

Small sample market research

Timothy Bock* and John Sergeant†

**University of Sydney* and †*Eureka Strategic Research*

Numerous market research studies are conducted with small samples of data. Many of these data are described as 'qualitative'. Researchers do not always appropriately take into account the size of the sample when drawing inferences about populations from small samples of data. In this paper we present some rules of thumb that may be employed when attempting to draw inferences from small samples, such as those typically used in qualitative research.

Introduction

'Our researchers speak both languages, qual and quant', states the owner of a research company, hoping to sustain its position in an increasingly specialised research industry. The past few decades have seen some researchers and research buyers view qualitative and quantitative research as viable alternative research methods and philosophies. Such a conclusion has merit only through its avoidance of conflict. The relativism that pervades the modern research industry is fraught with danger – buyers of research are seeking the truth, and the apparent absence of any agreed-upon criteria for establishing what is true should be a major ethical dilemma for the research industry.

The ethical dilemma is most obvious in political research, if only because it is in the public domain. The past decade has seen an unprecedented increase in the influence of market research, with polling increasingly driving the policies of Western governments and political parties. The irony is that, in their claimed desire to increase the democratisation of policy development, some politicians use qualitative research techniques that are unable to communicate the voice of the electorate accurately. According to Andrew McIntosh (Smith & Fletcher 2001, p. viii), 'those advising political parties in Britain have indeed misused focus groups, and neglected the proper use of survey research'. A similar conclusion has been reached in Australia (Collins 2000). What makes 'qualitative polling' so inappropriate? We would contend that the fundamental problem is not

one of qual versus quant, but one of sample size. The per-respondent costs of qualitative research are sufficiently high that in most situations where ‘quantitative’ results are required (such as predicting the consequences of policies, personalities and peccadilloes on preferences for political parties), it is not economically viable to conduct qualitative research with a robust sample size.

In this paper, we present the view that researchers and users of research should keep in mind the fact that qualitative research findings are typically drawn from small samples. We first discuss whether there is anything intrinsic to qualitative research that obviates concerns about sampling methods and sample size. We then present a list of the types of inferences that can be appropriately drawn from small samples. Finally, we present conclusions.

Qualitative research and small sample research

The vast majority of research involving small samples is commonly referred to as qualitative research. Indeed, in the absence of any agreed-upon definition of qualitative research, perhaps the single most obvious way of defining qualitative research is that it typically involves small samples. Nevertheless, we suspect that most researchers and users of research would find the depiction of qualitative research as being research involving small samples as rather missing the point.

If qualitative research is not simply research with small samples, then what is it? Perhaps the most widely held misconception is that qualitative research is research without numbers. This is essentially a negative definition, stating that qualitative research is research that is not quantitative. However, from the determination of the appropriate sample size (Griffin & Hauser 1993), through to analysis and reporting (Kelle 1995), some qualitative researchers have exhibited a decidedly numeric focus.

Gordon & Langmaid (1988, p. 2) typify what seems to be the view of many qualitative practitioners when they state that qualitative research ‘is centrally concerned with understanding things rather than with measuring them’. This carries with it an implication that quantitative research may well be concerned more with measurement than with understanding. A cursory review of the seminal quantitative marketing papers of the past half century reveals a preoccupation with gaining understanding (e.g. Guadagni & Little 1983; Goodhardt *et al.* 1984; Erdem & Keane 1996). Quantitatively inclined researchers simply see measurement as a necessary

step in gaining understanding. Indeed, it would seem impossible to draw any general inferences as to why consumers do things without *measuring* their behaviour, perceptions and the characteristics of the environment around them. While qualitative researchers may not employ the formal scales of quantitative research, their methods encompass measurement techniques. For example, the popular card sort of qualitative research measures the perceived similarity of objects.

Most practitioners and users of market research seem to view qualitative research as being a collection of techniques that are appropriate for solving certain types of problems. Different qualitative methodologies are seen as being appropriate for different types of problems (see Smith & Fletcher 2001). Similarly, among academic researchers, ‘consensus has gradually emerged that the important challenge is to appropriately match methods to evaluation questions and issues, and not to universally advocate any single methodological approach for all evaluation situations’ (Patton 1990, p. 492).

Tied to the advocacy of multiple research techniques is a more general view about the dangers of dogmatic adherence to any one research tradition. This ethos is perhaps best appreciated through the following parable of the fruits of qualitative research, which Patton (1990, p. 9) attributes to *Halcolm’s Evaluation Parables*:

There was once a man who lived in a country that had no fruit trees. This man was a scholar and spent a great deal of time reading. In his readings, he often came across references to fruit. The descriptions of fruit were so enticing that he decided to undertake a journey to experience fruit for himself.

He went to the market and asked everyone he met if they knew where he could find fruit. After much searching he located a man who knew the directions to the country and place where he could find fruit. The man drew out elaborate directions for the scholar to follow.

With his map in hand, the scholar carefully followed all of the directions. He was very careful to make all of the right turns and to check out all of the landmarks that he was supposed to observe. Finally, he came to the end of the directions and found himself at the entrance to a large apple orchard. It was springtime and the apple trees were in blossom.

The scholar entered the orchard and proceeded immediately to take one of the blossoms and taste it. He liked neither the texture of the flower nor the taste. He went to another tree and sampled another blossom, and then another blossom, and another. Each blossom, though quite beautiful, was distasteful to him. He left the orchard and returned to his home country, reporting to his fellow villagers that fruit was a much-overrated food.

Being unable to recognise the difference between the spring blossom and the summer fruit, the scholar never realised that he had not experienced what he was looking for.

This parable warns that measurement without understanding is, well, fruitless. It provides an argument for research techniques that help to foster understanding prior to the use of other, more measurement-oriented techniques.

With mixing and matching methods being a popular approach to using qualitative research (Patton 1990; Creswell 1994), a crucial question is: 'How does one know when to use which method?' By and large, researchers seem to use tacit knowledge when making such choices. Generally, the choice seems to come down to one of measurement. Quantitative methods are rejected when researchers believe that rating scales and the like are too blunt an instrument, or when there is, as yet, insufficient understanding of that which is to be measured.

While measurement is at the heart of most quantitative research, it should not be seen as synonymous with quantitative research. Statistical inference, and its subdiscipline of sampling theory, is also an essential part of quantitative research. To accept that qualitative research may provide a better understanding of the mind of the consumer is to make a statement about measurement; it in no way obviates the need for consideration of the broader issues of statistical inference. Statistical inference is concerned with the problem of 'drawing conclusions about a population on the basis of measurements or observations made on a sample of individuals from the population' (Everitt 1998, p. 163). In our experience, almost all qualitative research conducted in marketing attempts to draw an inference from a sample about a population. Clients do not watch focus groups to appreciate the performance of the moderator and the participants; they watch because they believe that the participants are in some way representative of the consumers in their markets. Therefore, the methods and concepts of statistical inference are of no less importance for qualitative research than for quantitative research. While statistical analysis is not without its own controversies (see Ehrenberg 1970; Lindsey 1999), to our knowledge no critic of quantitative research has ever attempted to prove that statistical inference is invalid, or to present an alternative.

Why are we making what to many would seem an obvious point? While there has been extensive debate about the differences between qualitative and quantitative research, this debate has been concerned largely with the difference between qualitative studies using small sample sizes and

quantitative studies employing large sample sizes. This literature is rich in its insights and implications for commercial practice (see Blyth & Robson 1981; Sykes 1990). However, it suffers from the drawback that it starts from the premise that the issue of sample size is not separate from the issue of qual versus quant. To our minds this is problematic, as the relative sample sizes of qualitative and quantitative research are at least in part driven by budget. By contrast, in this paper we focus solely on the issue of the types of finding that can be reached with small samples, irrespective of whether a qualitative or quantitative methodology is being employed.

This point is best appreciated with an example. The following conclusions are from a 'qualitative' study authored by Senior Researchers of one of Australia's largest suppliers of market research for one of the world's largest buyers of research:

For the majority of the younger generation (up to the young families), [brand name] plays only a minor role, and none in some cases, in the decision making process.... The [brand name] brand does, however, have some equity with the older generation.... All evidence points to ... a devaluation of the [brand name] brand over time.

While this study employed a sample of 30 consumers, the researchers have felt confident in generalising from subsamples to populations ('the majority of the younger generation'). They have also performed some form of (probably implicit) statistical test and concluded that age groups differ in terms of the role of the brand.

Is the conclusion right? It may be. However, the wording of the conclusions, and the absence of any caveats or limitations in the report, implies that it is a fact about the client's market. Statistical inference was invented to address questions of this type. Given the small sample size (30), and the empirical knowledge that there are rarely large demographic skews in the purchasing of established brands (Hammond *et al.* 1996; Ehrenberg & Kennedy 2001), it is likely that the conclusions were not warranted. They were certainly not warranted based upon the data at hand. This is not to say that the conclusions are incorrect. Conclusions can be reached without any research whatsoever and still be correct.

At a more fundamental level, it is interesting to note that the research conclusion quoted earlier also tacitly asserts that there exists a quantity corresponding to the strength of the role played by brand in purchase decision-making and that this quantity is measurable. More generally, it may be seen that any statement admitting of degree or strength is inherently quantitative. Thus the appeal to quantitative concepts in

ostensibly qualitative research is not limited to statements about differences between groups, which rely on the techniques of statistical inference.

Appropriate small sample findings

The impression to draw from the above discussion is not that much of what is called ‘qualitative research’ is invalid. Rather, it is that, when evaluating conclusions drawn from small samples, such as those commonly employed with qualitative research, we should be mindful of the difficulties inherent in any attempt to make generalisations about populations from small samples. The time and monetary constraints placed upon market research practices mean that small sample market research will remain widespread. However, as the informal data collation and analysis methods that are in widespread use prevent the application of formal significance tests to qualitative research results – although we note that Miles and Huberman (1984) advocate counting the frequency of different responses – we should be mindful of the types of conclusion that may be reliably drawn from small samples.

It is accepted that certain types of quantitative finding always require large sample sizes. For example, market segmentation studies generally require large samples (in the hundreds or thousands). Similarly, screening new product ideas often requires large sample sizes, to separate what may be small differences in preference that may equate to large differences in profitability. Is qualitative research less relevant for such research? Of course not. It may still be used to identify extremely unpopular products, and for hypothesis generation and questionnaire design.

When small sample sizes are being employed – say, when each subpopulation of interest has fewer than 30 respondents – we should be very careful to ensure that any inferences are appropriate given the data collection (of course, when a small sample represents a high proportion of our population, such concerns are less relevant). There are at least four types of valid findings that can be drawn from small *representative* samples:

- 1 *Lists of ideas.* A purist perspective on qualitative research sees it as the process of generating ideas and hypotheses. Lists of value to clients include lists about the types of benefits people desire from products, new product ideas and interpretations of the messages of ads. When

additional sample ceases to generate new ideas, such research is complete (Calder 1977; Lincoln & Guba 1985).

- 2 *'All' (or 'none') conclusions.* If (almost) everybody who participates in a small study says the same thing – for example, 'I don't understand this ad', 'the waiters are rude', 'I prefer the new packaging' – the conclusion is likely to be valid, based on established principles of statistical inference.
- 3 *'Some' conclusions.* Often the negative case of an all-or-none conclusion – the 'some' conclusion – is of most interest. The observation of only one contrary instance is sufficient to dispel a general conclusion. For example, while we suspect that the problem of researchers neglecting the effect of sample size when interpreting qualitative research is widespread, the only evidence we have presented in this paper is three examples. However, these three examples mean that nobody can argue that 'all' creators, buyers and users of qualitative research appreciate the constraints provided by small samples (i.e. 'some' clearly do not).
- 4 *Supporting the status quo.* In some situations there is substantial *a priori* evidence to suggest that a certain finding will be observed (e.g. Fornell 1995; Kalyanaram & Robinson 1995; Kalyanaram & Winer 1995; Uncles *et al.* 1995; Ehrenberg & Kennedy 2001). When qualitative research finds a pattern consistent with known empirical generalisations about markets – for example, if we found that few buyers of laundry detergent are devoutly loyal to a particular brand – the conclusion may be justified. This is also the case when the qualitative research confirms existing beliefs of managers (Smith & Fletcher 2001), although in such circumstances it is important to ensure that these beliefs were formed prior to the research, due to our ability to post-rationalise virtually any research findings, and are based on something more than speculation.

A paper about sampling would not be complete without a brief discussion about representativeness. It is possible that a convenience sample of the type commonly used in qualitative research will be representative. Unfortunately, it is equally possible that such a convenience sample will be highly unrepresentative – there is simply no way of having any justifiable confidence in the representativeness of a non-probability sample.

It has been argued that ‘purposive sampling’ is to be preferred to probability sampling when conducting qualitative research. Lincoln and Guba (1985, p. 202) state that purposive sampling

is based on informational, not statistical, considerations. Its purpose is to maximise information, not facilitate generalisation ... the criterion invoked to determine when to stop sampling is informational redundancy, not a statistical confidence level.

From a statistical perspective this statement is contradictory, because a confidence level is a measure of informational redundancy. A ‘tight’ confidence interval means that a researcher can be fairly confident about generalising from information that has been drawn from a sample, making the collection of further information unnecessary as it will probably be redundant. Two objectives are at the heart of purposive sampling: ensuring that all relevant types of people are included in the sample, and changing the sample structure during the research process. Both notions are consistent with sampling theory (see Sudman 1976; Singh & Chaudhary 1986). Ensuring that different types of people are covered in the sample is the goal of stratified random sampling. Indeed, quantitative researchers who always use the occasionally irrelevant age, gender and geography for stratification can learn much about stratification from qualitative researchers. Modifying the sample structure to reflect what has and has not been learned through the data collection is also highly desirable from a sampling theory perspective. It is referred to in sampling theory as sequential sampling; unfortunately, pragmatic considerations usually prevent this from occurring in both commercial qualitative and quantitative market research.

Conclusion

The human mind is trained to find patterns. As researchers, we like to find patterns. Clients like patterns. The great dilemma of marketing research is that patterns in data have two distinct causes. The first is that the patterns reflect real world phenomena of interest to our clients. The second cause of patterns is the research process. The smaller the sample size, the more likely that any pattern observed in data will reflect random variation in the characteristics of samples drawn from populations, regardless of whether the research is characterised as qualitative or quantitative. We should be mindful of this when presenting findings.

Researchers have long been mindful of the role of sample size when presenting the results of quantitative research. Even when quantitative studies have sample sizes in the thousands it is commonplace to indicate where there is limited ‘confidence’ in the results, such as through the provision of *p*-values and confidence intervals. When the sample size for a subgroup of interest is small – typically less than 30, 50 or 100, depending upon the researcher – it is common practice either not to present the data or to present it with appropriate caveats. We have identified four types of conclusion that may be drawn from small samples without the need to place caveats on the findings. In instances where results from small samples are presented, but are not consistent with the principles presented here, researchers have an ethical responsibility to inform clients of the limitations of the findings. A blanket caveat, such as ‘of course, these results are qualitative’, is not sufficient, as there are numerous findings that can be confidently made using small samples. To discuss whether all of the findings of qualitative research are or are not justified due to sampling considerations misses the point; as with quantitative research, in any qualitative study some of the conclusions may be drawn with greater confidence than others.

References

- Blyth, B. & Robson, S. (1981) Resolving the hard/soft dilemma. MRS Conference.
- Calder, B.J. (1977) Focus groups and the nature of qualitative marketing research. *Journal of Marketing Research*, 14, August, pp. 353–364.
- Collins, P. (2000) *The Bear Pit*. London: Allen & Unwin.
- Creswell, J.W. (1994) *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.
- Ehrenberg, A.S.C. (1970) Models of fact: examples from marketing. *Management Science*, 16, March, pp. 435–445.
- Ehrenberg, A.S.C. & Kennedy, R. (2001) There is no brand segmentation. *Marketing Research*, 13, 1, pp. 4–7.
- Erdem, T. & Keane, M.P. (1996) Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15, 1, pp. 1–20.
- Everitt, B.S. (1998) *The Cambridge Dictionary of Statistics*. Cambridge: Cambridge University Press.
- Fornell, C. (1995) The quality of economic output: empirical generalizations about its distribution and relationship to market share. *Marketing Science*, 14, 3/2, pp. G203–G211.
- Goodhardt, G.J., Ehrenberg, A.S.C. & Chatfield, C. (1984) The Dirichlet: a comprehensive model of buying behaviour. *Journal of the Royal Statistical Society – Series A*, 147, 5, pp. 621–655.

- Gordon, W. & Langmaid, R. (1988) *Qualitative Market Research: A Practitioner's and Buyer's Guide*. Aldershot: Gower.
- Griffin, A. & Hauser, J.R. (1993) The voice of the customer. *Marketing Science*, 12, 1, pp. 1–27.
- Guadagni, P.M. & Little, J.D.C. (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2, 3, pp. 203–238.
- Hammond, K.A., Ehrenberg, A.S.C. & Goodhardt, G.J. (1996) Market segmentation for competitive brands. *European Journal of Marketing*, 30, 12, pp. 39–49.
- Kalyanaram, G. & Robinson, W.T. (1995) Order of market entry: established empirical generalizations, emerging empirical generalizations, and future research. *Marketing Science*, 14, 3/2, pp. G212–G221.
- Kalyanaram, G. & Winer, R.S. (1995) Empirical generalizations from reference price research. *Marketing Science*, 14, 3/2, pp. G161–G169.
- Kelle, U. (1995) *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice*. London: Sage.
- Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*. Beverly Hills, CA: Sage.
- Lindsey, J.K. (1999) Some statistical heresies. *The Statistician*, 48, pp. 1–40.
- Miles, M.B. & Huberman, A.M. (1984) *Qualitative Data Analysis: A Sourcebook of New Methods*. London: Sage.
- Patton, M.Q. (1990) *Qualitative Evaluation and Research Methods* (2nd edn). Newbury Park, CA: Sage.
- Singh, D. & Chaudhary, F.S. (1986) *Theory and Analysis of Sample Survey Designs*. New York: John Wiley & Sons.
- Smith, D.V.L. & Fletcher, J.H. (2001) *Inside Information: Making Sense of Marketing Data*. Chichester: John Wiley & Sons.
- Sudman, S. (1976) *Applied Sampling*. New York: Academic Press.
- Sykes, W. (1990) Validity and reliability in qualitative market research: a review of the literature. *Journal of the Market Research Society*, 32, 3, pp. 289–328.
- Uncles, M.D., Ehrenberg, A.S.C. & Hammond, K. (1995) Patterns of buyer behavior: regularities, models, and extensions. *Marketing Science*, 13, 3/2, pp. G71–G78.