The strengths and weaknesses of research designs involving quantitative measures

Journal of Research in Nursing 2005 10: 571
DOI: 10.1177/136140960501000505

The online version of this article can be found at:
http://jrn.sagepub.com/content/10/5/571
The strengths and weaknesses of research designs involving quantitative measures

Wendy Walker  MSc Health Studies; Post Graduate Diploma in Adult Education, BSc(Hons) Nursing Studies, Diploma in Professional Studies in Nursing
Senior Research Fellow/Senior Lecturer
Faculty of Health and Sciences, Staffordshire University

Abstract  This paper presents a critical review of the strengths and weaknesses of research designs involving quantitative measures and, in particular, experimental research. The review evolved during the planning stage of a PhD project that sought to determine the effects of witnessed resuscitation on bereaved relatives. The discussion is therefore supported throughout by reference to bereavement research. Three levels of quantitative research are presented: descriptive, correlational and experimental. The findings suggest that experimental research is subject to a number of methodological limitations that may jeopardise internal and external validity of the research results and, consequently, limit their applicability for practice. Nurses are therefore encouraged to carefully consider the virtues of experimental designs, in their quest for evidence-based practice and in the planning of future research.

Key words  evidence-based practice, experimental research, positivism, bereavement, witnessed resuscitation

Introduction

In today’s political climate of demand for cost and clinical effectiveness (NHSE, 1996; Department of Health, 1997), nurses are increasingly expected to engage in evidence-based practice (EBP). The widespread movement to promote EBP represents a concerted effort to progress away from procedures based on tradition, ritual and routine, to a clinical service informed at every level by evidence that is scientifically derived (Ford and Walsh, 1994; Hicks, 1997; Hicks and Hennessy, 1997). Consequently, the perceived importance and profile of research has never
been higher and the expectation that research will deliver solutions to problems has never been greater’ (Balcombe, 1996: 1206). Notable health service research and development (R&D) initiatives include: funding for research to inform policy and practice (Salvage, 1998), strategies for the dissemination and integration of research such as the Cochrane Collaboration (Lefebvre, 1994) and the establishment of a National Institute of Clinical Excellence as a means of giving ‘new coherence and prominence to information about clinical and cost effectiveness’ (Department of Health, 1997: 3). According to Mulhall et al. (1998), emphasis on effectiveness and efficiency within the health service research and development programme, including discussions on the development of guidelines, give the highest regard to research designs that are less susceptible to bias. The quantitative approach to research is therefore likely to be embraced by those who support this point of view; being founded on the belief that the social world lends itself to objective forms of measurement (Cowman, 1993) and characterised by a set of orderly and disciplined procedures (Polit and Hungler, 1999).

Defining quantitative research

Quantitative research is depicted as the traditional scientific approach to research that has its underpinnings in the philosophical paradigm for human inquiry known as positivism (Polit and Hungler, 1999). Research driven by the positivist tradition is a ‘systematic and methodological process’ (Koch and Harrington, 1998: 884) that places considerable value on ‘rationality, objectivity, prediction and control’ (Streubert and Carpenter, 1999: 7). A distinguishing feature is the collection of numerical data (Jack and Clarke, 1998) that, in turn, can be subjected to statistical analysis (Carter, 2000a). Advocates of the quantitative approach are therefore described as objective scientists (Duffy, 1986) committed to the discovery of quantifiable information (Carr, 1994). Parahoo (1997) identifies three levels of quantitative research: descriptive, correlational and causal; causal referring to experiment as a research design.

Descriptive research

Descriptive research provides an account of the characteristics of individuals, groups or situations (Jack and Clarke, 1998) that may form the first stage of more complex designs (Clifford, 1997; Carter, 2000b). The overall aim is to ‘discover new meaning, describe what exists, determine the frequency with which something occurs and categorize information’ (Burns and Grove, 1999: 24). This is illustrated by Fraser and Atkins (1990) who carried out a telephone survey to identify survivors’ recollec-
tions of emergency nurse activities following the sudden death of a loved one. Using descriptive statistics, the authors reported on the frequency and percentage of nurse activities that either ‘helped’ or ‘would have helped’ survivors cope in this crisis situation. Recommendations for future research included expansion of the study to determine whether there is a relationship between the demographics of victims and survivors’ recollections of helpful and unhelpful nurse activities.

**Correlational research**
In correlational research, the investigator deliberately seeks to examine links (or relationships) between variables without introducing an intervention. The purpose is often to generate hypotheses that can be tested in experimental research (Parahoo, 1997; Burns and Grove, 1999). Evidence of this is seen in the study by Tye (1993). Data collected by means of a structured questionnaire were used to examine the relationship between age, length of professional experience and death education and their effect on qualified nurses’ perceptions of the needs of suddenly bereaved family members. One of the implications arising from this research was the need for education and training relevant to the needs of the suddenly bereaved. This lends itself to the application of an experimental design to determine its effects on staff perceptions.

**Experimental research**
Experimental research provides the framework for establishing a relationship between cause and effect (Roe, 1994, Mulhall, 1994). In experiments, the researcher as an active agent (Polit et al., 2001) uses deductive reasoning to prove or falsify hypotheses (Proctor, 1998). This involves manipulating an independent variable (cause) and observing the outcome on a dependent variable (effect) whilst attempting to hold extraneous variables constant (Newell, 1994). Similarity of subjects is ensured by ‘matching’ cases with respect to an infinite number of characteristics and allocating one from each pair to a control and experimental group on the basis of randomisation (Clifford, 1997; Hicks, 1998; Polit et al., 2001). Only observable facts are relevant and the techniques of inferential statistics produce precise numerical results (Hicks and Hennessy, 1997). A variety of experimental designs have been developed, ranging from the relatively simple before–after design, to the more complex multivariate factorial designs (Burns and Grove, 1999). According to Polit and Hungler (1999), random sampling, manipulation and control are the characteristics of ‘true’ experimental research. An approach to research that embraces these three virtues is the highly valued randomised controlled trial (RCT).
Robinson et al. (1998) carried out an RCT to determine the psychological effect (dependent variable) of witnessed resuscitation (independent variable) on bereaved relatives. Using manipulation, relatives of patients who required resuscitation were given the option to remain with the patient (experiment group) or were not given this choice (control group). The unit of randomisation was the patient undergoing resuscitation. Eligible resuscitations were defined and a chaperone was provided to give emotional support and technical information, adding control to the research design. Outcome measures included psychiatric and psychological morbidity at one and six months after the resuscitation event.

**Strengths of experimental research**

Experimental research is regarded by many as the optimum quantitative methodology for obtaining reliable information about treatment or intervention effect (McMahon, 1994; Mulhall, 1994; Sibbald and Roland, 1998; Donnan, 2000; Richardson, 2000; Polit et al., 2001). Moreover, the randomised controlled trial holds a superior status over other research methods as the ‘gold standard’ of evidence on which to base decisions about healthcare (Knipschild, 1993; Black, 1996; May, 1997). From a review of this literature, it is readily apparent that the power and strength of experimental research is related to control. This involves strict application of standardised procedures to reduce systematic bias and eliminate erroneous conclusions (Hicks, 1998; Burns and Grove, 1999). Control may be exerted in several ways, including: random sampling, inclusion/exclusion criteria, use of a comparison group, subject matching across groups, manipulation of the independent variable, single, double or treble blinding procedures, the use of precise measuring tools and the application of standardised statistical tests in the final analysis of data. The dividend of this control is the researcher’s ability to state with confidence that the outcome produced can only be attributed to the effects of the experiment (Duffy, 1985). This is in stark contrast to descriptive and correlational research, where less rigid approaches to explore and describe phenomena (as it exits) limits the extent to which firm conclusions can be drawn. In other words, the scientific and statistical rigour of experimental research maximises internal validity and increases the probability of generalising the findings beyond the study sample. Despite such exceptional advantages, the literature also points to several methodological limitations that may jeopardise internal and external validity of the research results and, consequently, limit their applicability for practice.
Threats to internal and external validity

Sampling
The sampling technique of experimental research relies on the development of explicit criteria prior to initiation of the study (Duffy, 1985). To qualify as a true experiment, the researcher is obliged to select a sample from the study population and allocate subjects to the various study groups on the basis of randomisation (Duffy, 1985; McMahon, 1994; Clifford, 1997; Polit et al., 2001). Problems may arise, however, in randomised controlled trials when potential participants are not prepared to opt for treatment on a random basis (Brewin and Bradley, 1989; Black, 1996; Silverman and Altman, 1996; Torgerson and Sibbald, 1998). Failure to achieve randomisation may limit the extent to which the study sample is representative of the parent population and, with it, generalisability of the study findings (Torgerson and Sibbald, 1998). Supporting this argument is the study by Stroebe and Stroebe (1989), who identified a range of what they call ‘accepter and refuser characteristics’ that could limit generalisations concerning health and recovery patterns among bereaved people. A further limitation is that the researcher may be unable to match patients with respect to certain physical, psychological and social traits (Hicks, 1998), all of which may have an influence on the outcomes of the research.

Recruitment
Successful random sampling also depends on a sufficiently large sample (Thompson, 1999). Difficulty in recruiting subjects to participate in a clinical trial is, however, remarkably common (McMahon, 1994). An important issue identified by Wilson and Rose (1998) is the role of ‘gatekeepers’ in the process of recruitment. For example, once an intervention or treatment becomes widespread, or in the absence of any preliminary evidence to suggest the intervention or treatment may be beneficial, it may be difficult to recruit participants or clinicians who are prepared to test alternatives (Black, 1996; Fairhurst and Dowrick, 1996; Getliffe, 1998; Sibbald and Roland, 1998). The latter could certainly apply to the phenomenon of witnessed resuscitation. Limited empirical evidence is available to guide decision-making in practice, and support for this intervention amongst healthcare professionals is not yet universal (RCN, 2002). Recruitment difficulties may be overcome by carrying out the study in multiple geographical locations (Burns and Grove, 1999), although Getliffe (1998) warns of their complexity and calls for stringent monitoring by the project management team. Power analysis can also be used to calculate and guide sample size. However, its use relies on having some estimation of the degree of change expected in the dependent
variable and is therefore limited to studies where research on the subject already exists (McMahon, 1994).

**Mortality, maturation and history**

Despite application of systematic and protocolised procedures, experimentation with humans is subject to a number of external influences that may dilute the study results. Once enrolled on a study, subjects may fail to comply with treatment or follow-up arrangements (Donnan, 2000), particularly if the study is a longitudinal research design (Watson, 1998). Mortality or attrition raises serious doubts about observed effects due to differences in the characteristics between the subjects who ‘drop out’ and those who remain in the study (Nieswiadomy, 1998). Changes within subjects rather than as a response to a treatment or intervention (maturation) may also have an intense effect on the study results (Haughey, 1994a; Oldham, 1994; Nieswiadomy, 1998). For example, a researcher may credit reduced symptoms of depression among bereaved relatives to the intervention of witnessing resuscitation, whereas changes in mental health may have resulted from personal adaptation to bereavement and loss, in spite of the intervention. Staying with this example, relatives may be exposed to a media campaign during the course of a study that portrays the stages of grief and gives insight into therapeutic interventions that seek to enhance mental health. When an event other than the experimental treatment or intervention influences the dependent variable, the study is said to be threatened by history (Haughey, 1994a; Oldham, 1994; Nieswiadomy, 1998).

**Hawthorne effect**

A further limitation of experimental research is that subjects may change their behaviour or respond in a specific manner simply because of awareness of being observed (Haughey, 1994a; Clifford, 1997). It is for this reason that the researcher may adopt what is known as a ‘double-blind’ technique in which neither the participants nor the experimenter knows which subjects receive the active treatment or intervention (Oldham, 1994). This procedure is viewed as a major strength in RCTs to avoid experimenter and subject bias (Hicks, 1998). Its use however, is restricted to studies where treatment or intervention can be disguised, for instance in clinical drug trials (Polit et al., 2001). Consider, for example, research to evaluate the effects of witnessed resuscitation. The relative as a ‘witness’, i.e. ‘a person present; one who sees or hears what happens’ (Oxford Current English Dictionary, 1990) cannot be blinded to the intervention. It also follows that those performing the intervention cannot be blinded for the same reasons. Clearly, this illustrates how efforts to strengthen external validity in experimental research are not always possible due to practical, professional or ethical reasons.
Reductionism
Experimental research relies heavily on the control and removal of contaminating variables within the internal structure of the study and is, therefore, by design, considered to be reliable (Duffy, 1985). However, in striving to attain consistent internal validity, the research location may become so false that threats to external validity increase (Haughey, 1994a). Evidence of this is seen in the study by Baker et al. (2000) who applied a total of 9 inclusion and 11 exclusion criteria in an attempt to meet the RCT’s demand for homogenous groups. However, the authors concluded that their results on family satisfaction with end-of-life care were not generalisable to the experience of families who did not meet their study criteria for enrolment. It is also argued that the inherent reductionism of experimental research is incongruent with the humanistic philosophy espoused for nursing practice (Munhall, 1982; Playle, 1995; Clark, 1998; Leininger, 1998) and that certain aspects of patient care are not amenable to manipulation, easily expressed or measurable (Corner, 1991; Parahoo, 1997; Mulhall et al., 1998; Polit et al., 2001). A case in point is the effects of widowhood on physical and psychological health (Stroebe and Stroebe, 1987). Clearly, we cannot manipulate widowhood. ‘People lose their spouses by a process that is neither random nor subject to research control’ (Polit et al., 2001: 178).

Measurement effects
For results to be meaningful, it is essential that reliable and validated outcome measures are used (Getliffe, 1998). In relation to bereavement research, Tomita and Kitamura (2002) identify 16 measurement devices available for the assessment of grief. However, their ability to differentiate between normal and pathological grief is questioned. Furthermore, while some instruments assess grief in general, others seek to measure grief in specific situations such as reactions to loss induced by sudden death (as may be the case for relatives witnessing a resuscitation event) or after miscarriage or foetal death. This therefore suggests that reliability and validity of research results also depend on careful assessment of the circumstances in which outcome measures will be applied.

Drawing upon the work of Cook and Campbell (1979), Burns and Grove (1999) also stress the importance of statistical conclusion validity in order to prevent distortions of the truth. According to Haughey (1994a), this type of validity can be threatened, for example if the criterion of statistical tests is violated. However, despite the importance placed on statistical significance, this may not represent the clinical or practical importance of the research results (LeFort, 1993; Hollis, 1994; Hicks, 1998). For example, research findings may indicate overall satisfaction with end-of-life care, but this may mean very little in terms of the real
needs of dying patients and their families (Hanson et al., 1997). Conversely, Richardson (2000) argues that many RCTs will produce only moderate differences in outcome, but this difference may be of clinical importance either to the patient or the degree to which study results lead to clinical changes in behaviour.

Ethical issues

Experimental research is subject to a number of ethical considerations, particularly when a double-blind trial is being developed (Clifford, 1997). Apart from the element of deception that is involved, participants allocated to the control group may be considerably disadvantaged (Hicks, 1998) especially when the outcome of treatment or intervention is uncertain or believed to be inferior to existing treatment regimes (Sibbald and Roland, 1998). On the other hand, where uncertainty exists about the effectiveness of current treatment or interventions, Hicks (1998: 22) argues that ‘it is self-evidently unethical not to subject these to rigorous RCT scrutiny, since without this, potentially harmful procedures may persist with all the negative consequences this may have for patient well-being’. There are also ethical problems associated with leaving some people untreated (Abbott and Sapsford, 1998). Presume, for example, that a researcher was interested in evaluating the effectiveness of a witnessed resuscitation protocol. It would certainly be unethical to withhold interventions for relatives in the control group who asked to remain with their loved one during the resuscitation process. This therefore requires measures to ensure that this group receives the intervention, irrespective of randomisation. Similarly, steps need to be taken to ensure that subjects in the control group receive the ‘normal’ or routine intervention (Haughey, 1994b). An ensuing dilemma, however, is that some participants may have received a less-than-beneficial treatment, when at the end of an experiment the outcome of one treatment or intervention is found to be significantly more effective (McMahon, 1994). Newell (1992) also has misgivings about patient participation in a trial that is insufficient in size to detect a better treatment, particularly as experimentation often involves some element of risk, including adverse effects that may only manifest in the longer term (Black, 1996; Hicks, 1998). When considering the effect that witnessing resuscitation may have on the grieving process, Offord (1998) and Fulbrook (1998) suggest that it is not beyond the realms of possibility for relatives to suffer from nervous shock or psychological injury and concerns for the medico-legal implications of this practice have been voiced (RCN, 2002). The ethics of experimental research therefore demands careful assessment of the risks and benefits that might be incurred and that this information is clearly articulated to prospective participants during the process of obtaining
informed consent. A general guideline is that ‘the degree of risk to be taken by those participating in the research should never exceed the potential humanitarian benefits of the knowledge to be gained’ (Polit and Hungler, 1999: 135).

**Conclusion**

Central to current healthcare provision is the call for nursing practice to be evidence-based, with the underlying assumption that patient care will be enhanced (Hunt, 1996; Hicks, 1998). Taking into account the facts and arguments presented in this paper, it seems reasonable to suggest that research designs involving quantitative measures can make a valuable contribution to the evidence-base required. Both descriptive and correlational designs have a key role to play in the development of new knowledge, generating questions and hypotheses that could form the basis of further research. However, at a time of increased demand for treatment and interventions that produce positive outcomes on health, experimental research and in particular the RCT is seen as the hallmark of scientific enquiry, in that it proffers the possibility of predicting cause-and-effect relations beyond reasonable doubt. Yet despite the application of rigorous procedures, including measures to control systematic error and bias, the use of experimental research is subject to a number of methodological and ethical concerns. Nurses are therefore encouraged to carefully consider the virtues of experimental designs in their quest for evidence-based practice and in the planning of future research.

**Key points**

- Central to current healthcare provision is the call for nursing practice to be evidence-based
- Quantitative research is depicted as the traditional scientific approach to research
- Research designs for quantitative studies include descriptive, correlational and experimental
- The randomised controlled trial is seen as the ‘gold standard’ of evidence on which to base decisions about healthcare
- Despite exceptional advantages of experimental research, nurses are encouraged to carefully consider its virtues
References
Walker The strengths and weaknesses of research designs involving quantitative measures


Journal of Research in Nursing 10(5)

Thompson, C. (1999) If you could just provide me with a sample: examining sampling in qualitative and quantitative research papers. Evidence-Based Nursing 2:3, 68–70.

Correspondence should be addressed to:
Wendy Walker,
Faculty of Health and Sciences,
Staffordshire University,
Blackheath Lane,
Stafford

582