiple

cases to be similar to the multiple subjects of an experiment or the multiple respondents in a survey study (Yin 2003). Rather, multiple cases should be considered as one would consider multiple experiments or surveys. Only if one accepts this view, it is possible to understand the scientific value of case study replication research. Basically, if one replicates a case study, on the one hand, it is possible to *attempt* to duplicate the exact conditions of the original case. On the other hand, it is possible to alter some of the previous conditions. For example, it is imaginable to alter the context of the case organization and therefore it is likely that new independent variables are investigated.

3.1 Yin (2003): The logic of multiple-case studies

Yin's logic of multiple-case studies assumes that case research is done on the basis of a rich theoretical framework. The main reason for using one or more theories is that it allows for the prediction of future events as well as the prediction of the absence of future events. Importantly, without a theoretical framework it is not possible to observe and reason the absence of future events, because then events simply do not happen. McClintock et al. (1979, p. 616) stress: "Only with preconceived expectations can non-events be identified".

Suppose, for example, the following IT outsourcing research situation whose investigation could be theoretically based on agency theory (for an assessment and review of agency theory, see, for example, Eisenhardt 1989a): Following agency theory, in an IT outsourcing relationship the service provider constitutes the agent and the client company constitutes the principal. One important proposition within the agency theory framework is: When the contract between the principal and agent is outcome based, the agent is more likely to behave in the interests of the principal (Eisenhardt 1989a, p. 60). Hence, if a contract is *not* outcome based, then it is very likely that an agent reduces his effort which may result in poor performance—in other words: the agent will behave opportunistically. Poor performance can be easily measured. For example, by using service level management tools that measure system availability and other metrics (Sturm et al. 2000). However, a case researcher who follows agency theory and who investigates an outsourcing relationship in which an outcome based contract is *not* in place would expect opportunistic behavior by the agent. Hence, the researcher would be able to observe not only events that actually happen but also non-events like the absence of opportunistic behavior. A scholar who does not base his research activities on agency theory (or another theory predicting the same future events or non-events like agency theory) would not be able to observe the absence of opportunistic behavior. As a consequence, this scholar would not be able to ask for the reasons for the absence of the opportunistic behavior. In contrast, the scientist who bases his research activities on agency theory would be able to identify reasons for the absence of opportunistic behavior. Trust, for example, that is established between the principal and agent is likely to eliminate (at least reduce) opportunistic behavior (for a review on IT outsourcing that deals with the role of trust among other things, see, for example, Dibbern et al. 2004). As a result of the observation of the role of trust in contractual arrangements, the scientist is able to extend agency theory by identifying trust as a cause for non-opportunistic behavior.

Yin's case study research framework primarily does *not* aim at describing and understanding the phenomena investigated. Rather, the main objective is the explanation of social phenomena—to be more precise: the investigation of cause-effect-relationships. From an epistemological point of view, therefore, Yin's case study framework is a concept that can serve as a methodological basis for explorative and/or explanative research. As a consequence, rigor methods concerning data collection and analysis are indispensable for this approach—concerning this matter, see, for example, an early article of Yin and Heald (1975) on the case survey method.

As far as the replication logic is concerned, Yin (2003) stresses that each case has to be carefully selected by the researcher so that it either (i) predicts similar results (a *literal* replication) or (ii) predicts contrasting results but for predictable reasons (a *theoretical* replication). Interestingly, Yin (2003, p. 48) states an IS example in his book to illustrate the concept of literal and theoretical replication, that is summarized analogously in the following:

> Consider the initial proposition that an increase in using computers in school districts will occur when such a technology is used for both administrative and instructional applications, but not either alone. To investigate this proposition in a multiple-case study design, three or four case organizations might be selected in which both types of applications are present. Then, the objective is to determine whether, in fact, computer use did increase over a period of time (literal replication). Furthermore, three or four additional cases might be selected in which only administrative applications are present. Here, the prediction is that little increase in computer use is taking place (theoretical replication). Finally, three or four other cases might be selected in which only instructional applications are present. And again, the prediction is that little increase in computer use is occurring (another theoretical replication). If this entire pattern of results across these multiple cases is found, then, the entire research study would provide substantial support for the initial proposition.

Dubé and Paré (2003) conducted a meta-study on positivist case study research. A total of 183 positivist case studies were identified and 109 case studies used a single-case design. Accordingly, the remaining 74 case studies used a multiple-case design. Out of these 74 studies, 9 used literal replication logic and 15 used theoretical replication logic (the remaining 50 studies did not specify their replication logic at all).

Finally, it has to be stressed that Yin was probably the first in the social sciences who attempted to relate case study research to typical quality criteria of quantitative research such as construct validity, internal validity, external validity, or reliability. Due to the fact that the multiple-case study design—which is a special case of replication research—is an important element of Yin's (2003) case study framework, replication strategies may affect these quality criteria (see, for example, Yin 2003, p. 34).

3.2 Lee (1989): How to allow for replicability

First of all, it is important to recognize that in contrast to Yin's (2003) replication framework, Lee's (1989, p. 40) analysis of case study replication logic explicitly assumes that the replication study is performed by an "independent investigator" and not by the researcher who has already carried out the preliminary study. Importantly, this assumption has considerable implications for the way of how to conduct case study replications. Lee points out that basically it is possible to replicate a case study in exactly the way that another researcher has performed it. This would involve to (i) apply the same theories to (ii) the same set of initial conditions in order to (iii) deduce the same predictions, and then (iv) test these predictions against the same observations made by the other researcher. Unfortunately, in IS research it is seldom possible to replicate the set of initial conditions. Lee (1989, p. 40) writes: "[A]ny observed configuration of individuals, groups, social structure, hardware, and software in a real-world setting is highly unlikely to recur and be observed again". As a result of the difficulty to replicate the initial conditions, it is not possible for an independent investigator to verify the case study findings of another researcher by attempting to replicate the exact initial conditions of a previous case.

However, Lee (1989) suggests another replication procedure that incorporates basically the same logic as Yin's (2003) theoretical replication framework. The independent investigator can apply the same theories as tested in the previous case(s) to a different set of initial conditions. As a result of this procedure, also the predictions will alter. And finally, such new predictions would presuppose different observations. Lee (1989, p. 41) substantiates his statements by a demonstrative example:

> "[T]he investigator would be working with a new prediction, 'Plato is mortal,' as opposed to the original prediction, 'Socrates is mortal"; even though the prediction would be different; it would still be the same theory being tested [The theory is that all humans are mortal; comment by the author of this article] … even though the observations in a particular MIS case study are non-replicable, the case study's findings (that a particular theory is confirmed or disconfirmed) would be replicable".

4 REPLICATION STRATEGIES FOR POSITIVIST CASE STUDY RESEARCH

As far as replication studies are concerned, various replication strategies (RS) can be applied. As outlined above, explanatory positivist case studies aim at testing the validity of theories (see, for example, Dubé & Paré 2003 and Yin 2003). Hence, if one conducts a replication study a basic objective is to verify the results of the original study concerning the validity of a theory. From this it follows that the theory investigated in a replication study has to be the same like the one investigated in the original study—to put it differently: In a replication study at least the most important dependent and independent variables and their interrelationships have to be held constant. In contrast to this fact, various other characteristics of a research process can be altered when a replication study is conducted. Three of them, namely "researcher", "context", and "methods" will be discussed in the following.

It is essential to outline that other researchers, for example, Berthon et al. (2002), consider "theory" as another characteristic of the research process that can be altered in a replication study. For example, in a replication study the same data could be collected as in the original case and used to test whether a particular theory explains the results of both cases better than the theory used in the original case. Following Yin (2003), such a procedure would address the issue of establishing internal validity of case study research.

4.1 Researcher

A replication can be either conducted by the same researcher like the original study or by another researcher. By whom a replication is conducted has been a central question in replication research for a long time. Karl Pearson, for example, already noticed in 1902 that "the high correlation of judgments [suggesting] an influence of the immediate atmosphere, which may work upon two observers for a time in the same manner" (p. 261). Considering this statement, an essential question arises: Are, for instance, five replication studies performed by a single research group (e.g. an IS department) as independent of one another as five replications each of which is conducted by a different researcher or research group? Probably not, Rosenthal (1990), for example, points out that in psychology the correlation of replicators that know each other (e.g. colleagues or a professor and her Ph.D. student) refers directly to a correlation of attributes. That is, the way colleagues think about the phenomenon investigated is often highly positively correlated. As a consequence, a correlation of data these investigators will obtain from their subjects in the experiments can be observed, too. The problem even gets worse, if one supposes that only one single researcher conducts one or more replications of her own original study.

However, the replication of a case study is not as simple as the replication of a laboratory experiment. Usually, when an experimental study is published, a lot of detailed information is available on the context of the study. That is, details about the task and the experimental conditions are provided. In addition, usually an entire section of an article is dedicated to the description of the method. To be more precise: Both the data collection and data analysis method(s) are normally outlined in detail. Hence, an independent investigator is able to conduct a pure replication (see Table 1).

In contrast, in case study research a pure replication conducted by an independent investigator is seldom possible because a detailed description of the case study context as well as the methods used is hardly available in the case study articles (Dubé & Paré 2003). However, this fact should not solely be attributed to the missing rigor of case researchers. Rather, there is always an ever-existing tension between the desire for detail and the need for brevity. Especially in case study research, when the data being analyzed are primarily words, not numbers that can be summarized easily in the form of statistical values, more space is needed to explain the methodology, results, and criteria for evaluating those results (Dubé & Paré 2003). Hence, a major question is whether high-quality, rigorous case research is amenable to the journal format (Inui & Frankel 1991).

Yin (2003) points out that in the past, case study research procedures have often been poorly documented, making external reviewers suspicious of the reliability of the case study. Hence, case study results often seem to be highly dependent on the characteristics of the researcher. However, a major point of positivist research in general and positivist case study research in particular is to remove the researcher as a variable. Consequently, the subjectivity of the research process is then low. Importantly, Yin (2003) developed tactics to overcome the subjectivity of the research process. He proposes to use a case study protocol to deal with the documentation problem. Additionally, the development of a case study database can positively affect the reliability of case study research.

However, to come back to the initial question—By whom is the replication conducted?—it is obvious that a pure replication (see Table 1) can be conducted more easily by a researcher who has already performed the original study, because all relevant information concerning the context of the case as well as the methods applied is available (in more detail). Finally, considering Yin's (2003) case study concept, the term "multiple-case design" is used to describe situations, in which a researcher who has already conducted the original study also performs the replication study.

	RS1	**RS2**	**RS3**	**RS4**	**RS5**	**RS6**	**RS7**	**RS8**
WHO conducts the replication? (Researcher)	HC	AL	AL	HC	AL	HC	HC	AL
WHAT is replicated? (Context)	HC	HC	HC	HC	AL	AL	AL	AL
HOW is the replication conducted? (Methods)	HC	HC	AL	AL	AL	AL	HC	HC
Objective of the replication study	Reproducibility				Generalization			
Type of replication study	Pure replication		Extension					

RS = Replication Strategy, HC = Held Constant (in the replication study), AL = Altered (in the replication study)

Table 1. Replication Strategies.

4.2 Context

A replication study can be either performed in the same context or in a different context. In the IS discipline "same" context usually means that the replication is conducted in an organization with the same characteristics like the organization of the original study. Obviously, in the IS discipline the replication study usually cannot take place in the same organization like the original study. Suppose the investigation of the implementation of an enterprise system. Barker and Frolick (2003), for instance, conducted a case study with a major soft drink bottler on enterprise resource planning (ERP) failure. If any other researcher (or Barker and Frolick themselves) had attempted to replicate the study with the same soft drink bottler one or two years after the original study, she would not have been successful. First, simply progress of time would have altered many ERP success (or failure) determinants (e.g. user acceptance). Hence, the conclusions drawn from the case study would differ from the original ones. Second, the soft drink bottler would probably not have been at the researcher's disposal a second time because the time invested in data collection (e.g. interviews) does not yield any profit for the case organization. Especially when a longitudinal design is chosen by the researcher (that implies a time-consuming research process), it is likely that the willingness to be case object is low in many organizations.

Considering all these arguments, it is an important issue to determine a detailed specification of those context characteristics that should be held constant in the replication study. Basically, which

characteristics should be held constant depends on the theory investigated. Suppose, for example, a case researcher who wants to investigate (i) the process of negotiating an IT outsourcing contract and (ii) the output of the negotiating phase (i.e., the structure of the contract). No doubt, the researcher should pay attention to the careful selection of the organizations (service provider and client firm) in which the replication study is taking place. If the objective of the replication is to reproduce the findings of the original study (see Table 1), and if the original study investigated negotiations between a multinational IT service company and a multinational client company, then it is clear that the replication study has to investigate the negotiation between two *multinational* firms, too—i.e., one major context factor, namely "multinational", is held constant. If the replication investigated a negotiation between a multinational service provider and a small firm employing only 30 persons, it would not be possible to determine whether a diverging negotiation process (in comparison to the original study) as well as varying contract structures are caused by an absence of a relationship between the independent variable (bargaining power) and the dependent variables (negotiation process and contract structure, respectively), or due to methodological issues (e.g. interviewer bias, see, for example, Salazar 1990).

However, despite the above-mentioned fact that a replication study usually cannot take place in the same organization like the original study, performing a replication in the same organization at a different point of time can be desirable. Consider, for example, if centralization/decentralization was a relevant variable in a theory and an organization changes from centralized to decentralized structure, this would be an ideal basis for conducting a theoretical replication. Additionally, whether a replication concerns reproducing or generalizing the original findings highly depends on the characteristics on which the cases differ and the theory being tested. For instance, if cases differ on a theory relevant variable (e.g. centralization/decentralization) this can be regarded as reproducing the original finding. That is, decentralization leads to X, in contrast, centralization does not lead to X. However, if the second case differs on a non-theory relevant variable, i.e., a context variable (e.g. industry, organization size, or country) it can be considered as establishing generalizability in a general sense.

4.3 Methods

A replication study can be either conducted by using the same methods for data collection and analysis or by using a different set of methods. Several data collection methods (e.g. interviews, documentation, observation, or questionnaires) can be used in case research. Dubé and Paré (2003) found that interviews (102 out of 107 case studies that elucidated data collection) and documentation (68 out of 107) are the predominant data collection methods, while the usage of observation (34 out of 107) and questionnaires (29 out of 107) is much lower. However, the interesting aspect that can be deduced from this data is that many researchers hardly use multiple methods for the analysis of a particular case. Unfortunately, many empirical phenomena cannot be analyzed properly by using only one data collection method. Importantly, Yin (2003) points out that using only one data collection method negatively affects construct validity.

As already mentioned, following agency theory, contractual outsourcing arrangements that are not outcome based favour opportunistic behavior by the agent. If, for example, a researcher wants to investigate opportunistic behavior, it is possible to simply conduct interviews in the provider organization and in the client organization. Unfortunately, one the one hand, employees in the provider organization will certainly not admit opportunistic behavior (even if it does happen) and, on the other hand, employees in the client organization will state the opposite. Hence, the researcher is confronted with a situation in which there is testimony against testimony. In such situations, it is therefore important in order to get reliable data to use another data collection method.

One manifestation of agent opportunism is the assignment of poor qualified service staff (Geoffrey 1998). Suppose that a hardly qualified trouble shooter calls on a customer. Clearly, the time to solve the problem will be longer in comparison to a highly qualified employee. Therefore, the invoice

amount to be paid for the service will be higher and this fact impairs the economic situation of the client firm. However, by analyzing documentations (for example the invoice on which the type of problem as well as the time to solve the problem are stated) opportunistic behavior can be found out (because a computer specialist is able to assess the reasonableness of the invoice amount). The example shows that empirical phenomena (here opportunism) sometimes cannot be investigated reliably with only one data collection method (here interviews). Rather, other methods (here analysis of documentations such as invoices) have to be applied in order to allow for a reliable measurement.

The type of analysis method to be applied is highly determined by the data collection method used. Considering the fact that in case study research a vast amount of data is usually only available in qualitative format, the central role of coding becomes visible. Coding data and then analyzing it statistically—the methods range from simply counting frequencies to complex time series analysis using Markov processes (see, for example, Yin 2003)—is of particular importance for positivist case research (Dubé & Paré 2003). However, usually case researchers do not publish their coding scheme and as a result it is not possible for other researchers to conduct a pure replication (see Table 1).

To sum up, considering the fact that the researcher, the context, and/or the methods can be altered in a replication study, eight (2x2x2) different case study replication strategies can be pursued (Table 1). As indicated in Table 1, the objective of replications in which the context is held constant is the reproducibility of the theory that was already investigated in the original study. In contrast, in those situations in which the context is altered, the research objective is to generalize the theory explored. Furthermore, in those cases where both context and methods are held constant in the replication study, a so-called pure replication is conducted. As already outlined, in contrast to natural sciences the chance to get pure replications published (at least in top journals) is low. Therefore, case researcher can also focus on extension studies that are characterized by the fact that at least one of the two factors "context" or "methods" is altered in the replication study.

5 CONCLUSION

This article has outlined the importance of replication research. In particular, scholars can use one of the eight strategies shown in Table 1 to replicate their own case studies or case studies performed by other researchers. Importantly, one of the most essential messages of this article is the fact that multiple cases should be considered as one would consider multiple experiments or surveys, and explicitly not as one would consider multiple subjects in an experiment or respondents in a survey (Yin 2003). Because only if one accepts this view, it is possible to understand the value of positivist case study replication research for scientific progress.

Finally, it is crucial to address the fact that, one the one hand, there is a reasonable quantity of case studies that were replicated by those researchers who already carried out the original study (Dubé & Paré 2003). On the other hand, however, an online-investigation in the SCOPUS database <http://www.scopus.com/scopus/home.url> on 10th of November, 2006, resulted in the identification of only one article that explicitly replicated a case study of a *different* researcher. Interestingly, this article deals with a topic that is of high relevance for the IS discipline. Dinh-Trong and Bieman (2005, pp. 481/482/491) investigated in their replication case study an open source development project and described their research effort as follows:

> "Case studies can help to validate claims that open source software development produces higher quality software at lower cost than traditional commercial development. One problem inherent in case studies is external validity—we do not know whether or not results from one case study apply to another development project. We gain or lose confidence in case study results when similar case studies are conducted on other projects … One or two case studies cannot conclusively determine the nature of OSS [open source software; comment by the author of this article] development. There are just too many differences between application domains, project participants, project support, and project lifespan. Understanding the nature of a software process such as OSS development will require many case studies. Our objective is to obtain further evidence to help determine whether or not the hypotheses represent general rules by examining other open source

> systems … The goal of this study … to better understand the nature of Open Source software development and to see if prior case study results can be replicated in a study of another system".

If one wants to classify Dinh-Trong and Bieman's (2005) replication case study by applying the classification scheme shown in Table 1, then it becomes obvious that they used replication strategy number 8 (RS8). The authors replicated a case study originally performed by Mockus et al. (2002), who investigated key requirements for the success of an open source software endeavor in two different case projects (Apache and Mozilla). Dinh-Trong and Bieman (2005) altered the case context (to FreeBSD project, an open source version of the Unix operating system) but used the same methods for data collection and analysis as Mockus et al. (2002). Hence, this replication case study constitutes a generalization and extension study. Dinh-Trong and Bieman (2005, p. 491) found that the data supported three hypothesis of the original study, one hypothesis was not supported and two other hypothesis were revised. In general, case study replication articles are currently hardly available in the IS literature (apart from studies using Yin's multiple-case design). To the knowledge of the author of this article, only Berthon et al. (2002) outline several sample articles for replication research in the IS discipline; however, these articles do not exclusively refer to case study research but to empirical research in general.

To sum up, in the introduction it was suggested that we would probably all agree that a person who concluded that all cats are black after seeing only one cat that happened to be black, was unintelligent. Keeping at this analogy, if the case results of Mockus et al. (2002) are the black cat, then the case results of Dinh-Trong and Bieman (2005) are a black-white striped cat. However, for scientific progress it is important that again and again scholars attempt to find another black or black-white striped cat—as long as a white cat is found.

Acknowledgments

I would like to thank Jens Dibbern for his valuable comments on an earlier version of this article. Additionally, I am grateful to two anonymous reviewers whose comments have improved this article considerably.

References

Amir, Y. and Sharon, I. (1990). Replication Research: A "Must" for the Scientific Advancement of Psychology. In Neuliep, J.W. (Ed.). Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 1990, 5 (4), 51-69.

Barker, T. and Frolick, M.N. (2003). ERP implementation failure: A Case study. Information Systems Management, 20 (4), 43-49.

Benbasat, I., Goldstein, D.K., and Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. MIS Quarterly, 11 (3), 368-386.

Berthon, P., Pitt, L., Ewing, M., and Carr, C.L. (2002). Potential Research Space in MIS: A Framework for Envisioning and Evaluating Research Replication, Extension, and Generation. Information Systems Research, 13 (4), 416-427.

Bornstein, R.F. (1990). Publication Politics, Experimenter Bias and the Replication Process in Social Science Research. In Neuliep, J.W. (Ed.). Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 1990, 5 (4), 71-81.

Bozarth, J.D. and Roberts, R.R. (1972). Signifying significant significance. American Psychologist, 27, 774-775.

Campbell, K.E. and Jackson, T.T. (1979). The role of and need for replication research in social psychology. Replications in Social Psychology, 1 (1), 3-14.

Carroll, J.M. and Swatman, P.A. (2000). Structured-case: a methodological framework for building theory in information systems research. European Journal of Information Systems, 9 (4), 235-242.

Dibbern, J., Goles, T., Hirschheim, R., and Jayatilaka, B. (2004). Information Systems Outsoucing: A Survey and Analysis of the Literature. The DATA BASE for Advances in Information Systems, 35 (4), 6-102.

Dinh-Trong, T.T. and Bieman, J.M. (2005). The FreeBSD Project: A Replication Case Study of Open Source Development. IEEE Transactions on Software Engineering, 31 (6), 481-494.

Dubé, L. and Paré, G. (2003). Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. MIS Quarterly, 27 (4), 597-635.
Eisenhardt, K.M. (1989a). Agency theory: An assessment and review. Academy of Management Review, 14 (1), 57-74.
Eisenhardt, K.M. (1989b). Building Theories from Case Study Research. Academy of Management Review, 14 (4), 532-550.
Gable, G.G. (1994). Integrating case study and survey research methods: an example in information systems. European Journal of Information Systems, 3 (2), 112-126.
Geoffrey, J. (1998). Consultants: IT Saviors or Slick Sellers? Upside, 10 (10), 110-120.
Greenwald, A.G. (1975). Consequences of prejudice against the null hypothesis. Psychological Bulletin, 82, 1-20.
Hendrick, C. (1990). Replications, Strict Replications, and Conceptual Replications: Are They Important? In Neuliep, J.W. (Ed.). Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 1990, 5 (4), 41-49.
Herriott, R.E. and Firestone, W.A. (1983). Multisite qualitative policy research: Optimizing description and generalizability. Educational Researcher, 12, 14-19.
Hubbard, R. and Vetter, D.E. (1996). An Empirical Comparison of Published Replication Research in Accounting, Economics, Finance, Management, and Marketing. Journal of Business Research, 35 (2), 153-164.
Inui, T.S., and Frankel, R.M. (1991). Evaluating the Quality of Qualitative Research: A Proposal Pro-Tem. Journal of General Internal Medicine, 6 (5), 485-486.
Kuhn, T.S. (1962). The structure of scientific revolutions. University of Chicago Press, Chicago.
Kupfersmid, J. (1988). Improving what is published. American Psychologist, 43, 635-642.
Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.). Criticism and the growth of knowledge (pp. 91-196). Cambridge University Press, Cambridge.
Lee, A.S. (1989). A Scientific Methodology for MIS Case Studies. MIS Quarterly, 13 (1), 32-50.
Lee, A.S. (1991). Integrating positivist and interpretive approaches to organizational research. Organization Science, 2 (4), 342-365.
McClintock, C.C., Brannon, D., and Maynard-Moody, S. (1979). Applying the Logic of Sample Surveys to Qualitative Case Studies: The Case Cluster Method. Administrative Science Quarterly, 24 (4), 612-629.
Mockus, A., Fielding, T., and Herbsleb, D. (2002). Two Case Studies of Open Source Software Development: Apache and Mozilla. ACM Transactions Software Engineering and Methodology, 11 (3), 309-346.
Neuliep, J. W. and Crandall, R. (1990). Editorial Bias Against Replication Research. In Neuliep, J.W. (Ed.). Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 1990, 5 (4), 85-90.
Orlikowski, W.J. and Baroudi, J.J. (1991). Studying Information Technology in Organizations: Research Approaches and Assumptions. Information Systems Research, 2 (1), 1-28.
Pearson, K. (1902). On the mathematical theory of errors of judgment with special reference on the personal equation. Philosophical Transactions of the Royal Society of London, 198, 235-299.
Popper, K.R. (1972). Conjectures and refutations. Harper, New York.
Rosenthal, R. (1990). Replication in Behavioral Research. In Neuliep, J.W. (Ed.). Handbook of replication research in the behavioral and social sciences. [Special Issue.] Journal of Social Behavior and Personality, 1990, 5 (4), 1-30.
Salazar, M.K. (1990). Interviewer Bias: How it affects survey research. AAOHN Journal, 38 (12), 567-572.
Sturm, R., Morris, W., and Jander, M. (2000). Foundations of Service Level Management. Sams, Indianapolis.
Walsham, G. (1995a). The emergence of interpretivism in IS research. Information Systems Research, 6 (4), 376-394.
Walsham, G. (1995b). Interpretive case studies in IS research: nature and method. European Journal of Information Systems, 4 (2), 74-81.
Walsham, G. (2006). Doing interpretive research. European Journal of Information Systems, 15 (3), 320-330.
Yin, R.K. (2003). Case Study Research: Design and Methods. Third Edition. Sage Publications, Beverly Hills.
Yin, R.K. and Heald, K.A. (1975). Using the Case Survey Method to Analyze Policy Studies. Administrative Science Quarterly, 20 (3), 371-381.
Zmud, R.W. (1995). Editor's Comments. MIS Quarterly, 19 (4).