ISSUES WITH DATA COLLECTION METHODS IN CONSTRUCTION MANAGEMENT RESEARCH

Kate Carter[1] and Chris Fortune[1]

[1]*School of the Built Environment, Heriot-Watt University, Edinburgh EH14 4AS*

The effectiveness of data collection is vital to the overall quality of research. A review of data collection methods was carried out on construction management research to establish trends. A survey was administered to a sample of 650 housing associations using two data collection tools, the traditional postal survey and a web-based survey. Response rates, and the dimensions of time and cost were compared to measure the effectiveness of each method. The web is a relatively untapped resource for construction management research. Literature on web surveys argues the advantages in terms of reduced time and cost and potentially higher response rates. It is suggested that it could assist in making data available quicker, cheaper and in greater quantities. This can only be a benefit to research. A review of ARCOM proceedings, Refereed Journals and Postgraduate research shows limited utilisation of the web as a research tool. The range of data collection methods commonly adopted in both quantitative and qualitative research was identified. There is a common theme of low response rates which may lead to less than rigorous analysis. The results of the survey comparison illustrate the differences between a traditional approach to data collection and the use of modern technology. There are concerns in the use of the web in research. Sample selection and traceability become less controllable. Access to the web is traditionally seen as a limitation to participation. These factors are being addressed by the new web technology and obstacles to the use of the web are slowly being removed. The approach to data collection is fundamental to the conclusions that may be drawn from a piece of research. In understanding the mechanisms associated with data collection researchers are able to use modern technology to take the drudgery out of the process. Potentially more time can be spent in designing research and analysing the results than is typically spent in collecting data.

Keywords: data collection, research methods, quantitative and qualitative research.

THE PIVOTAL ROLE OF DATA IN EMPIRICAL RESEARCH

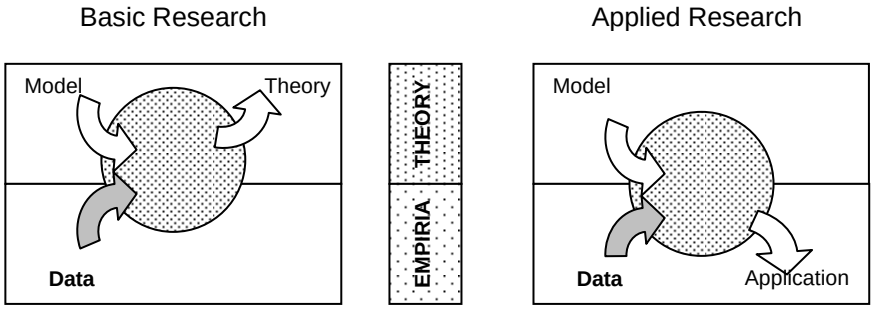
Empirical research involves the observation of real world experiences, evidence and information. In a research context this evidence and information is referred to as ‘data’. On its own data has no real meaning. It is only when it is interpreted that meaning can be derived. Empirical research relies on the existence of a research question, data and the analysis of that data. The question must be capable of being researched or answered with data (Punch 1998). The validity and the quality of data are important concepts sometimes not given adequate attention. The quality of data is in a direct relationship with the quality of the research. Poor quality data will lead to poor quality research.

There is a process of constructive alignment between data and the research concepts that must be observed when designing the method of data collection (Figure 1). This alignment and its success or otherwise underpins the quality of the research. This

[1] email@university.ac.uk

paper intends to discuss the pivotal role data plays in the research process and the ways in which it is collected.

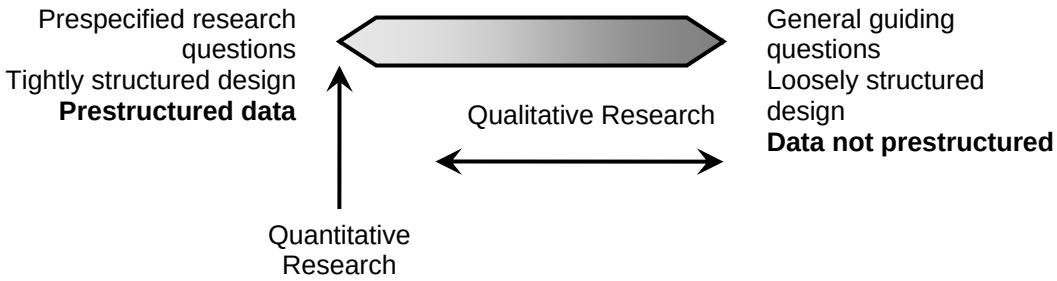
Figure 1: Position of data in the research process



The nature of data in research is directly related to the philosophical viewpoint of the research. Locke (1649) one of the founders of modern day empiricism stated that "*No man's knowledge here can go beyond his experience*". Empirical research is founded on the assertion that knowledge may only be gained through experience and the induction of that experience. It is this experience that is interpreted in the form of research data.

Punch (1998) describes research as lying on a continuum between pre-specified and unfolding (see Figure 2). Data ranges from prestructured to not prestructured. The data may be quantitative or qualitative but the presence of data is an essential part of empirical research. Typically quantitative data would be found to the far left of this continuum while qualitative data occupies a much greater range.

Figure 2: The nature of data in the research continuum (from Punch 1998)



The concept of quantitative data is one of quantity, and it is expressed numerically. Table 2 is an example of quantitative data. The use of numbers brings a structure to data and essentially involves the use of measurement, either counting or scaling (i.e. 0% to 100%). The main problem associated with quantitative data is that of adequate measurement.

Qualitative data is empirical information that is not numerical. It can lie anywhere along the continuum from prestructured to not prestructured and takes the form of people's words or the researcher's description of observation or experience (Sapsford and Jupp 1996). Mason (1996) argues that qualitative data is generated rather than collected. Interviews, documents, visual images can all be used as a source of data, but it is the researcher's epistemological position that determines how that data is generated.

There are many methods to collect data. It is important that the most appropriate method is selected for a particular piece of research. Some methods of data collection or generation are set out in table 1.

Table 1: Methods of gathering empirical data (from UIAH 2004)

	Explorative research	**Research as Revision of a Model**	**Hypothesis-based Study**
The study of **inanimate objects**	Documenting objects	Gathering Data for Analysis	Experiment
Methods for **observing** people, animals or objects	Non-systematic observation	Systematic observation	Experiment "Staged" (simulated) incident observation.
Methods for **asking questions**	Focused interview	Questionnaire	Experiment The role-playing method.
The study of **documents and secondary material**	Hermeneutical analysis of letters, Conversation sampling	Study of letters and other documents	Indirect study (e.g. of deposits or wearing). Ex post facto - study of existing files

A study carried out by EIRASS into the effects of data collection methods identified factors that influence data quality and validity (Ettema *et al*. 1996). Type of population, sample control, non-response, type of questions, complexity of questionnaire and available resources are some features affecting the value of the data. The study noted that there is limited research into data quality. There is also an increasing sophistication in model development coupled with the use of data which is not being critically assessed. It is clear that careful consideration must be given to data collection and how it fits into the overall research process.

The first section of this paper will examine the role that data plays in construction management research. The survey in terms of data collection and the vast resource of the internet will be considered for its impact on data collection. A comparison of traditional and web-based survey techniques is used to discuss the benefits and problems associated with the internet as a data collection tool. The paper goes on to discuss the possibilities for qualitative research presented by the internet and concludes with some thoughts on the importance of well thought out data collection.

APPROACHES TO DATA COLLECTION IN CONSTRUCTION MANAGEMENT RESEARCH

Loosemore, Hall and Dainty (1996) conducted a survey of publications in the refereed journal *Construction Management and Economics* between 1983 and 1993. This revealed a predominance of quantitative data collection and analysis in construction management research. 57% of the articles published used a quantitative methodological approach. Only 8% were based on qualitative research and 13% used a mixed methodology. The remaining papers were classified as “non-research” papers. Analysis of papers published recently was carried out to establish the change in approach to data collection over the last ten years. The analysis was conducted using the framework suggested by Bryman (1992) and employed in the study by Loosemore *et al.* This framework classifies quantitative data collection as methodologies using experimentation, surveys, structured interviews or questionnaires. Observation,

unstructured and semi-structured interviews, diaries, projective techniques, verbal protocol, documentary inspection and unstructured questionnaires were all classed as qualitative data collection. Paper in *ARCOM Proceedings* and postgraduate construction management research at Heriot-Watt University were analysed. The results are found in Table 1. There has been an increase in qualitative research and the classification of research as discussion or mixed methodology has increased. This suggests a greater but not absolute confidence in qualitative research methods and the use of qualitative data.

Table 1: Research Approaches in CME, ARCOM Proceedings and HWU Postgraduate Research 2001-2003

	CME 1983-1993 (%)	**ARCOM 2000**	**ARCOM 2001**	**HWU 2001-2003 (%)**
Discussion papers	22	45	38	9
Quantitative methodology	57	29	28	42
Qualitative methodology	8	6	19	18
Mixed methodology	13	20	15	31
Total	100	100	100	100

The use of quantitative research approaches remains predominant within construction management research and this reinforces the idea that the majority of research is still using a rationalist or scientific approach. Any new researcher will be guided by the culture of their discipline. Their supervisors, colleagues and peers will be instrumental in the choice of research approach and methods made by someone embarking on a research career. In construction management there is a strong culture of quantitative research. This is often attributed to the origins of construction management research lying in the engineering discipline (Edum-Fotwe *et al.* 1996, Seymour and Rooke 1995).

Quantitative Data Collection

Quantitative data collection methods include gathering data using measurement techniques or equipment, systematic observation and the questionnaire survey. The use of the survey is evident in much research. In a review of recent CME publications it is clear that surveys are still a common data collection tool. 16 out of 29 papers used a survey to collect data for the research. Half of these used primary data collection. The survey is often a tool to collect quantitative data, although not exclusively so (see discussion later on the use of the survey in qualitative research). The choice of a questionnaire for data collection is guided by several factors. Most importantly will be the epistemological position that the researcher holds. Empirical research requires the linking of data to concepts. A questionnaire can be used to prestructure data very effectively. It is used to collect data that accurately describes a situation. Precise answers can be sought and easily comparable data is achievable. Most empirical research depends on comparison to establish conclusions (Sapsford and Jupp 1996). The time and finance allocated to research is often very limited, especially in “non-funded” research. With limited resources to conduct the research the option that attracts least cost and minimum effort in terms of time will be chosen in most cases. A questionnaire may be conducted face-to-face, over the telephone or self-administered. The self-administered questionnaire will be the cheapest and quickest method of

obtaining data from a large survey population and is therefore the most common choice to researchers under time and financial constraints.
The average response rate in recent CME publications was between 30% and 40%. According to Root and Blismas (2003) surveys within the construction management discipline should be achieving 85% return rates. This of course is a generalisation based on earlier research by Dillman (1978) but the achieved return rates are certainly nowhere near the perceived return rate of a well conducted survey. This does not bestow confidence on the collected data and by association the results of the research. Non-response is a major cause of error in surveys (Sapsford and Jupp 1996). One may question the validity of the research if the research process is in some way introducing a bias to the data. This raises several issues to do with the way in which data is collected.
It was stated that the most important aspect in the choice of questionnaire as a research tool is the epistemological position of the researcher. This may be true but all too often the questionnaire will be chosen not because of its appropriateness for generating research data, but because of its relative speed and inexpensiveness for data collection. Regardless of the choice of survey tool, whether web-based, paper based, a telephone survey or an interview, a fundamental requirement for data collection is to provide evidence. The conclusions must follow validly from the evidence (Sapsford and Jupp 1996).

Qualitative Data Collection

With the increase in qualitative research in construction management there has been a general broadening in the way in which data has been collected. Qualitative data is often generated in interview, focus group, participant observation and existing data. Qualitative data is non-numerical and usually takes the form of people's words or the researcher's description of what has been observed or experienced. Qualitative data can range from prestructured to not prestructured. There is a diversity in qualitative research that reflects the complexity and range of data that may be generated. Punch (1996: pp148) states that *what can count as useful data and ways of collecting them is very wide.* It should be stated that qualitative data is not synonymous with qualitative research.. The questionnaire whilst the mainstay of quantitative research can also be used to generate qualitative data. Open-ended questions and attitude scaling are typical methods for generating qualitative data from a questionnaire survey. Because of the nature of qualitative data there is less control over the format that it takes. The quality of the data does not rely on a measurement tool for its collection and therefore is more difficult to assess for its usefulness to a piece of research. Qualitative data may be collected in response to predetermined parameters or generated as part of an unfolding piece of research. The researcher is responsible for making sense of the data.

THE USE OF THE WEB IN COLLECTING DATA

The web has provided a new medium for research. The tools available allow research data to be collected more quickly, for less cost and it is suggested that a better rate of return can be achieved. Web-based surveys have increased in use since 1995 and the supporting technology is improving with development. The tools are only recently being used in construction management research. There are few examples of its use. Yet it has been successfully employed in consumer research and commercial applications for a number of years. Increasingly academic research is relying on the internet for gathering research data. In a survey conducted in 2002 (Roztocki and Morgan) of academic attitudes to web surveys, the majority believed that web surveys

are more efficient and the overall quality of response was *at least as high as traditional paper surveys.*

Many of the problems associated with web surveys are common to all types of survey. One area that is particularly relevant to web-based surveys is sample selection. It is relatively simple to create a web-based survey and place it on the web so that anyone can access it. The problem is moderating who accesses and responds to it. A probability based method is used to focus on a particular sample group. There are several ways of achieving this. Intercept surveys target visitors to a website and use systematic sampling to invite every *n*th visitor to participate in a survey. Low response rates to these surveys raise concerns over non-response bias, and also about self-selection (Couper 2000). List-based samples and Pre-recruited panels use email invitations to participate and it is suggested that this form of sampling will result in a higher response rate (Cook *et al.* 2000).

All surveys are subject to sources of error. Web-based surveys are no different. Typical errors are found in coverage and sampling, non-response and measurement. Coverage and sampling error are a result of the difficulty in identifying a sample that is representative of the population and that has equitable access to the technology to enable them to participate in the survey. Non-response can occur for a variety of reasons. Participants can drop out of the questionnaire process at any point. Poor graphics and the “flow” of the questionnaire itself are two main sources of non-response. Measurement error is a function of poorly worded questions (Manfreda *et al.* 2002) and is associated with all types of survey.

One of the main advantages of web-based surveys over paper surveys is the data is in electronic format and is collated as the returns are received. This contributes to having data available at a much earlier date than is possible with a paper survey. This function of web-based surveys is useful for large scale studies and assists in reducing errors with in-putting of data.

Using the web to generate qualitative data

A questionnaire that includes open-ended questions and opportunities for capturing qualitative data can be conducted via the web as readily as a quantitative questionnaire. There is little distinction in the technological requirements. It is also possible to use the web to conduct interviews, collect documentary evidence from a wide range of sources and even observe participants. The main issue surrounding the generation of data using the web is that of quality. Quality of data can be improved through rational and logical research design, but the web and its relative immaturity for collecting data has its own problems.

The web has a culture that allows for stop-start interaction. There is a tendency for web users to abandon one activity in favour of another.There is some research that shows that people act differently when using the web and are more likely to reveal personal details. This may be due to the anonymity associated with web use.

"Some modes of research are more inclined to richness than others. It is perfectly feasible to ask open-ended questions on-line and you often get a fuller, more articulate answer. On the other hand, we have found that the first big dip in respondent numbers often comes with the first open-ended question. They can tick boxes quickly, but they can't be bothered to type something. Lists and ranking questions are difficult to do on the telephone but are easier on-line." (Perrott 2001)

Delphi Groups can be carried out using the internet. A combination of web and e-mail works together to deliver detailed and extensive qualitative insights. It is argued that this is likely to be a more effective way of generating qualitative data than is possible with conventional conference room focus groups.

Sampling is a recurrent issue with web-based research regardless of the theoretical foundation. There are problems with traceability and consistency. Qualitative research tends to avoid the probability sampling of quantitative studies and uses *purposive sampling* (Punch 1996). This means deliberate selection of participants for a study. This aspect of qualitative research makes the web a more suitable medium in terms of data traceability compared with random sampling common to many quantitative studies.

Ethical Issues

Ethical consideration must be given to any piece of research. There are particular areas of concern for web-based research. The Association of Internet Researchers (AOIR) (2002) has produced a guide for ethical decision-making and internet research. This goes through various key issues. The most prominent issue is one of privacy. Web research can be conducted with relative ease across the entire globe and hence is carried out in countries with differing ethical and privacy standards. Participants in European Union must provide *unambiguous consent* (AIOR 2002:pp7) while in the USA there are far less stringent requirements. The researcher must be aware of these issues within the survey population.

EXAMPLE OF PAPER BASED VERSUS WEB-BASED SURVEY

A study of Registered Social Landlords (RSLs) was carried out to establish the understanding of sustainability and partnering within the social housing sector. The research tool was the questionnaire. This was chosen for the following reasons.

1. The survey population is very large. There are over 2300 housing associations within the United Kingdom. One of the main aims of the survey was to establish trends across the whole social housing sector.
2. The survey was designed to establish if factors coming out of an earlier qualitative study could be seen to represent the whole population of housing associations.
3. The use of a questionnaire focused the participant's thoughts on particular issues emerging out the vast area of sustainability and partnering.

Development of the questionnaire

The survey was conducted amongst the UK population of RSLs. 990 housing associations involved in development of social housing were selected for inclusion in the survey. It was decided to conduct a small experiment in research methods. One third of the participants were invited to complete a traditional paper survey the remaining two thirds were asked to complete the same survey which had been converted to a web-based survey using Dreamweaver (copyright© 1997-2002 Macromedia). All participants were selected from a national database of social housing providers. The participants selected for the web-based survey had provided email addresses and contact was made through this address. All participants were invited to take part under the title of 'Development Manager'. Due to resource constraints, it was not possible to address the survey to individuals with the organisations. It is believed that this might have contributed to a lower response rate.

Comparison of cost, time and response rates

The paper survey was developed following a piece of qualitative research in the Scottish social housing sector. The questions were developed using the emerging results from this study. The paper based questionnaire was piloted and refined prior to its broad circulation. This resulted in six weeks of development. The web-survey took three weeks to develop into an HTML application. So in total the web survey would have taken nine weeks from start to finish. The survey in its two forms was conducted

using a random samples from the same population. The response rates in the paper and web-based surveys were 35% and 42% accordingly. These rates represent an average return rate in construction management research. The comparison of costs for each type of survey are slightly misleading. Although in this case the costs for a web-survey were nil, there was significant use of resources to develop the web survey. To establish a relative comparison figures were sought from an independent Web-survey company charges in the region of 68p per respondent for (SySurvey 2004). Although not conclusive the comparison demonstrates that cost and time savings can be achieved and increased response rates are possible.

Table 2: Comparison of web-based and paper-based surveys

	Paper-based survey	**Web-based survey**
Time (development)	6 weeks	9 weeks
Time (response)	8 weeks	4 weeks
Response rates	35%	53%
Cost (per participant)	84 pence	Nil (68 pence)

DISCUSSION

We as researchers must strive to conduct research using the best quality data available or our research is fundamentally flawed. There is a large amount of empirical research conducted in construction management. The concepts and conclusions of this research depend on the existence of empirical data. The quality of this data has a direct bearing on the quality of the research and the conclusions that may be drawn.

There are a wide range of data collection methods and approaches available to the researcher. The survey or questionnaire remains predominant in construction management research although there is a significant increase in qualitative research methods and process driven quantitative data measurement techniques, e.g. cost measurement and productivity studies. In designing a piece of research it is vital that the data collection is given adequate thought to enable the collection or generation of data that provides gravitas to the research. The sophistication of models developed in the construction management sector is not always being accompanied by adequate data collection methods to enable true testing of these models. It may be that research projects do not reach fruition because of inadequate data, limiting the usefulness of the research. The research sector has a significant role to play in improvements in construction management as an industry but its effectiveness must be reduced if good quality data is not being collected to support research projects.

Qualitative research is emerging and the data collection methods associated with qualitative research are becoming more common and also more accepted in construction management research. With the increased use of focus groups, unstructured interviews, work diaries, etc. the techniques are becoming more sophisticated at generating data. Trends show an increase in these methods and it might be that a further increase in qualitative data will lead to further development of these techniques.

The web has provided a new way for data collection. There is a wealth of opportunity in generating and collecting data using the web and associated tools. Both quantitative and qualitative data may be collected on the web and many traditional collection techniques can be adapted to work on the web. Two of the main advantages of the web are the vast range of data readily available and the access to survey populations for collecting data. Web-based data collection has its problems. There is little control over what is placed on the web and any data that is found there is not automatically

reliable. Also there is less control over a survey population and potentially bias in the data, caused by technical proficiency and accessibility to the internet.
Data collection or generation has many associated problems. At each stage of data collection there is the potential to introduce bias to the research. The quality of data is determined by the careful consideration of issues affecting data collection. Each piece of research is unique and therefore data collection must be considered for the specific criteria of the research. There is no one right method for data collection and a research project may reach a different conclusion if an alternative data collection method is used. It must be said that the most relevant data collection techniques are used. Whether the research is quantitative or qualitative the method and approach to data collection remains important. There are many issues to be considered in the design of research and data collection must be given a central position in the research process to enable high quality and relevant research to emerge.

REFERENCES

AIOR (2002) *Ethical Decision-Making and Internet Research* http://www.aoir.org/reports/ethics.pdf

Bryman, A (1992) *Research Methods and Organisational Research* Routledge, London

Cook, C., Heath, F. and Thompson, R.L. (2000) *A meta-analysis of response rates in web- or internet-based surveys.* Educational and Psychological Measurement 60: 821-836.

Couper, M. (2000) *Web Surveys: A Review of Issues and Approaches*. Public Opinion Quarterly 64: 464-494.

Dillman, D (1978) Mail and Telephone Surveys: The Total Design Method. New York, Wiley

Edum-Fotwe, F.T.; Price, A.D.F. and Thorpe, A., (1996) *Research Method versus methodology: Achieving quality in scholarly research for construction management*, Proceedings of the 12th Annual Conference of ARCOM, pp 428-437

Ettema, D, Timmermans, H, Van Veghel, L (1996) *Effects Of Data Collection Methods in Travel And Activity Research* Eindhoven: EIRASS http://www.bwk.tue.nl/urb/eirass/report.htm

Locke, J (1649) *An essay concerning human understanding.* 22nd ed. London, 1812

Loosemore, M; Hall, K and Dainty, A., (1996) *Innovation and courage in construction management research*, Proceedings of the 12th Annual Conference of ARCOM, pp 418-427

Manfreda, K; Batagelj, Z and Vehovar, V (2002) *Design of Web Survey Questionnaires: Three Basic Experiments* Journal of Computer-Mediated Communication 7 (3) http://www.ascusc.org/jcmc/vol7/issue3/vehovar.html accessed 20/6/2004

Mason, J (1996) *Qualitative Researching*. London: Sage Publications Ltd

McCaffer, R and Edum-Fotwe, F, (1999) *Moving on from the crossroads: an agenda for research development in construction management*, Proceedings of the 15th Annual Conference of ARCOM, Liverpool John Moores University, pp 3-11

Perrott, N (2001) On-line research: Talking to Nicky Perrott http://www.mori.com/pubinfo/nep/online-research.shtml accessed 20/6/2004

Punch, K (1998) *Introduction to Social Research*. London: Sage Publications Ltd

Root, D and Blismas, N (2003) *Increasing Questionnaire Responses from Industry*, Proceedings of the 19th Annual Conference of ARCOM, Brighton University, pp 623-631

Roztocki and Morgan (2002) *The Use of Web Based Surveys in Academic Research in the Field of Engineering* American Society of Engineers National Conference 2002

Sapsford, R. and Jupp, V. (1996) *Data Collection and Analysis*. London: Sage Publications Ltd

Seymour, D. and Rooke, J. (1995) *The culture of the industry and the culture of research*, Construction Management and Economics 13, pp 511-523

SySurvey (2004) Online Surveys: Pricing http://www.sysurvey.com/pricing/pricing.asp accessed 20/6/2004

Yun, G and Trumbo, C (2000) *Comparative Response to a Survey Executed by Post, E-mail, & Web Form* Journal of Computer-Mediated Communication 7 (3) http://www.ascusc.org/jcmc/vol6/issue1/yun.html accessed 20/6/2004