Generalising from Case Studies

Emma S. Wikfeldt
Halmstad University

ABSTRACT

The generalisability of case study findings is heavily criticised in the scientific community. This study attempts to answer to what extent generalisation is possible, through a literature review. Resources were collected by searching in databases and in reference lists. A presentation of arguments from both sides will follow, finding that generalisation is possible to almost the same extent as quantitative research, if done correctly and carefully, with great concern and accuracy.

Keywords:
  generalise, external validity, case study, small-N

In the scientific world, discord regarding methodology is inevitable. Proponents of the qualitative and the quantitative methods respectively, expresses clear differences in opinion when the question of validity is put on the table. Arguments against quantitative research consider its superficiality, whereas assertions against qualitative research stems from a concern for reliability. However; when it comes to case studies, perhaps the most criticised element is generalisability (Firestone, 1993; Firestone & Herriott, 1983; Gerring, 2007; Hägg & Hedlund, 1979; Sharp, 1998; Tellis, 1997; Woodside, 2010). Another element of criticism is external validity, with which generalisability is frequently associated (Byrne & Ragin, 2009; Huberman & Miles, 2002; Robinson & Norris, 2001).

A case study is the act of extensively studying something in particular, with the aim of amplifying one’s understanding of said thing (Ruddin, 2006). Baxter & Jack (2008) continues this description:

Case study research is more than simply conducting research on a single individual or situation. This approach has the potential to deal with simple through complex situations. It enables the researcher to answer “how” and “why” type questions, while taking into consideration how a phenomenon is influenced by the context within which it is situated. (p. 556)
Case studies usually include a small sample of participants in a relatively small controlled environment (Tsang, 2014). Yet another description is presented by Simons (2015):

Case study is the study of the singular, the particular, the unique, whether that single case is a person, a project, an institution, a programme or a policy. (p. 175)

It can also be categorised as an intensive study (contrary to an extensive study) meaning it investigates a small sample but a multitude of variables (Jacobsen, 2002). Intensive studies are well placed in generating plenty of detailed and relevant data, which are not collected out of context. This is one of the strengths that make the intensive case study very valuable. The internal validity is very high in this type of research (Jacobsen, 2002).

Yin (2013) considers case study generalisation as “an effort to generalize from a small number of cases to a larger population of cases” (p. 325). It is an attempt “...to learn from one [case] and understand many [cases]”, says Campbell (2003, p. 15). Ruddin (2006) comments that “[w]ithout generalization, we could not interact with our world in a coherent manner ...” (p. 799). Robinson and Norris (2001) continues with saying that generalisation is part of our cognitive capabilities, and is therefore ineluctable.

Despite this, generalising case results to other contexts than the one studied, seems like a difficult task (Firestone, 1993). Results from a single case are considered by many as illegitimate for developing scientific theory. For instance, Mintzberg (2005) remarks, “If there is no generalizing beyond the data, no theory. No theory, no insight. And if no insight, why do research?” (p. 10).

Many researchers are frightened by the thought of drawing conclusions from case studies, especially studies of single cases, due to the issues stated above. Kennedy (1979) acknowledges this, and states that the statistical research has its very own set of evaluation rules, which gives a clear and objective instruction of how the data should be interpreted. The author enunciates the significance in developing rules for how to pursue inferences for case studies, since the existing guidance are perceived as enigmatic compared to the statistical one. They seem to require more subjective evaluation than what is acceptable in the scientific community (Kennedy, 1979).

Qualitative intensive research is praised for its descriptive accuracy. Even though there are actions researchers can take to amplify the external validity and replicability of their case study, “[g]eneralizability is clearly not the strength of qualitative research”, as Firestone (1993, p. 16) so excellently explains it. He continues by conveying the importance of assessing the claims of case study conclusions

precariously, since the required level of subjective estimation cannot be logically legitimised.

To what extent are generalising from case studies justified? An answer to this question is attempted here through a literature review; therefore exclusively containing secondary sources. The literature describing the theoretical framework was found by searching in databases, such as Sage, Web of Science, Science Direct, Scopus, et cetera. The pursuit for relevant references was conducted using keywords such as, among others, **generalise qualitative research** and **case study inference**. The majority of the sources are scientific papers, published in a variety of journals. Once a paper was found useful for this study, a scan of the bibliography was made to possibly find other useful sources of information within.

My choice to assemble the study in this way stems from the nature of my purpose. The purpose is to examine the already established theoretical structure, which is formed by earlier conducted scientific research. This research is most readily accessible through databases, in the form of research papers and other literature reviews. Continuing the investigation via the bibliographies of papers that was found, is a good way of finding other sources (Jesson, Matheson & Lacey, 2011). This strategy is also motivated by not wanting to waste time and other resources in the process.

Following this introduction, arguments from both sides of the spectrum will be presented, and ultimately an integral conclusion will be indicated, which will attempt to answer the question posed above.

**ARGUMENTS FOR GENERALISING**

Authors such as Flyvbjerg (2006, 2001) and Ruddin (2006) defend the concept of qualitative generalising. Both deny the inability of cases to supply means for scientific development; deeming it a misunderstanding. The reasoning avers that case inference is all too often considered synonymous with statistical inference; the latter mainly used in quantitative research (Ruddin, 2006; Yin, 2012). Striving not to confuse the two is important, Ruddin says, to avoid undermining the strength and validity of the case study results. By drawing conclusions from individual cases “[w]e do not infer things ‘from’ a case study; we impose a construction, a pattern on meaning, ‘onto’ the case …” (p. 800, emphasis added).

The difference in purpose and end goal between case studies and quantitative studies is also important to distinguish. To conduct case studies is to attempt the creation of hypotheses (Firestone, 1993; Ruddin, 2006; Woodside, 2010; Yin, 2012) rather than quantitatively stating statistical facts. Cronbach (1975) identified the
hypotheses as “working hypotheses” (p.125), meaning that every case study result found will contribute in accepting or denouncing the hypotheses, and possibly in building new ones. Kennedy (1979) says that distinctive case studies will never find a conclusive answer, but will instead only find confirming or disconfirming answers.

While statistical findings are mainly generalised to populations, cases have a tendency to generalise to other circumstances and situations, with the help of in-depth analytic investigation (Yin, 2012). As Yin says, case studies are not intended to generalise “from samples to universes” (p.18) as for instance qualitative surveys are. The claims made when generalising from cases cannot be considered as “proof” in a statistical sense. Rather, they build theoretical premises which function as tool to make assertions about situations akin to the one studied. Likewise, if further case studies show resembling outcomes, they can be said to support the hypotheses and therefore be a part of constructing the theory (Yin, 2012). The phenomenon discussed here is branded with the term “analytic generalization” (Lincoln & Guba, 2000, p. 112; Yin, 2012, p. 18).

“It is important to realize that nonstatistical [sic] arguments need not be invalid”. That is what Kennedy (1979, p. 664) claims. He further on substantiates the importance of single case studies. An evaluator will pick and choose from a multitude of cases based on how well they resemble his or her own situation. This makes the single case study very valuable, he says. However, as stated in the introduction, Kennedy deems the rules of qualitative results evaluation to be ambiguous, and has thus presented a number of rules to abide by for optimal objective generalisation.

Firstly, one must assure that one examines an ample range of attributes within the case(s). Secondly, common and/or similar attributes between the different samples must be abundant. This helps the generalisability of the results. Thirdly, unique attributes between the samples must be few. Unique attributes namely inhibits the study’s generalisability. Lastly, the attributes ought to be relevant. This could be achieved by evaluation; which contains establishing what attributes can be directly connected to the research topic (e.g. a child’s ability to read), what attributes are indirectly linked to the preceding attributes, based on previous experience (e.g. the parent’s education), and what other scientists have contemplated to be relevant, not necessarily supported by evidence.

Kennedy expressed above his defence for single- or few-case studies. Meanwhile, those studies are often met with antagonism, which commonly has been labelled the small-N (Gerring, 2007).

Additionally, Tellis (1997) reports that a single case can be very useful in characterising an extreme case; that is, a case finding that takes the theory premises to an extreme level. The extreme case could be one which either confirms or denounces the underlying theory. This is a point that Yin (2012) continues to elaborate on. The
case findings may corroborate or completely reject the theoretical baseline, or it may create brand new hypothesis that commences the construction of new frameworks.

Furthermore, authors present the idea of a critical case (Flyvbjerg, 2006; Firestone, 1993), which shows researchers a way to more credibly generalise. Flyvbjerg describes the phenomenon such as:

“If it is valid for this case, it is valid for all (or many) cases.” In its negative form, the generalization would be, “If it is not valid for this case, then it is not valid for any (or only few) cases.” (p. 230)

A classic example of a critical case is the “black swan”, instated by Karl Popper in 1982. Even though Popper did not label his theory as such, it nevertheless holds the characteristics of a critical case. He firstly put forth the statement “There is a white swan”, which is the result of discovering a swan and observing its whiteness. If one were to inductively evolve the assessment, we get the inference “All swans are white”, which makes the original case observation a critical one. Yet, it is consequently falsifiable, since a single observation of a swan that is not white would make the statement deductively untrue. If one would observe the presence of a black swan, the observation would instantiate the existence of a mistake in the previously made assessment.

A CRITICAL PERSPECTIVE ON GENERALISATION

Arch Woodside (2010, p.21) will inaugurate this section with his diagram, portrayed below.

![Figure 1: Research methods' attainment of objectives.](image)
The figure depicts case study research as one with high accuracy and low generalisability. It is also evident that a statistical quantitative survey is on the opposite side – with the potential of generating high generalisability, but may be at risk for displaying low accuracy in the results. Authors such as Oskamp (1965) claims that researchers regularly are overconfident of the generalisability (and accuracy) of their findings.

Furthermore, the English poet William Blake astutely expressed in 1808 his view on generalisation:

To Generalize is to be an Idiot. To Particularize is the Alone Distinction of Merit — General Knowledges are those Knowledges that Idiots possess. (www.izquotes.com)

Blake indicates that the act of particularisation is one to prefer over generalisation, and so do many others. Lincoln and Guba (2000) claim that: “The trouble with generalizations is that they don’t apply to particulars” (p.27). A well-known metaphorical joke regarding statistical research may also help to enunciate this position:

Did you hear about the statistician who had his head in an oven and his feet in a bucket of ice? When asked how he felt, he replied, "On the average I feel just fine." (www.maths.usyd.edu.au)

The joke, although somewhat hyperbolic, plays on the tendency of statistical research to miss out on important, unique, and particular cases when they generalise data.

Continuously, Gomm et. al. (2000) says that for an analytic generalisation of a case study to be good, the analyst must acknowledge all different and similar factors between previously conducted case studies and the currently studied one. However, due to the flaws and limitations of the human brain, such as being prone to habitual thinking, it is easy to disregard important differences and similarities in the investigation, meaning that a logically reasoned discernment is unlikely to occur. The authors continue, arguing that good generalisations in case studies are often insular, and can therefore give a narrow-minded view of the general circumstances. Knowledge that is too general is thus contributing to an oversimplified worldview (ibid.).

Cronbach (1975) also takes a critical approach to generalisation, arguing that generalisation shouldn’t be the main goal of scientific research. When a scientist collects data and analyses it, he acknowledges both controlled variables and uncontrolled variables; the latter containing situational circumstances, personal characteristics, and particular events – all unique for the situation. Therefore, every
time the findings are to be applied to a different situation, it needs to be carefully evaluated almost as extensively as the first time; thus making generalisation in the original study redundant.

Campbell & Stanley (1966) express their view of generalisation by deeming it unjustified. They state it is an extrapolation of the results found; meaning one makes a subjective judgement while generalising and merely assumes that the relevant axioms are true. “Logically …. we cannot generalize at all. But we do attempt generalization by guessing...” (p. 17). When we generalise from case studies, we make presumptions about unproven laws, according to the authors; hence we are not objective.

Firestone and Herriott (1983) are clear proponents of multisite case studies; that is, case studies including multiple cases and cross-references between them. This is primarily to escape what they call “radical particularism” and hence make the study gain generalisability. They argue multisite case studies provide better description and generalisability than single-case studies do. This is also embraced by Jacobsen (2002), as he labels multisite case studies as the perfect balance between generalisability and data relevancy. However, Firestone and Herriott (1983) continues with saying there is a danger in conducting cross-site comparisons, due to the possibility of differences in structure and practices between different sites – especially if the sites are located in different regions of the world. These discrepancies could contribute to incorrect and skewed results along with researcher biases (Ibid., see also Parker et. al., 2008). Therefore, Firestone and Herriott encourage standardisation between different regions and sites in an effort to eliminate these risks.

Finally, we will investigate the connection generalisability has to the concept of external validity. A study’s finding inhabits a high external validity if it is well applicable to other realms than the one itself is situated within. This definition is not far from the goal of case study generalisation; where the findings are supposed to be pertinent to other akin situations and settings. The two concepts are therefore related significantly.

Gerring (2007) brings up external validity, saying that:

To be a case of something broader than itself, the chosen case must be similar (in some respects) to a larger population. Otherwise—if it is purely idiosyncratic (unique)—it is uninformative about anything lying outside the borders of the case itself. A study based on a nonrepresentative [sic] sample has no (or very little) external validity. (p. 248)

What the author claims here is that good representativeness is essential for good external validity. A couple of years before, Gerring (2004) stated that “[s]ingle-unit designs often fall short in their representativeness...” (p.348). This means, he
continues, that the findings cannot be assumed true for considerably bigger arrays of unstudied units.

**CONCLUDING THOUGHTS**

Obvious is that differentiating views are abundant regarding this topic. Many objections towards case study generalisation seems to stem from a deliberative worry of presenting an inaccurate picture of reality; a fear that is entirely appropriate to have, since the prominent purpose of science itself is to paint an accurate picture of reality. Evidently, it is easy to misstep and therefore be unsuccessful in presenting authentic results. The rules of qualitative generalisability that Kennedy (1979) suggested could most definitely help plenty of scientists reaching better generalisability and external validity, and, perhaps equally important, help them trust the notion of case study generalisation.

Yin (2012) has conveyed that analytic generalisation is inferior to statistical generalisation when dealing with case studies. In this way, we can generalise from a few cases, or even a single case. Contrary beliefs, such as the small-N problem, declare that the very nature of the case study, with small samples, makes it invalid for generalisation. Some even declare that researchers usually are overconfident in doing their cases studies (see Oskamp, 1965); implying that generalising is a detached way of making inferences, which can do more harm than good to science and our understanding of the world. When generalising case study results, researchers must practice a high level of accuracy and cautiousness to succeed; due to the lesser extent of evaluation rules (Kennedy, 1979), the risk of the case not being representative (Gerring, 2004), and probability of biased and subjective influences interfering with the scientist in the analysis (Firestone, 1993). In short, for generalisation to be justified, it must be done carefully and with much consideration.

It is inappropriate to generalise case studies to a population, and eligible to generalise it to theory, as we have learned from Yin (2012). Therefore, it could be suggested that “All swans are white” is not a suitable inference to case research. Since it is falsifiable (see Popper, 1982), we cannot be confident that every single swan in the world has a white colour. What we *can* say is: since this case study shows that this swan is white, other cases of swans are likely to be white as well. Thus, we build a theory that states *swans are white*. When we conduct other studies on swans, we will likely get results confirming the theory’s premises; results which will reinforce the strength of the theory’s validity. It is also possible for us to find unique cases, such as swans that are black, brown, or pink, even; which in turn will undermine the theory, disconfirming it. Through the course of the research, the theory will develop, change multiple times, and perhaps be completely dismissed in the end. Ultimately, what is
going to be the determining factor of the theory’s fate is whether or not it represents reality. We might consider that it could be likened with some sort of Darwinian natural selection, where only the best corresponding theory survives. The process described just now can be connected to the notion that case study helps in building theory and hypotheses by finding confirming or disconfirming results, as stated in Cronbach (1975), Firestone (1993), Kennedy (1979), Ruddin (2006), Woodside (2010), and Yin (2012).

Another objection is that generalisation is an unsound way of stating inferences, and that, instead, we should strive to use particularisation to a greater extent. The author suggests that the unique case is valuable in its own way, in spite of being non-statistical (Kennedy, 1979). Simons (2015) says case studies are the study of the particular. William Blake states that particularisation is the one and only way of judging something of its true nature. With these statements in mind, we might be able to suggest that the case study is the optimal way of studying the world. Yet, mankind is found to constantly seek generalised answers in her surroundings. Maybe we have a need to generalise, even though it may seem to be the non-optimal decision? Such questions are yet to be answered.

In summary, the literature, although discrepant, suggests that case studies can be generalised close to the same extent as statistical studies, if done correctly. It is clear that the research on case studies is far from as sophisticated as the research on statistical studies. Therefore, this paper will end in an invigorative suggestion to continue the case study research, particularly in the field of generalisation and external validity. What makes case study results valuable? How can researchers be more realistically confident in their case study findings, and how can their strength and veracity be conveyed to others, so that more find the results trustworthy? Answers to these questions and supplementary ones are hopefully to be found as the advancement of science proliferates.

REFERENCES


