



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Building Engineering

journal homepage: [www.elsevier.com/locate/job](http://www.elsevier.com/locate/job)

# How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge?



Ray Galvin\*

Department of Architecture, University of Cambridge, 1-5 Scroope Terrace, Cambridge, UK

## ARTICLE INFO

## Article history:

Received 6 November 2014

Received in revised form

4 December 2014

Accepted 6 December 2014

Available online 24 December 2014

## Keywords:

Qualitative data

Semi-structured interviews

Building and energy consumer research

## ABSTRACT

Research in building energy consumption often uses semi-structured interviews to produce qualitative data on consumer beliefs, attitudes, practices and skills. A survey of 54 recent papers in six prominent building and energy journals shows that the samples are typically small, but inferences are often made for interventions in the light of the findings, on the assumption that these are somehow transferable to wider populations. It is often asked 'how many interviews are enough' to produce reliable results. Theoretical literature on this theme has avoided a straightforward statistical critique, and justified the practice with appeals to precedent, the special nature of qualitative personal data, and a limited pool of empirical work. This paper reviews this literature and presents a statistical approach, based on binomial logic, to critiquing and supporting the practice of semi-structured interview research in the building and energy field. The approach developed offers a set of straightforward criteria which researchers can use to estimate the reliability of their findings and inferences from the qualitative data produced in semi-structured interviews.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

It has become common in recent years to gather data on human attitudes and behaviour in building energy research through interviews with building occupants and other relevant actors. Examples of this kind of research are found in the main journals which deal with technical aspects of energy in buildings, including *Building and Environment* [1–8], *Energy and Buildings* [9–21], *Energy Efficiency* [22–33], *Energy Policy* [34–45], *Sustainable Cities and Society* [46–49] and *Building Research and Information* [50–63]. Generally in such projects, interviews are recorded, transcribed, and analysed according to a coding method which identifies key themes or concepts which bear closely on the projects' research questions.

The data so gathered is generally claimed to be 'qualitative' rather than 'quantitative'. It provides information about what sorts of relevant things are happening and how, rather than the number, proportion or magnitude of relevant factors. For example, [11] observed hundreds of buildings in a medium-sized city to estimate the proportions of dwellings where certain occupant behaviours were impeding energy-efficient ventilation. This was *quantitative* data. The author of the study then interviewed 21 households in the city to find out what sort of practices, attitudes and difficulties occupants claimed were determining their ventilation practices.

This was *qualitative* data. It gave no information about the prevalence of any particular attitude or practice within the city, but offered clues as to what these attitudes and practices were, and how they were situated within other related discourse and practice.

A sample of 21 represents a small number of households in a city of 120,000 homes. The question arises: how many interviews are enough? What is the minimum number of interviews required in a study in the field of energy consumption in buildings, to provide reliable information on the qualitative features of human attitude, practice or behaviour that bear on such studies' research questions?

Researchers have to decide how much time and funding to invest in data gathering. Most qualitative interview data is recorded, transcribed and meticulously analysed, often by several researchers to minimise bias. Interview format usually consists of prompts which lead interviewees to speak widely, on and around the topic area, to allow for new ideas and insights which the researcher would not have otherwise thought of. This demands different research skills from the technical, engineering-based skills which building energy researchers are often trained in, yet this technical background knowledge is also an important ingredient in enabling the interviewer to pick up on leads which arise in the interview.

The research also has to be credible for a critical, academic audience. Such an audience can rightly ask the following: were

\* Tel.: +44 7758 832 415.

E-mail address: [rg445@cam.ac.uk](mailto:rg445@cam.ac.uk)

there sufficient interviews to ensure a full coverage of issues; are the issues that arose in such a small sample truly representative of the relevant population; what are the limitations of the methodology and how can these be quantified?

This paper investigates this question from a statistical point of view. To the author's knowledge no such study has yet been offered in any sphere of science in which qualitative interviews are used. Those which address the issue avoid the question of a straightforward, statistical analysis of the validity of small sample, qualitative data and tend instead to rely on precedents from studies conducted in past years. This is interesting because in quantitative statistical analysis, where large samples are generally used, sample size is routinely calculated or, if response numbers cannot be controlled, the limits of the study's statistical power are calculated. In this field it is accepted that sample size depends on factors such as reference population, kind of analysis, confidence level required, sampling technique etc., as well as the type of survey and how it is conducted. There are clear mathematical rules. These seem to be lacking for small sample qualitative interviews.

The paper is structured as follows. Section 2 identifies the main types of information which qualitative studies in energy and buildings seek to produce. Section 3 reviews social science literature which has attempted to address the question as to how many interviews are enough. Section 4 offers a statistical approach to the reliability of small sample qualitative data. Section 5 offers discussion, conclusions and recommendations for research practice and further research.

## 2. Qualitative studies in building and energy research

### 2.1. The studies considered in this study

In order to ascertain the scope of the types of information which researchers attempt to obtain from interviews in the field of energy consumption in buildings, papers in this field which use interviews were examined in six academic journals in which such papers are frequently published (see list in Section 1). In all but one of these journals an attempt was made to identify all the papers which have used this methodology in this field over the past 10 years. As the sixth journal, *Building Research and Information*, publishes a greater number of such papers, a random selection of 13 such papers was made. The number 13 is arbitrary but was chosen because this was the highest number of such papers found in any of the other journals. This gave a total of 63 papers.

Most of these papers deal with specificities of consumption behaviour and attitudes in buildings and thereby report on interviews with occupants, though some report on interviews with relevant experts in the building industry. It was difficult to draw a clear line between these and policy-focused papers, but as a rule of thumb only papers dealing with consumption and occupant behaviour and/or attitudes were included. This is by no means an exhaustive list, as there are other journals which publish papers on similar issues, hence some approaches may not be represented in this selection. It is noteworthy that the last (63rd) paper to be selected, i.e. [59], used interviews for an aim that had not been found in any other paper. This raises the possibility that not all approaches will be fairly represented in this analysis, as it is conceivable that a novel approach might have appeared in the 100th or 500th paper if the search had been extended. This point is not trivial because it also bears on the question of 'how many' items must be examined to get a full (or 'saturated') picture of a data field – the issue which forms the substance of this paper.

9 of the 63 papers either did not record their interviews or were unclear about this. Most of these 9 reported a very high number of interviews, mostly by telephone, and appeared to have used 'interviews' in order to complete pro forma questionnaires. These were excluded from the remainder of the analysis.<sup>1</sup> The remaining 54 all used a 'semi-structured' approach, in which key starter questions were asked or prompts were given, so that respondents could talk freely around the themes and possibly introduce ideas or claims which had not been anticipated by the interviewer.

### 2.2. Types of findings in these papers

There were three main types of findings researchers were looking for in these papers, defined here as thematic, statistical and discursive. 'Thematic' findings have to do with *what* is happening. Here the interviews seek to discover what specific behaviours, attitudes, beliefs, practices, skills, situations, or combinations of these, are represented among their interviewees. This can be further divided into two branches: 'grounded theory' as defined by [73], in which the researcher is seeking to identify every possible theme or combination of themes which might emerge, and an approach based more directly on the investigators' research questions, which more precisely define the scope of the things being looked for.

The issue of 'saturation' emerged repeatedly in these papers. Saturation is defined in this literature as the point at which no new relevant information is forthcoming, even if more people are interviewed. Most of these papers claimed to have reached saturation by using just the number of interviews they conducted, or to have conducted more than were necessary for saturation to occur.

From a statistical point of view the outcomes of this type of research can be classified as binary (or binomial), in that each outcome is either found or it is not found. Hence the overriding research question of this paper is as follows: *How many interviews are needed, to ensure that all the relevant themes which are present in the relevant population have emerged in the interviews, given that a theme may be present in the population from which interviewees are selected, but not have (yet) emerged in the interviews conducted so far.*

A more precisely targeted research question therefore arises: if a particular theme is present in proportion  $R$  of the population, what is the probability that it will emerge at least once within a given number  $n$  of interviews?

The corollary of this question is: if a particular theme is present in proportion  $R$  of the population, how many interviews are required in order for there to be a 95% probability<sup>2</sup> of it emerging within these interviews?

The second type of finding sought in these papers is more overtly statistical, namely the proportion of the interviewees who attest to a theme that has emerged (e.g. [9,18,57]). At this point these are not strictly qualitative studies (though all those found here claim to be), as they seek to *quantify* their findings.

From a formal statistical point of view, this is still a binomial issue but more complex than that outlined above. The relevant research question is: *If a theme is found in proportion  $p$  of  $n$  interviews, what is the 95% confidence interval for the proportion of the population in which this theme is found?*

<sup>1</sup> These are the papers referenced [4,23,25,26,30,31,33,46,47] in the bibliography.

<sup>2</sup> The figure 95% is chosen because this is a commonly accepted level of statistical confidence for social science research. In significance testing it corresponds to a  $p$ -value of 0.05.

Download English Version:

<https://daneshyari.com/en/article/283821>

Download Persian Version:

<https://daneshyari.com/article/283821>

[Daneshyari.com](https://daneshyari.com)