So you want to do research? 2: developing the research question

Keith A Meadows

In the first paper of this series (Meadows, 2003), the main phases of the research project were mapped out. Heading this process was identifying the research question. Identifying and formalizing the research question are where many problems with research occur. Without a well-defined and specific research question or hypothesis, even if the research is carried out to the end, it is unlikely to provide much information. A tightly-focused research question or hypothesis dictates what data are collected, how they are collected and analysed and provides a context for the results.

The aim of this paper is to discuss the process of developing a research question from the initial idea through to the final research question, which is defined in operational terms, is unambiguous, measurable and worthy of the research effort.

Developing the research question

Most research originates from a problem or question identified from the literature or arising out of the experiences of the researcher relating to some general problem. For example, it might have been observed that non-attendance rates at an outpatient clinic appear to be very high for a particular group of patients and the purpose of the research would be to explore the reasons why. Alternatively, the researcher may want to consider whether the implementation of a programme, e.g. to facilitate a patient-centred approach to the management of a specific chronic illness, has beneficial outcomes in terms of patient's wellbeing compared to no programme, or which rehabilitation programme has better outcomes for patients than another following a myocardial infarction (MI).

Reviewing the literature

The first essential stage in developing the research question should be reviewing the literature. By reviewing the literature at an early stage, existing research associated with the problem can be identified and will assist in drafting the raw research question. Having decided on the raw, or rough, question to be addressed by the research, two questions will need to be answered to refine the research question(s). First, what kind of information needs to be collected to answer the question? Second, how is this information collected? In other words, the design of the study or research needs to be decided and the right methodology to fit the type of question needs to be selected. A comprehensive review of the literature helps to answer these questions by identifying earlier research and the methodologies used.

An effective way of identifying relevant literature is through the use of electronic sources (*Table 1*). Relevant literature can be searched over specified years, using keywords, by study area, and by author. Medline and EMBASE provide coverage of the literature in many health-care areas but do not record all publications from all medical journals. While Medline has a wide coverage of English language journals, EMBASE provides an increased coverage of articles in other European languages.

Keith A Meadows is Head of research and development, City and Hackney Primary Care Trust, London and a Visiting Lecturer, St Bartholomew School of Nursing and Midwifery, City University, London

Email: keith.meadows@chpct.nhs.uk

ABSTRACT

The development of the research question for a study can be where a lot of research fails. Without a well-defined and specific research question or hypothesis, findings from the research are unlikely to tell us very much. Developing a tightly-focused research question or hypothesis defines how and what data is collected and analysed and provides a context for the results. This article, the second in a series of six, focuses on the process of developing a research question or hypothesis from the initial idea through to the final research question, using examples to illustrate the key principles. Approaches to reviewing the literature, including hand searching and the use of electronic sources, are described together with their different strengths and weaknesses. An overview of the deductive and inductive approaches to research are described, as well as the underlying rationale of the null hypothesis and one and two-tailed tests. Finally, issues around the feasibility of the study, including, cost, time and relevance, are discussed in relationship to developing the research question or hypothesis.

The comprehensiveness, or recall, of a database search depends on the search strategy applied. Searches with high recall may have low precision, i.e. while they retrieve a large number of articles, many of these might be inappropriate (NHS Centre for Reviews and Dissemination, 1996). These problems can be minimized, to a degree, through the use of search strategies and knowledge of the relevant terms used for indexing. However, such skills are not easy to acquire and it is therefore highly recommended that an information scientist is consulted at an early stage of the review process.

Scanning the reference lists of retrieved articles, and the bibliographies of systematic and non-systematic review articles, can identify additional literature not recovered by the search. Whenever possible, key journals in the field should be searched by hand to identify articles which may have been missed or overlooked in the various database searches. Hand searching is also useful in identifying recent articles which have not yet been cited or indexed electronically (NHS Centre for Reviews and Dissemination, 1996).

Published literature can provide a rich source of research ideas through either extending or refining previous research. Other sources of ideas include:

- The Research Findings Register (http://tap.ukwebhost.eds.com/doh/refr_web.nsf/Home?OpenForm) provides a summary of research projects funded by the NHS, including information such as the research question, methods and findings. The research projects listed are fairly current, which is sometimes not the case with journal articles.
- The National Research Register is a database of about 80 000 ongoing and recently completed research projects funded by, or of interest to, the UK's health service (www.update-software.com/national/)
- The NHS Centre For Reviews and Dissemination (www.york.ac.uk/inst/crd/) holds abstracts of quality-assessed systematic reviews, economic evaluations of health-care interventions and publications and projects by a variety of health-care technology assessment agencies
- The Cochrane Library holds freely-accessible abstracts of Cochrane Reviews (www.update-software.com/cochrane/).

Conference proceedings can be a useful source of information of research in progress and completed. They are generally an unreliable source of data, so where possible, reports from the author(s) should be obtained before any reference is made to the study (NHS Centre for Reviews and Dissemination, 1996).

Results of studies which have been published in reports, conference proceedings, and discussion papers or other formats, and which have not been indexed on the main databases, is known as 'grey literature'. The identification of grey literature is difficult, although the libraries of specialist research centres, research funding organizations and societies may provide a useful source.

It is important to stay focused on the relevant issues and concepts when undertaking a literature review. Whenever possible, the researcher should concentrate on that literature which is credible and has undergone peer review, although identifying other types of literature can be helpful.

Table 1. Online sources of bibliographic information

Medline	www.ncbi.nlm.nih.gov/PubMed/
CINAHL	www.cinahl.com
ENB	enb-search.ulcc.ac.uk/cgi-bin/hcdsearch
Psycline	www.psycline.org
ERIC	www.eric.ed.gov
Metcrawler	www.metacrawler.com/info.metac/dog/index.htm
JISC	www.jisc.ac.uk
Cochrane Database	www.update-software.com/Cochrane/default.htm
British Medical Journal	www.bmj.com
BIDS	www.bids.ac.uk/
OMNI	omni.ac.uk
EMBase	www.embase.com
Athens	www.athens.ac.uk

Research design

Research methodology can be broadly categorized as either qualitative or quantitative. The aim of qualitative research is to help understand social phenomena in a natural, rather than an experimental, setting with emphasis on the meanings, experiences and views of the participants (Pope and Mays, 1995); i.e. to determine 'why', rather than 'how many'. Anastas and MacDonald (1994) refer to qualitative research as 'flexible designs', which are interpretative, using ethnographic or qualitative approaches and with less pre-specification as to what information is required. Often in flexible designs, the design evolves and develops as the research continues.

Quantitative research, which includes experimental research such as randomized control trials (RCTs) and surveys, are examples of what Anastas and MacDonald (1994) refer to as 'fixed research designs'. Quantitative research is generally under-

'The aim of qualitative research is to help understand social phenomena in a natural, rather than an experimental, setting with emphasis on the meanings, experiences and views of the participants, rather than providing quantified answers to a research question.'

taken using a model that requires the need to know exactly what to do before collecting and analysing the data. Quantitative research has a reliance on statistical analysis and generalizations and is often theory driven (Robson, 2002). Therefore, before starting, some idea is needed of what area is to be dealt with, what information is to be gathered, and how.

Types of research question

Why the question is being asked, and the type of information to be collected to answer it, determines the type of research design to be used. For example, differences in the experience of mental illness across different cultures or ethnic groups could be investigated using standardized survey techniques and quantitative analysis (i.e. fixed design) to describe differences in rates of illness such as incidence and prevalence. On the other hand, by taking a qualitative approach (flexible design) the researcher would gain insights into, and an understanding of, the factors and experiences from the analysis of the narrative, obtained from in-depth interviews, which underlie the identified differences (Nazroo and O'Connor, 2002). There are three broad types of questions which research can address:

1. Descriptive

The purpose of the study is to describe what is going on or what exists. For example, differences in rates of illness or the proportion of outpatients sampled with various opinions about the level of service provided. Both quantitative and qualitative research methods can be applied to answer descriptive questions.

2. Relational

This is when a study looks at the relationship between two or more variables, e.g. the proportion of men and women with negative attitudes on the level of service provision, or the proportion of women with eating disorders compared to men. In these examples, the relationship between gender and negative attitudes and eating disorders are respectively being examined. Only quantitative or fixed research design can answer these types of questions.

3. Causal

This is when a study is designed to examine whether one or more variables significantly causes or affects one or more outcome variables. For example, does an intervention programme improve patient self-reported wellbeing? Only a quantitative or fixed research design can answer these types of questions. For more information of the different types of research questions see Fink (1995), Robson (2002).

The research hypothesis

An alternative to the research question is the research hypothesis. In contrast to a research question, an hypothesis is a prediction, which is phrased in operational terms as to exactly what the researcher thinks will happen in the study and is generally linked to some underlying theory (Gilbert, 2001; Robson, 2002). The broad subject matter could, for example, be to test the hypothesized effects of some treatment on some measurable outcome.

Generally, a hypothesis is set in the context of some theory. The theory need not be some 'blue-sky' scientific theory and not all studies will necessarily have, or require, a hypothesis; the study might be designed to be exploratory, or to explore some area more thoroughly in order to develop some specific hypothesis or predictions that can be tested in future research. On the other hand, a research question may be asked which is unrelated to any particular theory, but addresses a particular problem.

There are two approaches to establishing facts: the inductive and deductive methods. With inductive reasoning, the focus is shifted from specific observations to making broader generalizations and theories. This could be done by making a number of specific observations, e.g. observing that quality of life improves in patients receiving psychotherapy, and from these make generalizations about the benefits of psychotherapy on quality of life. It is possible to test a number of hypotheses to test these generalizations and develop a theory. However, a major limitation of the inductive approach is that it is not possible to prove such generalizations beyond the specific situation under study. On the other hand, deductive reasoning involves working from the general to the specific. For example, a researcher may be developing a theory around the effectiveness of psychotherapy on the quality of life of patients. He/she would develop a number of specific hypotheses to test. As each of these hypotheses are tested and confirmed, this provides growing support for the theory. For a more detailed discussion on hypothesis testing, see Argyrous (2000).

Research question or hypothesis?

Choosing whether the study will ask a research question or test a hypothesis is, of course, dependent on the purpose of the study. If a prediction is to be tested, which is related to some underlying theory, a hypothesis for the study needs to be developed. However, if a description of what exists in the study population is needed, or an examination of relationships between variables and/or factors, or examination of the causes and effects of a new treatment, the relevant research question needs to be asked. How

this question is asked is dependent on whether the researcher is undertaking a quantitative or qualitative research study.

Operationalizing the research question

We know that the type of information obtained from a study is dependent on the question or questions asked, and the type of questions asked has an important bearing on whether the research methodology chosen is qualitative or quantitative.

Because of the nature of qualitative research, with its tendency to focus on the narrative so as to have a better understanding of social phenomena in natural rather than experimental conditions, the questions that are asked in qualitative research are different from those in quantitative research.

Often in qualitative research, the research questions evolve from a conceptual framework which is built by the researcher, either graphically or in narrative form, showing the key factors, variables and concepts to be studied and their presumed interrelationships. The framework can be theory-driven, simple, descriptive, and causal or common-sensical (Miles and Huberman, 1994). For a more detailed discussion on conceptual frameworks and research questions in qualitative research, see Miles and Huberman (1994).

In the early phases of setting up a quantitative, or fixed-design study, the research question will be a preliminary one, which will be too vague or broadly phrased to enable specific data to be collected and analysed to answer it. A comprehensive review of the literature will be essential to identify what previous research has been undertaken and which methods have been used to answer the question. This information will provide the basis for developing the research question so that what is being asked, and how it is being asked, is made explicit. This is known as 'operationalizing' the research question.

Operationalizing the research question involves first, identifying the concepts referred to in the research question and second, converting these concepts into operational definitions and expressing them as measurable indicators. In other words, having identified the 'concepts' in the research question, the researcher must determine how these concepts will be measured. Determining what data need to be collected and the best approach to collecting the information, will then follow on. A concept has been described as:

> **'a label we put on a phenomenon that enables us to link separate observations and to make generalizations (e.g. depression, quality of life, disability). A convenience, a name we give to observations and events.' (*Trochim, 1999*)**

A concept can range from the concrete to the abstract and the degree of that abstraction will have a significant impact on the ease and availability in the selection of the indicator(s) to measure the concept(s). The more abstract the concept, the greater the difficulty. For instance, the concept 'physical functioning' could be considered to be more concrete and understandable and, consequently, more easily defined than 'self-esteem'. Physical functioning could be operationally defined in terms of ability to climb one flight of stairs, walk more than a kilometre or 100 metres. Yet for self-esteem there is less clarity and perhaps agreement as to what it is and its constituent components. These stages can be summarized as a simple strategy for operationalizing a research question:

- Formulate the research question
- Identify each of the concepts or main ideas in the research question and list these as Concept 1, Concept 2, and so on
- Review the literature, search for and list identified operational terms for each of the concepts
- Identify and list measurable indicators for each concept
- Rewrite the research question in an operational form.

Example 1

The first raw question could be:

'*Which treatment is better for treating psychiatric disorders?*'

In its present format there is no way to test the question – which treatments and psychiatric disorders are of interest? However, following a review of the literature it may be possible to move on to a more structured question such as:

'Is psychotherapy more beneficial than psychiatric medication for people with a psychiatric disorder?'

This question is testable, but in its present form is too general. There are three key concepts: psychotherapy, psychiatric disorder and psychiatric medication (*Table 2*). To operationally define these, it needs to be clear as to what is meant by psychiatric disorder (i.e. acute or chronic; mild or severe), what type of psychotherapy, what type and dosage of medication will be given, over what period of time, and how would the outcome be measured?

Working from *Table 2,* a more specific question could be:

'Will there be a statistically significant improvement in mildly depressed women as assessed on the

'Often in qualitative research, the research questions evolve from a conceptual framework which is built by the researcher, either graphically or in narrative form, showing the key factors, variables and concepts to be studied and their presumed interrelationships'

Beck Depression Inventory (BDI) after 6 months of cognitive therapy compared to fluoxetine (Prozac)?'

This question operationally defines both the type and degree of illness, against an objective measure or measurable indicator – the BDI (Beck et al, 1961), as well as the time period and treatment. In addition, by stating a statistically significant improvement, it can be determined whether any difference found will not be by chance, rather than just random improvement.

Example 2

'Does rehabilitation improve quality of life of patients post-MI ?'

Again, the question is testable but needs to be more precise. First, a definition is needed of what kind of rehabilitation is provided. Also, there is no mention of a control group. There may be an improvement, but it may also be found that quality of life improves in patients not attending a rehabilitation programme, so a control group is needed to compare and measure the effectiveness of any changes. What does the rehabilitation programme comprise? How long after the completion of the rehabilitation programme is it of interest to find improvements in quality of life? What does quality of life mean?

Quality of life is an ill-defined concept and can cover a number of non-health as well as health-related dimensions including psychological wellbeing, pain, social activity, role and physical functioning. Many so-called quality of life measures are, in fact, measures of health status. As a consequence, the researcher needs to be specific in predicting which areas to expect changes to occur, e.g. social and role activity, and psychological wellbeing.

The population of interest also needs to be clarified – are we thinking about all patients who have suffered MI? Anxiety levels arising from participating in a rehabilitation programme might, for example, be higher for patients with a second MI compared to first MI, therefore possibly reducing any beneficial effect on psychological wellbeing. What if patients with diabetes were to be included? Will they be more concerned over increased risks of acute complications such as increased hypoglycaemic episodes? Age may also be a key factor of outcome. The length of the programme needs to be defined. Too short, and any potential improvement may be missed. But, what is meant by improvement? What levels of improvement are being looked for and compared, to justify the programme's effectiveness?

Research questions must be written to avoid ambiguity; they must be specific as to what is an answer, they must be answerable with the data available and the method used to collect the data and, finally, must be worthy of the research being undertaken. For a more detailed discussion on developing research ideas and research questions, see Robson (2002).

Operationalizing the research hypothesis

Types of hypothesis

As with the research question, the hypothesis must be operationalized so that what is being predicted, and how, is made explicit. In the earlier example of the effectiveness of cognitive therapy on mildly depressed patients, assume a theory is being developed around the effectiveness of cognitive therapy. The hypothesis could be, for example:

'That there will be a statistically significant improvement in mildly depressed women compared to men as assessed on the BDI after 6 months of cognitive therapy compared to fluoxetine.'

This is almost identical to the research question with the exception that the hypothesis is written as a prediction by the use of the word 'that'. In all other respects, the level of specificity is the same.

The formal way of setting up the hypothesis test is to formulate two hypothesis statements. The first describes the prediction, e.g. that there will be a statistically significant improvement in mildly depressed women compared to men as assessed on the BDI after 6 months of cognitive therapy compared to fluoxetine. The hypothesis making the prediction is known as the alternative hypothesis (sometimes notated as HA or H1). The only other possible outcome to the prediction is that the variables will not be related, i.e. that there will not be a significant improvement in mildly depressed women compared to men as assessed on the BDI after 6 months of cognitive therapy compared to Prozac. This second type of hypothesis is called a null hypothesis and represents the alternative hypothesis (sometimes notated as HO or H0) (Siegal and Costellan, 2001).

Table 2. Operationalizing the concept (after Trochim, 1999)

Concepts	Concept 1	Concept 2	Concept 3
	Psychotherapy	Psychiatric disorder	Psychiatric medication
Operational terms	Cognitive therapy	Mildly depressed	Fluoxetine
Measurable indicators	6 months	Beck depression inventory	

Alternative and null hypotheses are needed because it is much easier to test the null hypothesis than the alternative. For example, it only needs to be established that there is a significant difference between the rate of detecting tuberculosis in a screened group, compared to non-screened group, to reject the null hypothesis and accept the alternative hypothesis, i.e. the prediction is supported. It does not need to be specified by how much the detection rate differs, just that there is a statistically significant difference.

Usually in research, the researcher is trying to find support for the alternative hypothesis but in some studies, the prediction might be that there will be no difference or change. In such cases, the researcher is trying to provide support for the null hypothesis, rather than the alternative.

One-tailed and two-tailed hypotheses

The alternative hypothesis example above, where the direction of the outcome was specified as a statistically significant improvement in mildly depressed women compared to men, is called a one-tailed hypothesis (Argyrous, 2000). When a hypothesis has been written without specifying a direction of the outcome, this is called a two-tailed hypothesis. A two-tailed hypothesis can be used when the researcher is unsure, or not sufficiently confident, to make a prediction about the direction of the outcome. In this case, the hypothesis would be written to reflect this uncertainty. For example:

'That there will be a statistically significant improvement between mildly depressed men and women as assessed on the BDI after 6 months of cognitive therapy compared to fluoxetine'

Here the hypothesis has been written as predicting a significant difference in outcome between men and women without specifying which group will improve most. In other words, the direction of the outcome has not been hypothesized. The importance of whether a one-tailed or two-tailed hypothesis is tested, is the statistics used to test the hypothesis. This will be discussed in the fourth article of this series, but for a more detailed discussion of hypothesis testing, see Argyrous (2000).

Using the question or hypothesis

Finally, when developing the research question or hypothesis, the researcher will need to consider whether the study is feasible or not, e.g. he/she may have to make a trade-off between rigour and practicality. Will the researcher have the resources to complete the research? Can the required cooperation be obtained from other staff and patients to carry out the project in the specified time? Is the length of the project reasonable? Can the project be completed in the time available? What is the possible impact on day-to-day work and colleagues? Is the research relevant?

With ever-increasing demands on research funds, research must be timely, relevant and cost-effective. The chosen research topic might be of great importance to the researcher, but where does it fit into factors such as current practice, priorities and needs? All research will require ethical approval before starting and issues concerning patient consent will also need to be addressed. ■

Anastas JW, MacDonald ML (1994) *Research Design for Social Work and the Human Sciences*. Lexington, New York

Argyrous G (2000) *Statistics for Social and Health Research*. SAGE Publications, London

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) Inventory for measuring depression. *Arch Gen Psych*. 4: 561-571

Cooper G (2001) Conceptualising social life. In: Gilbert G, ed. *Researching Social Life*. SAGE, London

Fink A (1995) *How to Ask Survey Questions*. SAGE Publications Inc, Thousand Oaks, CA; London

Guildord J P, Frucher B 1(1978) *Fundamental Statistics in Psychology and Education*. 6th edn. McGraw-Hill Inc, London

Meadows KA (2003) So you want to do research? An overview of the research process. *Br J Comm Nurs* **8**(8): 369–75

Miles M B, Huberman (1994) *Qualitative Data Analysis: An Expanded Source Book*. 2nd edn. SAGE Publications Inc, Thousand Oaks, CA; London

Nazroo J, O'Connor W (2002) Introduction. In: O'Connor, Nazroo J eds. *Ethnic Differences in the Context and Experience of Psychiatric Illness: A Qualitative Study*. The Stationary Office, London: 9-16

NHS Centre for Reviews and Dissemination (2001) *Undertaking Systematic Reviews of Research on Effectiveness. CRD Guidelines for the carrying out or commissioning reviews*. CRD Report No 4. 2nd edn. University of York, York

Pope C, Mays N (1995) Qualitative research: reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ* **311**(6996): 42–5

Robson C (2002) *Real World Research: A Resource for Social Scientists and Practitioners*. 2nd edn. Blackwell, Oxford

Siegel S, Castellan NJ (1988) *Nonparametric Statistics for the Behavioural Sciences*. McGraw-Hill Inc, London

Trochim B (1999) *Centre for Social Research Methods*. http://trochim.cornell.edu/ (accessed 18 August 2003)

KEY POINTS

- All research should have a specific and well-defined research question or hypothesis.
- Reviewing the literature is an essential phase in the development of the research question or hypothesis.
- The research question and the information needed to answer it determines the research design.
- Research should be timely, relevant and cost-effective.