Transforming Research Methods in the Social Sciences

Published by Wits University Press

Transforming Research Methods in the Social Sciences: Case Studies from South Africa.
Project MUSE. muse.jhu.edu/book/88982.

For additional information about this book
https://muse.jhu.edu/book/88982

For content related to this chapter
https://muse.jhu.edu/related_content?type=book&id=3356598
Introduction

Longitudinal research designs are commonly employed within the social sciences. Typically, such designs are used where the focus is on stability and change: for example, in developmental, social and cognitive psychology. Longitudinal designs are predominantly defined by the element of time, since the emphasis is on data collected at different time points, generally from the same participants. In this chapter, we describe and critique the main longitudinal research designs, together with factors that may impact on the quality of data and suggestions for how to counter these. Illustrative examples of longitudinal research are drawn from our project, the Road and Aircraft Noise Exposure on Children’s Cognition and Health in South Africa (RANCH-SA) study (Cockcroft, Seabi, Goldschagg & Greyling, 2013; Seabi, Cockcroft, Goldschagg & Greyling, 2012, 2013; Seabi, Goldschagg & Cockcroft, 2010a, 2010b).

The RANCH-SA study took advantage of a naturally arising experiment when the old Durban International Airport was decommissioned and a new airport built in a different location. In this project, we were interested in whether air traffic noise exerted a noticeable impact on school children’s reading comprehension, attention and working memory. A control sample was drawn from children attending schools within the same sociodemographic context, but who were not exposed to aircraft noise. Data were collected from the same participants in wave 1 (when the airport was operational) and waves 2 and 3 (when the airport had been closed). Environmental sound levels of aircraft taking off and landing over the noise-exposed schools were measured prior to the airport’s closure. Follow-up measurements were made on later visits to schools when the airport was decommissioned (Cockcroft et al., 2013; Seabi et al. 2010b; Seabi et al., 2012, 2013). We refer to this study throughout the chapter to provide practical examples of theoretical issues related to longitudinal research designs.

Theory and purpose of longitudinal design

Longitudinal research is different from cross-sectional and correlational research in terms of design, the data yielded and the methods of analysis used (Menard,
While cross-sectional research is useful for determining the direction of relationships between variables through correlations, it cannot establish causal patterns. In contrast, longitudinal research, if carefully designed, can address questions concerning causation.

A key feature of longitudinal research is that data are collected at two or more points in time. Another feature of such designs is that the participants from whom data are collected at time one are the same participants from whom data are collected at time two, and at further time points. There may be some addition or deletion of participants over time (discussed later) but the intention is to retain the same participants for the duration of the study. Consequently, the analysis of the data will entail a comparison between these time points. This enables the detection of change or difference from one time point to another, and for causal patterns to be determined. For example, the RANCH-SA study entailed the collection of data across three time points/waves: year one when the airport was operational, year two when the airport had been decommissioned and year three as a follow-up. In all years, the same children were assessed, commencing when they were in Grades 4, 5 and 6 and concluding when they were in Grades 6, 7 and 8. Thus, the RANCH-SA study had both cross-sectional and longitudinal components in its design (Cockcroft et al., 2013; Seabi et al., 2010a, 2010b; Seabi et al., 2012, 2013).

Longitudinal research may be descriptive, explanatory or both. Descriptive longitudinal research seeks to illustrate how a phenomenon changes over time. In such research, the researcher attempts only to describe the form of change over time (e.g. whether it is linear or non-linear). In contrast, explanatory longitudinal research attempts to identify the underlying cause of the change process through the use of one or more predictor variables. Both descriptive and explanatory aspects are important to consider when conceptualising longitudinal research. It is usually necessary to first have an accurate and detailed description of the change trend before attempting to explain it. Therefore, before the first wave of data is collected, it is valuable to conceptualise the form of change in the variables of interest and then to formulate the theoretical causes of change in those variables. In addition, the nature of the relationships among independent, dependent and/or mediating variables must be carefully specified. This requires precision about which variables are expected to change or why they are changing (Ployhart & Vandenberg, 2010). Such information would be based on a thorough review of available literature in the area. In our study, we had to theorise about why and how noise may affect different elements of cognition in school children. Thus, we had to know about typical developmental trajectories of memory, reading comprehension and attention in children in Grades 4 to 8. Then, we had to hypothesise about how noise may change these trajectories and indicate the nature of that change. We also had to consider how the particular demographic characteristics of our sample (e.g. 44% came from low socioeconomic backgrounds, all attended English-medium schools, 49% spoke English as their mother tongue, 45% spoke English as a second language, 40% were female) could impact these developmental trajectories (Cockcroft et al., 2013; Seabi et al., 2010b; Seabi et al., 2012, 2013).
Types of longitudinal design

Menard (2008) describes four basic longitudinal designs, discussed below.

**Total population design**

In this design, the total population is measured at each time point of the study. An example of this type of design is Census data collection. The population members may change slightly over time because of attrition (participant dropout) and addition (new participants joining the population). However, if the time points between data collection are not too long, the majority of participants will generally stay the same.

The next three designs are subsets of the total population design as they use samples drawn from the total population. The difference between the remaining three designs is in the extent to which the same or comparable participants are measured or studied from one time point to the next (Menard, 2008).

**Repeated cross-sectional design**

Here the same data are collected at two or more time points from different but comparable participants. For example, if the RANCH-SA study had instead collected data in each year from children at different noisy and quiet schools in Grades 4 to 6 who were similar to the original group in terms of home language, school language, socioeconomic status, age and gender, this would have constituted a repeated cross-sectional design. Such a design meets only one of the criteria of longitudinal research provided at the outset of this chapter, namely the collection of data at different time points. It fails to meet the criterion that such data be collected from the same participants, and so purists do not regard this as a longitudinal design (Menard, 2008; Nesselroade & Baltes, 1979).

**Revolving or rotating panel design**

In this design, data are collected from a sample of participants either retrospectively or prospectively for several time periods; then some participants are dropped and replaced with new participants (hence ‘revolving’). The revolving aspect deals with potential attrition and threats to validity resulting from repeated measurement in prospective studies, or with periods of extended recall in retrospective studies. Retaining a core set of participants over several measurement time points allows for short-term measurement of change, and separate, as well as comparative, analyses of each set. Replacement of the group of participants that was dropped in a measurement period with a new but comparable group allows for analysis of long-term patterns of aggregated change, similar to the analysis in total population and repeated cross-sectional designs. Since some participants are subject to repeated measurement and others are not in such a design, comparisons are permitted that can show whether the repeated measurement may be producing bias in the data. For example, the development of rapport and trust between researcher and participant over time may facilitate the disclosure of information that did not occur at the first measurement point. Alternatively, the knowledge that participation requires the completion of long and tedious tests may result in lowered motivation and cooperation at
the second and subsequent measurement points, which may be reflected in data differences (Ployhart & Vandenberg, 2010).

**Longitudinal panel design**

If participants are not replaced and data are collected over several time points, the research is using a longitudinal panel design. There are two types of longitudinal panel designs, retrospective and prospective. In a prospective panel design, data collection occurs at two or more time periods, where each time period entails a series of measurements. Researchers who take a strict view of longitudinal research argue that this is the only true longitudinal design, as it allows for both intra- and inter-individual changes to be measured (Menard, 2008). The RANCH-SA study employed a prospective longitudinal panel design, since data were collected at three separate time points, using the same participants (from noise-exposed and quiet schools); the same measures of reading comprehension, attention, working memory and annoyance reactions to noise were administered at each point (Cockcroft et al., 2013; Seabi et al., 2010a, 2010b; Seabi et al., 2012, 2013).

Prospective panel designs have several strengths, including that they allow for data to be collected concurrently with the event under investigation; they allow for continuous measurement of events and changes that would be too difficult to measure in retrospective studies; and prospective constructs, such as goals and aspirations, can be measured and their actualisation later on can be determined. Limitations of this type of design include participant attrition and/or non-response; repeated data collection from the same participants may affect their responses, making them less representative of the typical individual; they are expensive in monetary and human resource terms; and they require long time periods (from years to decades) (Scott & Alwin, 1998). Although this type of design is most effective for studying causal relationships, it is important to remember that not all relationships are causal. Causal relationships are quite difficult to determine due to a variety of extraneous and confounding variables that exist in a given sociocultural environment. This means that unless a true experimental design is used, causality can only be inferred, never proven (Singer & Willett, 2003). Despite these limitations, prospective longitudinal designs are still regarded as ideal, since they are the most reliable and valid way of collecting longitudinal data.

In retrospective panel designs, data collection may occur only once at a single period, but the data are collected for two or more periods (prior to, or during, the period in which the data are collected). For example, commonly in such designs, participants are asked to recall episodes and events from their lives (hence ‘retrospective’) and are then also assessed at the time of data collection. Since data are collected at one period only, retrospective panel designs are generally less expensive and time consuming to undertake than prospective designs. Some limitations of this design are the fallibility of human memory (i.e. respondent recall) and selection biases (since only survivors are interviewed) (Scott & Alwin, 1998).

Other longitudinal designs are generally variants of these four designs – for example, the accelerated longitudinal design. In the latter design, several age cohorts are sampled, and then longitudinal data are collected for each cohort.
Section One: Quantitative methods

The aim is to study age-outcome trajectories for a broad age span over a relatively short duration of months rather than years (Menard, 2008). The particular design that is used has implications for the types of data analyses that are possible, an issue taken up later in this chapter.

Factors that may affect the quality of longitudinal data

Prior to designing a longitudinal study, it is important that the researcher is aware of issues that can impact negatively on the quality of the data collected. By outlining these, it is hoped that researchers will be able to make well-informed decisions about ways of minimising their effects.

Attrition and non-responding

Attrition (also referred to as panel mortality), or the loss of randomly assigned participants or data following the first wave of data collection, is inevitable in longitudinal research with human participants. A particular type of non-response can affect data quality in several ways. The first is through the reduction in sample size, which becomes a particular concern for studies with small subgroups within the population. If cell sizes become too small, the types of analyses that can be performed are restricted (Laurie, 2008). Secondly, attrition can affect the generalisability of findings if the remaining participants differ in some way from those who dropped out, as this would produce a final sample that is not representative of the originally sampled population. This is referred to as differential or systematic attrition and can lead to attrition bias (Gustavson, Von Soest, Karevold & Roysamb, 2012; Laurie, 2008).

While attrition in longitudinal studies tends to be random, particular groups have been identified that are more likely to drop out, such as younger individuals who are more geographically mobile (Laurie, 2008). Other causes of attrition in longitudinal studies are due to inability to contact participants and/or refusal to participate in subsequent waves of the research. The frequency of data collection may also affect attrition. While it is easier to keep in touch with participants over short time periods, this also escalates the costs of the study and reduces the longitudinal inferences that can be drawn from the data. Also, participants may experience fatigue through repeated interviewing/testing. The length, complexity and format of the measures that participants have to undertake, as well as how relevant the content of those measures is to participants, are also factors in attrition (Gustavson et al., 2012).

Several methods have been proposed for reducing attrition. One is the use of refreshment samples – new, randomly sampled respondents who are given the questionnaire/measurements at the same time as a subsequent wave of the panel – which offer information that can be used to identify and adjust for potential bias due to attrition (Deng, Sunshine Hillygus, Reiter, Si & Zheng, 2013). Ensuring that additional contact details of next of kin are collected besides those of the
participants will help with tracking should they move away during the course of the data collection. Regular keeping-in-touch exercises with the participants, through a phone call, email or letter, help to maintain a sense of belonging. These exercises include thanking respondents for their participation and informing them of the study findings. Laurie (2008) provides a set of additional ways in which attrition can be minimised in longitudinal survey panel designs. These include refusal conversion, which refers to instances where participants who have initially refused to continue participating are contacted again at a later period to determine whether anything can be done to encourage participation (i.e. to convert their refusal to participate into acceptance). The reasons for refusal to participate are varied, such as being too busy, disliking the interviewer, interview questions and/or tests, and some of these can be adjusted. Another method of encouraging participation and reducing attrition is to use incentives. However, there are mixed views about their effectiveness. Most evidence suggests that some incentive is better than none, and that monetary incentives given unconditionally in advance of the interview/testing are most effective in increasing responding (Laurie, 2008). If participants have to travel in order to participate, it seems morally appropriate to reimburse their travel expenses. A concern when using incentives is that poor-quality data may be collected from less cooperative participants who would not have participated if no incentive had been offered. In the RANCH-SA study, monetary incentives were inappropriate and financially unviable. Instead, we provided each participant with refreshments during data collection, as well as a small token in the form of a pen.

Since attrition is unavoidable in longitudinal studies, the best way to deal with it is to always report the total number of randomised participants who complete and who do not complete the study’s protocol. Usually, attrition is reported as a simple descriptive statistic (e.g. percent total attrition) on the presumption that lower rates imply better science. This is because best-evidence behavioural interventions require retention rates of 70% or higher in each component (e.g. experimental and control) for an intervention with positive outcomes to be considered meaningful. Attrition does not necessarily signal bias, nor does it confer methodological flaws (Amico, 2009). The primary concern in evaluating threats caused by attrition is the extent to which participants or their data are missing due to random or non-random factors. There are multiple causes for missing data: participants may complete some but not all of the measures; some may complete the measures but leave out certain sections; and some may not complete any measurements at a given time point. It is useful, when reporting attrition, to distinguish between loss occurring because of discontinuation of participants and loss because of requested withdrawal, non-compliance to participation in the treatment/experiment, or failure to obtain data from assessments/measures despite continued service use (in intervention-type studies). For randomised controlled trials, assessments of differential attrition should focus on loss of participants in the treatment component relative to loss in the control component. Our RANCH-SA study suffered a fairly high attrition rate. During the first wave of data collection, a cohort of 732 learners (mean age 11.1 years) participated. A second wave took place a year later after the relocation of the
airport and 650 of the original participants were retained. The third wave of data collection occurred the following year with a significantly reduced sample of 378 participants (mean age 13.1 years). There was high attrition of participants in the third wave because the participants had moved from primary to high school (i.e. new schools) and some principals did not grant permission for their schools to participate in the project. In addition, there was bad weather on the assessment day, which resulted in many learners staying home.

Respondent recall
Respondent recall can be a challenge in retrospective studies as it relies on the accuracy of participants’ memories, and inaccuracy in recall increases as the length between the event and the recall increases. Other factors that influence the accuracy of respondent recall are the saliency of the to-be-recalled event to the participant and the frequency of occurrence. Consequently, when designing a retrospective longitudinal study, the lengths of recall time should ideally not be too large. Other techniques for improving respondent recall include the provision of a landmark event to clearly mark the beginning of the recall period and a double-question bounded recall procedure, where participants are first asked to recall events for a longer time period (e.g. the previous eight weeks), and then for the time period of interest (e.g. previous four weeks) (Grotpeter, 2008).

Sample size
An inadequate sample size can limit the utility of the data collected. The appropriate sample size for a longitudinal design may be determined by using a target (usually set at 95%) for the power of a statistical test to be applied once the sample is collected, or by using a confidence level to determine how accurate a result is likely to be and what the degree of certainty or error would be. With quantitative longitudinal designs, generally the greater the sample size, the greater the chance of identifying a statistically significant result. However, very large samples also carry dangers in that minute deviations from the null hypothesis can emerge as statistically significant, even if they have no practical significance. It is important, therefore, to calculate the minimum number of participants and time points needed in order to answer the research question(s), and to have some idea of the kind of change from the null hypothesis required to be practically important, before commencing the study. In planning a longitudinal study, it is critical to estimate the power of the study to detect certain effects, such as treatment effects in intervention studies, as studies with low power cannot be repaired after the data collection. The possibility of a 30% attrition rate over time should also be considered when calculating the optimal sample size (Amico, 2009; Laher & Botha, 2012).

Number and spacing of measurement points
Another factor that may affect the quality of longitudinal data is the number and spacing of the repeated measurements. The minimum number of repeated measures for a longitudinal design is reported to be three (Chan, 1998). A limitation of two time point studies is that any change from time one to time
two is by default linear, making it impossible to determine the form of change over time. With a difference between two times, it is not possible to determine whether change was consistent or delayed, or whether it plateaued and subsequently changed. The second limitation to using only two time points is that true change and measurement error are confounded. One may erroneously conclude that there was a true change between time one and time two, when measurement error may have depressed scores at time one and raised them at time two (Rogosa, 1995; Singer & Willett, 2003).

**Secondary data**
Many longitudinal studies use secondary data. This eliminates the cost of data collection, but introduces other issues concerning the variables and how they were measured and coded in the database, which limits the types of questions that can be asked of the data. For example, the database may be based on interviews or tests that are different from those that the researcher would ideally have used. The other issue is that the researcher has no control over the quality of the captured data.

**Threats to reliability and validity**
When tests of psychological constructs are used in longitudinal studies, the reliability of those constructs should be considered. This refers to the extent to which consecutive measurements of this construct yield the same result, given that the underlying score on the construct has not changed (Cronbach, 1984). Every score contains a true score as well as a certain amount of error, which may include bias (i.e. systematic error, such as an intelligence test that consistently underestimates all tested individuals’ IQ by 15 IQ points, thus reducing the validity of the measure but not its reliability) and random error, which is not consistently due to over- or underestimation (and which reduces both validity and reliability). An example of the latter would be a person’s IQ score that reflects cognitive ability, administrator effects (such as experience in administering the test), environmental factors (noise, room temperature), personal qualities (how the participant is feeling that day), as well as the participant’s familiarity with IQ tests (test wiseness). The aim of research is to maximise true score variance relative to error variance. While some biographical variables can be measured with high levels of reliability, such as participants’ highest completed grades, birth dates and gender, other variables cannot always be measured with the same level of reliability, such as attitudes, perceptions and socioeconomic status (due to the highly variable ways in which this construct is measured). For example, in the RANCH-SA study, socioeconomic status was assessed by the percentage of participants who were eligible for free meals at school (44%), since there is a significant correlation between the free school meal ratio and a range of Census indicators representative of socioeconomic status (Cockcroft et al., 2013; Seabi et al., 2010b; Seabi et al., 2012, 2013; Stansfeld et al., 2005). In South Africa, the criterion for a child to be eligible for a free school meal is that the child’s caregiver receives a government social grant.

As with socioeconomic status, measures of cognitive ability may have a large error component. Multi-item measurement is advised where a large error component is anticipated, as with the cognitive measures used in the RANCH-SA study.
Since these errors are presumed to be due to random factors, they should cancel each other out, and thus multi-item measures will give a more reliable estimate of the participants’ true scores on these cognitive abilities than single-item measures would (Menard, 2008).

Longitudinal designs are vulnerable to several threats to validity, such as sample selection, attrition, construct operationalisation and regression to the mean. Some of these have already been addressed. In terms of construct operationalisation, the difficulty is that it may be possible to operationalise the same construct in different ways across the lifespan. Whenever the way in which a construct is measured changes in longitudinal research, there is no certainty whether the change results from change in the actual construct being measured or from change in the measurement of the construct. When conducting lifespan studies, the same measurement may not be valid at different life stages. For example, if we continued the RANCH-SA study into adulthood, we would need to change the measures of reading comprehension, memory and attention from those appropriate for children and adolescents to those appropriate for adults. Thus, it is important to be aware that the change of measurements in itself may induce a change in performance. For this reason, longitudinal research emphasises consistency of measurement whenever possible, and caution when measurement occurs over periods of rapid developmental change, such as infancy, early childhood and adolescence (Rogosa, 1995).

Longitudinal data analysis

Prior to undertaking any analyses, it is important to understand how missing data can be dealt with, since attrition is inevitable in longitudinal data collection. The practice of simply replacing missing data with the mean is not recommended as it distorts the relationship between the missing variable and other variables in the analyses. While somewhat better, conditional mean imputation (i.e. imputing the mean for similar individuals) is also problematic as it fails to account for the variation in the predicted value within the subgroups defined by the variables used to group observations. Foster and Krivelyova (2008) and Muthén, Kaplan and Hollis (1987) discuss a range of useful techniques for dealing with missing data, such as imputation, partial deletion and interpolation.

The ensuing analytic tools used for longitudinal research will depend on the data that have been collected, the sample characteristics, as well as the particular research questions that the study set out to answer. Consequently, there are many quantitative and qualitative analytic techniques that can be utilised, but it is beyond the scope of this chapter to detail all of them. Below, we provide an illustrative example of the statistical analyses of data from a longitudinal panel design with an experimental and control group, drawing on our experience from the RANCH-SA study.

The first stage in analysing longitudinal data is to provide a basic description of it, generally derived from the descriptive statistics – for example, describing the typical performance of each sample at each time point, often in terms of
means, standard deviations and frequencies. Thereafter, analyses focus on comparisons of differences both within (intra-individual) and between (inter-individual) groups. In terms of intra-individual comparisons, separate analysis of the control group can allow for typical developmental trajectories to be studied. Such an investigation would compare the performance of the control group at each wave of the study. These data, together with existing research, would allow for a priori hypotheses about the direction of these developmental trajectories to be established. Inspection of individual and overall developmental patterns may also contribute to the choice of analytic models. In such a single-group analysis, it is valuable to first explore whether the control population exhibits any of the post-intervention changes in trajectories that are hypothesised to be due to treatment/experimental conditions. For example, in the RANCH-SA study, we first had to determine the trajectory of reading comprehension, attention and working memory for the control (non-noise-exposed) group, across the three waves of the study (i.e. did these improve, deteriorate or remain stable?), before we compared this group's performance with that of the noise-exposed control group (Cockcroft et al., 2013; Seabi et al., 2010b; Seabi et al., 2012, 2013). If such changes are found for the experimental group alone, they are more clearly attributable to the treatment and not to developmental issues. Thus, the experimental group should also be analysed separately. Here, the basic trajectory form (linear, non-linear) may be investigated. The treatment/experimental condition may induce curve shapes different from those in the control group. Change within each group may be qualitative and/or quantitative and determining this would depend on how the data were collected and measured. Event history analysis is a useful set of techniques for describing, analysing and predicting the timing of qualitative change (see Menard, 2008, for detail).

After the within-group comparisons have been conducted, between-group analyses are performed where the control and experimental groups are compared. For example, in the RANCH-SA study, the main research question was whether or not aircraft noise has an impact on reading comprehension. Statistically significant differences were found between the experimental and control groups on reading comprehension \((F(1) = 8.416, p < .004, d = 0.17)\) and working memory \((F(1) = 12.731, p = .001, d = 0.23)\), in favour of the control group (Cockcroft et al., 2013; Seabi et al., 2013). Evidence that reading comprehension was affected by the air traffic noise came from a significant improvement in the reading comprehension of the children attending the noise-exposed schools in the two years following the closure of the airport (Seabi et al., 2012, 2013). Historically, such between-group comparison methods in longitudinal designs include multiple regression analysis and analysis of variance. These methods have advantages, including the ability to examine potential interaction effects. A subsequent question in the RANCH-SA study was whether aircraft noise and home language interact to influence participants’ performance on reading comprehension. A statistically significant interaction was found \((F(1) = 25.621, p = .0001)\), with a large effect \((d = 0.95)\), suggesting that the negative effect of noise on reading comprehension was worse for second-language English speakers (Seabi et al., 2012). However, these methods of difference testing have limitations as they only
analyse mean-level changes and treat random differences between individuals as error variance. Variants of the traditional longitudinal designs described earlier may contain nested data structures, often in the form of various panel designs, which are effectively analysed using hierarchical linear modelling, mixed effects modelling, random effects modelling or random coefficients modelling. These longitudinal data analyses are based on statistical procedures that combine the various nested components. For example, an expected trajectory over time may be created in which the expected values are maximum likelihood estimates and formal tests of these hypotheses are run. Longitudinal structural equation models (SEMs), such as latent curve analysis and latent variable SEM, are also valuable. These formalise theoretical ideas into statistical models against which the data are compared to determine which provide the best fit (McArdle, 2012).

Other analyses include time series analyses, which are generally used with aggregated data from many participants at many time periods – for example, a school’s performance on tests of literacy and numeracy over a five-year period. This type of analysis may be used with data from a total population design. The simplest of these analyses track changes over time in a chart with percentages or absolute change depicted between data points (Menard, 2008). The causes or correlates of these depicted changes are typically analysed with linear regression models.

In the social sciences, the data are often psychological self-report measures, rater-captured measures or psychological tests of a range of human functions, such as attention, memory, reading comprehension, personality, emotions, attitudes and perceptions. When the timespan of the study covers many years or involves participants who are experiencing rapid changes (infants and adolescents), the continued validity of these measures over time is important (as discussed earlier). The tests that were designed to tap reading comprehension in our sample of Grades 4 to 6s are unlikely to measure this construct in the same way in the same sample when they reach Grade 10. The problem of the changing relationship between measures and the latent variables underlying those measures is linked to measurement/factorial invariance. Factorial invariance is evident when the factor structure remains the same across multiple time points (or multiple populations). Confirmatory factor analysis is most commonly used to investigate factorial invariance (Millsap & Cham, 2012).

Thus, there is a range of techniques for analysing longitudinal data, a few of which have been described. The choice of analytic technique should be informed by the research design, sampling characteristics and research questions.

Qualitative longitudinal research

Longitudinal research is typically associated with quantitative studies, and this is seen in the considerably larger literature devoted to quantitative than to qualitative longitudinal research (Holland, Thomson & Henderson, 2006; Saldana, 2003). However, qualitative longitudinal studies are equally important, as they may uncover ‘why’ and ‘how’ questions that quantitative studies cannot always
address, or they may uncover interesting data that quantitative studies may ignore as statistically non-significant. In the social sciences, qualitative longitudinal designs often follow an ethnographic approach, with attention to the investigation and interpretation of the process of change over time in social contexts. Such designs aim to capture the interplay between time and sociocultural context, offering a nuanced understanding of phenomena that evolve slowly, such as the personal experience of illness (Carduff, Murray & Kendall, 2015).

As with quantitative designs, their qualitative counterparts involve at least two waves of data collection over an extended period (typically a year or longer) using the same participants. The theoretical approach (e.g. grounded theory, phenomenology) that guides the research question(s) and which informs the data collection and analysis is what defines the study as qualitative rather than quantitative. While quantitative longitudinal research focuses on data from groups or aggregations of people, qualitative longitudinal research tends to focus on specific individuals. Thus, there are fewer participants than in quantitative designs as the focus is on generating rich, detailed information. This feature makes qualitative longitudinal research just as time and resource intensive as quantitative longitudinal research (Saldana, 2003).

Qualitative longitudinal research allows participants to reflect on the presence or absence of changes experienced since the previous interview. In this sense, qualitative longitudinal data are iterative, so that follow-up interviews often draw on what was learnt previously to understand what has changed about specific events, periods or feelings in order to tell a story over time (Carduff et al., 2015). To ensure that reliable data are captured, care should be taken that the interviewing style is not too repetitive or that respondents become conditioned to particular questions over time.

Ethical issues in longitudinal research

Several qualities of longitudinal designs introduce ethical complexities. These relate to the prolonged duration of the data collection, the often assertive methods used to achieve and maintain high follow-up rates, issues related to compensation for participation, and the potential ambiguity regarding the relationships between researchers and participants (Lessof, 2016; Scott, 2005). Since longitudinal studies rely on repeated follow-up of participants, often over long periods, researchers generally go to great lengths to retain participants. Respect for participants must be maintained, no matter how pressured researchers feel to maintain high follow-up rates. Contacting participants in a manner beyond the scope of what was provided in the consent form – for example, by asking a school or business to help locate a participant – would be ethically inappropriate and violate confidentiality and privacy, unless the researcher has specifically asked permission to use those routes should the participant be unavailable at subsequent waves of the study. In this regard, care must be taken to avoid disclosing either the names of other contact persons or the nature of information provided by these individuals or organisations. Any cross-communication may only be
done with the permission of the participant. When participants cannot be found using the original tracking data, alternative methods such as the use of public domain records and databases may be useful. The ethics of using such databases have not been well documented, and it is best to include this approach in the informed consent. During the initial informed consent, participants must be fully informed of the methods and intensity of follow-up procedures that will be employed, and notified of their importance to the study; and researchers should be sensitive to the timing of each point of contact, and, if it is not ideal for the participant, should arrange a more convenient time. Participants have the right to terminate their involvement in the study at any point, and researchers are obliged to accurately and clearly communicate the effect that disengagement will have on the participant in a manner that is not coercive (Lessof, 2016).

Incentives are often used in longitudinal studies to encourage participants to return, and to pay for their time and any costs that they may incur, such as travel expenses. Researchers should be aware that, particularly when working in low socioeconomic contexts, a high level of compensation could be a source of hidden coercion. However, this needs to be balanced with evidence that a cash incentive produces higher rates of follow-up contact and generates higher levels of satisfaction related to participation. Participants often view cash as a more valued, dignified and convenient form of payment for their time than alternatives, such as food or gift vouchers. Consequently, the cash amount should be reasonable compensation for the participants’ time and effort, and not excessive (Cottler, Compton & Keating, 1995).

Another ethical issue particular to longitudinal studies is that it is often difficult for researchers to maintain neutrality when they are in regular contact with participants over long time periods. The purpose of neutrality is to ensure the objectivity and integrity of the data. By expressing disappointment, disapproval or other emotions, the researcher may inadvertently shape the information that a participant provides, particularly in interview data. Balancing a non-judgemental attitude with appropriate levels of interest and concern is important in determining a participant’s willingness to participate over the duration of a longitudinal study (Lessof, 2016).

What is considered best practice in ethical research can change over the duration of a longitudinal study. For example, technological developments for data capture, analysis and linkages to administrative data may raise new ethical questions that were not foreseen at the outset of a study (Lessoff, 2016; Scott, 2005). Consequently, informed consent procedures should be repeated at each wave to account for such developments.

Conclusion

The purpose of longitudinal research is to describe and explain patterns of change. Quantitative studies attempt to establish the direction and magnitude of relationships between variables, often with a focus on the causal nature of such relationships (Menard, 2002, 2008). Thus, the emphasis is on change
within participants over time (growth trajectories), and often also on differences between participants, commonly a control and experimental group (Bollen & Curran, 2006; Singer & Willett, 2003). Qualitative longitudinal studies aim to capture the rich, graduated interplay between time and sociocultural context, which evolves slowly (Holland et al., 2006). Planning and implementing longitudinal studies requires careful consideration of many factors, such as the phenomenon under investigation, sampling, measurement intervals, potential threats to validity and reliability, appropriate choice of data analysis, as well as cost and duration of the study. Longitudinal studies introduce ethical issues pertaining to the boundaries between participants and researchers due to the long-term nature of contact, the type of compensation given for participation, and the appropriateness of methods used to trace participants. One of the primary applications of the findings of longitudinal studies is in policy development, evaluation and change. For example, results from the RANCH-SA study may be used to inform urban planning laws regarding the location of airports relative to centres of learning.

References
Section One: Quantitative methods


Longitudinal designs: The RANCH-SA study


