Sage Research Methods

Beginning Quantitative Research

For the most optimal reading experience we recommend using our website. A free-to-view version of this content is available by clicking on this link, which includes an easy-to-navigate-and-search-entry, and may also include videos, embedded datasets, downloadable datasets, interactive questions, audio content, and downloadable tables and resources.

Author: Malcolm Williams, Richard Wiggins, Paul R. Vogt
Pub. Date: 2024
Product: Sage Research Methods
DOI: https://doi.org/10.4135/9781529682809
Methods: Data mining, Quantitative data collection, Sampling
Keywords: attitudes, social problems
Disciplines: Sociology, Criminology and Criminal Justice, Business and Management, Communication and Media Studies, Education, Psychology, Health, Social Work, Political Science and International Relations
Access Date: June 27, 2024
Publisher: SAGE Publications Ltd
City: 55 City Road
Online ISBN: 9781529682809

Starting With the Basics: From Research Problem to Variables

Research problems

The philosopher of science Karl Popper often said that we are students of a problem, not a discipline. All interesting and useful science begins with a problem to be solved. In social research, this is often a 'social problem', for example, why do young women not enter engineering professions in some countries but not others (Godfroy Genin & Pinault, 2011)? Does teenage drug taking lead to later relationship problems (Newcomb & Bentler, 1988)? How many homeless people are there in a particular city (Williams & Cheal, 2001)? Sometimes the problem may be more methodological, for example, can new social media methods better predict crime levels than traditional surveys (Williams et al., 2016), or it may 'test' some existing **theory**, such as migrants have greater resources in their originating society than non-migrants (Musgrove, 1963). An interesting characteristic of social research is what is a problem in one time or place may not be a problem in another time or place. So often research is comparative to test whether a 'problem' is location or time-specific. For example, see Zmerli and Hooghe (2011).

Not all problems can be successfully tackled through social research, either because (as in all science) we simply don't have the methodological capability or they are not problems that can be resolved with the resources available. Research questions often arise out of perceived social issues or problems that are considered of sufficient importance to a particular society to be the focus of social research, and consequently, it is these which often attract funding (more of this later), so what is a 'problem' (Bulmer, 2015)? Rarely are topics of interest completely novel; they usually have a history and quite frequently research has already been conducted and a body of knowledge exists in the research literature (we will say more about this below).

How do you identify a research problem? For many researchers, you just don't have to, because you have been hired to work within a research team as a researcher or doctoral candidate and are therefore investigating an identified research problem, but if you are a novice social researcher, needing to come up with your own ideas, there are a few things to consider.

Often, the problem will start off as something which interests you. So has it interested others? Has there been

other research, either directly in the problem area or in a related one? How might this differ from what you want to research, either in terms of the population of interest, the sample or the methods used? Could you, at least partially, replicate a previous study? Sometimes your problem might be too 'broad', and you may wish to pick one specific and 'researchable' aspect of it. For example, each year one of us regularly supervises two or three undergraduate/masters research dissertations on the problem of 'homelessness'. You can't just research the 'problem of homelessness', but rather a very specific part of the problem needs to be identified. What have been the 'routes' into homelessness in a particular city (and how might that differ from other locations)? What form does being homeless take, in a particular location? What are public attitudes toward rough sleepers? And so on.

Research questions

Research problems need to be expressed in the form of research questions that can be 'operationalised' into **hypotheses**, then measurements. A research question is simply a statement of what we want to know. It should be expressed fairly briefly, no more than a sentence or two. It may be quite a specific question that directly represents the research problem, such as why are there fewer women in senior academic roles than men. It may be a very localised and small-scale study, or one which aims to make generalisations across countries, comparisons between countries or through time. Some research problems, particularly those that are explored in larger studies, may have more than one research question, in order to explore a quite broad area of concern or interest. The above problem, concerning women entering engineering, is a problem for certain countries, but it leads to many specific research questions, such as: Are educational experiences that 'gender' engineering present in countries that have larger numbers of women entering the profession? Do universities have positive action programmes? Are subjects 'gendered' in schools?

Some research questions just cannot be answered, or answered in the context of their being posed. For example,

> *Do young people feel less optimistic about their future?* This is far too vague. In principle, something like it could be answered, but defining *optimism*, and *future*, would be necessary – and indeed what counts as a young person. Moreover, to be able to say anything, there would need to be a comparison with the attitudes of previous cohorts of young people (or indeed people in other age groups).
> *Should people become vegan?* The problem here is it begins with a value statement, which implies

either pro- or anti-veganism. But attitudes toward a vegan diet/lifestyle, amongst vegans and non-vegans, could be researched.

Does exposure to violent media images lead to violent behaviour? There are all sorts of problems here. What counts as a violent image? What counts as violent behaviour? But mostly isolating the media images from other stimulants to violence is impossible. Actually, in a large enough study, with enough data, the effects of other stimulants might be controlled for, but this is a question beyond the resources and abilities of most new researchers.

Why do people commit suicide? This is, in principle, unanswerable, because they are already dead! Analysis of suicide notes might provide some reasons, but the reasons might actually be latent. Unsuccessful suicides are often not intended to be successful, so this is not a group that can answer the question successfully.

The above examples show that some things just cannot be done, but the areas of interest may be amenable to research, if answerable questions can be posed.

Whilst the initial framing of the question or questions may be broad-brush, soon it is necessary to turn them into research hypotheses and then measures, 'operationalised' within a survey or experiment – in other words, turned into measurements that can empirically answer the question. The population of interest must also be defined. So, for example, an exploration of student views on research methods teaching would lead us to ask which subject, or subjects, are the students studying? Are they undergraduate or postgraduate, in one university in one country or a range of universities in one, or more, countries? The simple rule is that a sharp question will provide a sharp answer. The opposite, of course, is true!

The role of theory in research

Quite often, the research question is grounded in a theory, which you are testing, either in an amended form or in a new context. All research implies theoretical assumptions, but good research sets out by making those assumptions clear. So what is a theory?

A 'theory' is a proposition, or set of propositions, about relationships between phenomena. It proposes that something is, or isn't the case. In particular, if a theory is true, it implies how the world must be, or acts in certain ways. A theory should have testable consequences. In other words, it should predict or forbid certain things to happen and should be able to derive measurements from those predictions that will allow us to

say whether or not something is the case, or how much or how many of something. Finally, the scope of a theory can vary, and this should be written into it to begin with. Some theories might be very localised, and others have a much broader scope. An example of the former might be that a particular local public transport initiative will reduce traffic congestion. In this case, local factors will play an important role, but nevertheless in formulating such a localised 'theory', one should draw on data from other locations. It might well turn out that one's localised theory is not local at all. An example of more ambitious theory testing might be in the exploration of the social composition of recent migrant refugees from troubled Middle Eastern countries, in order to assess their potential for economic and social integration. You may theorise that the majority of them come from an educated or merchant class in their home country, and if you did you would be testing the theory of migratory elites, proposed by Frank Musgrove, in the 1960s, which broadly states that migrants in a developed, or relatively developed, society will be those who have comparatively greater resources to permit migration (Musgrove, 1963). Now, it may be that you are able to confirm that this was indeed the case, or if you discover that the migrants were not from such classes, the theory would be falsified – at least in this instance, and you would need to show why this was the case and possibly amend the theory to take new localised circumstances into consideration.

Theory comes in different forms. You will see many publications under the heading 'social theory'. Much, though not all, of this constitutes 'grand societal theories' grounded in philosophical assumptions that have a broad explanatory reach, for example, Giddens (1993) 'structuration theory', or Bauman's (1999) 'liquid modernity'. Indeed, whilst these theories may be intuitively plausible, it is very difficult to derive propositions from them that may be empirically tested in order to show whether, or to what extent, they are correct. Some theories with a wide societal or historic reach are indeed testable, but these are usually expressed in clear terms, which define their scope and limits. Indeed, the migratory elites theory, mentioned above, is one such example.

Then, at the opposite end of the scale comes what we might term *experiential theory*, because it is suggested by our direct experiences of the world. Such 'theories' are often hunches based on what we see, hear or people tell us. For example, one of us believes that the readers of certain right-wing UK newspapers are more selfish in their everyday manners, and this belief arises from observations of such people's behaviour on train journeys. Now, this may or may not be true, but whether it is, or not, there is not a scrap of prior theory that has informed this hunch, and in order to turn it into good scientific research, one would first need to find evidence of such prior theory and possibly actual empirical studies. Finally, having done this, it would need to be tested through the derivation of empirical measures. But a great deal of research, rightly or wrongly, does

begin from such hunches, or politicians' ideas or fads. The former British prime minister David Cameron set at least two such hares running in 2010/2011, with his vague concept of the 'Big Society' and his wish to measure happiness![i]

Most theories that social researchers develop or test are 'theories of the middle range'. This concept was first suggested by the US sociologist Robert Merton (1968) and has since been developed as 'middle-range realism' by the British sociologist Ray Pawson (2000).

Middle-range theory consolidates propositional statements and observed regularities and shows how these might be tested empirically. As in the natural sciences, the 'scope' of the theory is stated – that is, what would count as a confirmation that the theory was right and what would count as a falsification, to show it to be wrong. For example, recent work in nursing research (Elo et al., 2013) is concerned with well-being from the viewpoint of people themselves – in this case an elderly population in Scandinavia. Though the concept of well-being is commonplace, in this case it was not defined in advance: the definition was based on the experiences of the study subjects. It began with the hypothesis that the environment was considered as a source of well-being, with the elderly seen as fulfilling their needs and the environment as a resource that contributes to well-being. The research had four stages: (1) the creations of concepts were described inductively through concept synthesis, (2) relationships between the concepts were examined to set up a hypothetical model, (3) hypotheses were set up to verify the concepts and to test possible models and (4) the verification and presentation of the theory (Elo et al., 2013). Note that within this process two things are going on. Firstly, a theory is developed from a concept of well-being that was not predefined but hypothesised as originating in people themselves. This was then empirically tested and a new theory created, which itself was capable of being tested under other circumstances.

Variables

Once we have a clear research question, have reviewed the literature in order to ascertain what previous empirical or theoretical work has been done and have located our question in an appropriate theoretical framework, we must operationalise the research question or questions into research hypotheses and then into **variables** that can be measured. This is sometimes called 'descending the ladder of abstraction'.

Quantitative research is usually 'variable-based'. A *variable* is a feature or aspect of your topic that can take

on different values – typically referred to as a characteristic of your unit of analysis (person, neighbourhood etc.). A variable tends to differ depending on its relationship to one or more other variables, for example, the level of unemployment by region.

Variables are usually classified as independent, sometimes called the predictor variable (the presumed cause in a relationship) or dependent (the assumed effect), sometimes called the outcome variable. So, for example, labour unrest in a region (dependent/outcome) could be related to a decline in employment in that region (independent/predictor), or the voter turnout for particular parties or factions in a region (dependent/outcome) could be related to the degree of voter dissatisfaction in that region (independent/predictor).

But the independent–dependent variable relationship is rarely a direct one, and other kinds of variables usually intervene. A *moderator* variable is one that shapes or influences the relation between an independent and a dependent variable. For example, we may be interested in the relationship between labour unrest (dependent) and a decline in employment (independent), but the relationship may be *moderated* or influenced by which part of the country your respondents reside. In this sense, region would be regarded as a moderator.

Moderator variables are typically distinguished from *mediating* variables. These are variables which 'transmit the effects of another variable'. For example, parents 'transmit' their social status to their children directly, but they may also do this by exercising choice over the kind of education their child receives (Goldthorpe, 2016).

Finally, there are confounding variables, whereby there are two independent variables, whose effects cannot be separated. For example, if a teacher *X* used textbook A in a class and teacher *Y* used textbook B and students were given tests of their learning achievement, the 'independent' variables (the textbooks and the teachers' teaching effectiveness) would be 'confounded', because there would be no way to tell whether any difference observed between the classes, in achievement (the dependent variable), was caused by either or both of the dependent variables.

To return to the labour unrest example, a theory may specify the relationship between the dependent, moderator and independent variables and will constitute the theory to be tested; that is, unemployment causes labour unrest, because it depresses wages, and each element of the theory is represented by a variable.

But it may not be that simple. It is unlikely, for example, that even in industries that have experienced wage decline there is not always labour unrest. Further variables will need to be added and tested in the 'model' (more on models in *The SAGE Quantitative Research Kit*, Volumes 8, 9 and 10). It may be that in some industries labour is not 'organised' in trade unions, or the enterprises may be very small, or employ more casual

labour. We now consider variables as being ascribed to a case. **Cases** are typically individuals who are the subjects of our inquiry.

Cases

On the face of it, a 'case' is quite intuitive, it is usually an individual from whom we can obtain information. This information might be derived from official statistics, or other secondary sources, and comprise things like age, sex, occupation, address, record of severe illness (e.g. cancer) and so on. Alternatively, such information, alongside other attributes, behaviours, attitudes and beliefs, may have been collected via a survey. But all of this information will relate to a 'case' and be operationalisable into variables. But in order to conduct analyses between cases, we will need to have the same information for each case. So, in the simplest of examples, in order to ascertain any difference between males and females on a question of interest, we must have information of each 'case' about whether they are male or female. If we do not have this information, then this is called 'missing data'. Analyses can be conducted when some data are missing, but this requires particular procedures which make assumptions about the mechanism of 'missingness'. We discuss this further in Chapter 7.

Though the case is usually an individual person, it does not have to be. It may be a country, a company, a school or an historical event. Some researchers, nowadays, might draw their cases from more than one level and combine the analyses at each level, say an individual school pupil, a school and a locality. This is called multilevel analysis and is explained in Volume 9.

Whether one chooses individuals, schools, countries and so on as the 'case', in the analysis itself these are then described as the 'unit of analysis'.

Research hypotheses and measures

The term *hypothesis* has been mentioned above a few times. Logically, hypotheses and theories are equivalent, they are each propositional statements, but hypotheses are used in two ways in social research and science more generally. Firstly, and in the sense above, they are statements which make specific testable links between theories and measurement, and secondly, they are statistical propositional statements. A sta-

tistical hypothesis is initially posed negatively and is called the 'null hypothesis' (its opposite is usually called the 'alternative hypothesis'). If something is found to be statistically significant (i.e. the finding is not due to chance), then the researcher *rejects* the null hypothesis and *accepts* the alternative one. Statistical hypotheses may be derived from research hypotheses. Statistical hypotheses and hypothesis testing are described in detail in Volume 3.

Theories help us to generate fairly informal research questions, that will guide the research, but they need to be firmed up into research hypotheses that will propose what should happen, or is forbidden from happening. These hypotheses are tested and may confirm, partially confirm or falsify a theory (Stinchcombe, 1968).

Hypotheses are testing a theory, and to do this they must propose measures. Assuming that you can identify a causal effect, the hypotheses must propose independent variables that can explain that effect. There does not have to be one single independent variable, because as we noted above there may also be variables which moderate or mediate the relationship of an independent and a dependent variable. But what is important is that the research hypotheses clearly state what relations between these variables are expected. Some good practice in formulating research hypotheses might be something like this:

- Are you clear in the language you are using to formulate the hypothesis?
- Are you specifically representing the key elements of the theory in the hypotheses?
- Does the hypothesis include both an independent and a dependent variable? Have these been clearly identified?
- Are there moderating or mediating variables that must be taken into account? Could some variables be confounding?
- Does the hypothesis explain what you expect to happen during your research?
- Can the hypothesis be successfully tested through the methods proposed (e.g. a survey, a secondary analysis or an experiment)?

Here is a brief example from counterurbanisation research. Counterurbanisation is a 'middle-range' theory, which aims to explain why people move from urban (often cities) to rural or 'peri-urban' areas (Champion, 1994, 2001). One of us conducted research on counterurbanisation migration to Cornwall, UK. The original theory proposes that migrants will themselves become economically better off in their destination and that their destination will also benefit the community economically. Neither was the case in Cornwall, so the research aimed to find out why. Here are two of the research hypotheses and their measures. There were more hypotheses, and the one given below is a somewhat simplified account (Buck et al., 1993; Williams &

Champion, 1998):

> *Hypothesis 1*: migrants move to Cornwall primarily for its environmental attraction, rather than economic advantage.
> Operationalised through independent variables, such as attitudes to environment, current economic resources, employment attitudes and intentions. Also, there were proposed moderators, such as age, education and social class.
> *Hypothesis 2:* migrants will have amassed sufficient economic assets to be able to move to Cornwall.
> Operationalised through independent variables, such as housing status and employment prior to migration. Moderating variables included region of origin, household composition and size.

Identifying cases, units of analysis, research hypotheses and variables is not the end of the matter, for there are further rungs on that ladder of abstraction! The variables must now be turned into measures in a survey or experiment. There is much more about this in Volumes 2, 3, 4, 6 and 8. But for the moment, it is worth noting that the measures that will operationalise your variables will have to be derived by you in a new survey (though you may use those that others have previously used), but increasingly researchers are turning to the secondary data sources. In this case, you have to try to find existing measures, that were used to obtain the data for the original survey, that will be appropriate for the variables you want to measure. Sometimes, a 'proxy' can stand in for what you want. For example, the UK censuses do not measure housing quality, but they do measure whether or not a dwelling has central heating. The absence of central heating is a reasonably good proxy for poor-quality housing. However, these censuses do not measure income at all, so to find out about individual socio-economic status a whole lot of variables together, such as social class, occupation, location, education and so on, must be used.

Description and explanation and causes

Description and explanation are the cornerstones of science, and all explanation begins with some kind of description. National censuses, official statistics and other large-scale surveys are descriptive in nature. That is, they take a number of phenomena that are considered useful to measure and measure them. This does not mean that the questions asked do not have a purpose, as we will explain in more detail in the final chapter; all research, however 'neutral' it may appear to be, is for a purpose and is socially located. Moreover, descriptive surveys may then be utilised by others to provide explanations. For example, UK Longitudinal Census data

can be used to provide detailed descriptions of change over time, in things like housing, migration, change in socio-economic status or household change, but the data do not themselves provide explanation, something more is required – a theory which purports to explain such change, but can then be tested using the descriptive data. Some research examples can be found at www.indigo-sandbox.ucl.ac.uk/celsius/research/columns/research-projects.

Before a researcher can move to trying to explain something, she must be sure her data are adequately describing the phenomenon in question. This description has two dimensions. Firstly, the measures themselves must measure those things they purport to measure (more of this in Volumes 4 and 5). Secondly, if one is using sample data, then the sample must adequately describe the population – again this will be discussed in more detail in Volume 2.

So, description precedes explanation, and a study may only be descriptive. An explanation is the answer to a why, or how, question. Why is there less social mobility in country *X* as compared to country *Y*? How do women negotiate male working environments? Often explanations are answers to causal questions, but causal questions – though certainly answerable with quantitative data – are rarely straightforward. We will say more about this, in relation to choosing a research design, in Chapter 4, whilst a particular approach to causal analysis, in quantitative research, is the topic of Volume 10.

There is one last thing to say about explanation and description. Explanations are predictions, indeed logically they are isomorphic (Schlipp, 1991, p. 556) – that is, logically each implies the other. A prediction that labour unrest will follow from high levels of unemployment implies an explanation. But more strongly, a successful explanation that in context C^1 labour unrest did follow from high unemployment, but was moderated by wage reduction, provides a prediction to be tested in context C^2.

Causes

Explanation usually implies causes. In quantitative research, we answer causal questions either through experimental or survey methods. Now, experiments in the social world cannot be the same as those in the laboratory, say for chemists or physicists. Firstly, because the social world is more complex, and secondly, the laboratory is often too much of an artificial environment to produce results that can hold much beyond the laboratory setting, so social researchers must conduct their experiments in the field – that is, in the natural social setting. But never mind, the logic is the same. The population is divided into two samples, the first is the

experimental group, in which there is some kind of 'treatment', say an educational or health intervention with this group, and the second is the 'control' group, with whom there is no intervention. If other environmental effects can be accounted for, then a change in the experimental group is said to be caused by the intervention. But there is much more to it than that, and the logic and practice of experiments is the topic of Volume 6. Indeed, in many circumstances even experiments in the field are not possible and causal explanation must be sought through analysis of survey data.

But be forewarned that there is no concept in these volumes more troublesome than causality. As the philosopher Nancy Cartwright remarked, 'One word, many things' (Cartwright, 2004). There is disagreement amongst both researchers and philosophers about what a cause is. One common formulation goes like this:

> To attribute cause, for *X* to cause *Y*, three conditions are necessary (but not sufficient): (1) *X* must precede *Y*, (2) *X* and *Y* must covary (i.e. associated together) and (3) no rival explanations account as well for the covariance of *X* and *Y*.

Causal relations may be simple or multiple. In simple causation, whenever the first event (the cause – represented by the independent variable) happens, the second effect always follows (represented by the dependent variable). However, simple causation is all but absent in the social world, and multiple causation is all but universal. Multiple causes may be such that any one of several causes can produce the same effect. Homelessness, for example, may be 'caused' by many different antecedent conditions, and homelessness itself causes other things, for example, ill health. Furthermore, the same conditions in one time or place context may cause different things.

An important divide in thinking, in quantitative methods, is the extent to which causal inference can be derived only from multiple probabilistic associations in models or the extent to which causes can be 'interpreted' from models (Byrne, 2002; Freedman, 2011; Williams, 2021). These two approaches can be found in the other volumes in this *SAGE Quantitative Research Kit*. In Volume 6, the emphasis, following Campbell and Stanley (1963), derives from the logic of the experiment. Did the treatments within the experiment make a difference (internal validity), and to what extent is a causal effect generalisable to other populations, contexts or settings (external validity)? In Volume 10, a quite different approach is taken where the simple cause–effect model is superseded by a complexity approach which sees outcomes nested in a vast array of antecedent conditions, in which a simple causal model would under-describe complex and emergent social processes and mechanisms.

Amongst the authors of this volume, one of us (Paul) takes a position, rather closer to the first approach, and the other two of us (Dick and Malcolm) take something of an intermediate position, which sees causes as complex, but also amenable to statistical analysis.

But let us try to simplify for the moment. Causal explanations are grounded in the variable analysis we described above. When developing causal explanations from survey data, one must move beyond the association of two variables, though they often start there. Indeed, it is often said that statistical correlation is not causation (though as Matthew McBee notes in Volume 10, it is always the case that causation implies correlation).

Two variables are often statistically associated – that is, their occurrence together is seen to be more than chance. For example, a society with large numbers of homeless people may be one with high unemployment, but there again it may not. This raises all sorts of interesting questions. Why, for example, is it that there is an association in one place (*A*) between homelessness and unemployment but not in another (*B*) (Bramley & Fitzpatrick, 2018)?

What is the direction of causality in *A*? Does unemployment cause homelessness, or is it the other way around? Or (and in fact this is closer to the truth) the cause operates in both directions over time?

But in place *B*, there is relatively little homelessness, despite high unemployment. Finally, there is somewhere called *C* where there is low unemployment but high levels of homelessness.

Now, let us assume your measurement of homelessness and unemployment is the same in *A*, *B* and *C*; then the explanation (which will be different in each case) will require the addition of new variables in an explanatory model.

Notice, we have used the word *model* here to indicate a statistical model, usually with three or (a lot) more variables. In linear models, causal explanation is derived from the best 'fitting' statistical model. At this stage, that may sound mysterious, but in Volumes 9 and 10 much will be revealed in the matter of causal explanation and the kinds of models we use to achieve it!

One last note of caution when one searches for causes in quantitative data. It is the dog that did not bark! By this we mean that causal attribution must ultimately rest on that which is measured. If it is not measured, it cannot contribute to the 'cause' in the model. And it is often the case that something important was not

measured! This is referred to in the literature as 'omitted variable bias' (Riegg, 2008). There are also models (structural equation models) which aim to take account of variables that cannot be measured directly by defining 'unobserved' or 'latent variables'. These are introduced in Volume 9. We will return to causes in Chapter 4.

Conclusion

In this chapter, we have briefly outlined some of the methodological issues that a researcher must consider at the beginning of the research process. There are commonalities across all research: that you must begin with a research question and that inevitably you are testing some kind of theory, whether it is local, informal or experiential, or a much more formal and established theory. You must descend the 'ladder of abstraction' from the broad brush of your initial question or questions down to specific measures – or the adoption of existing measures in secondary sources.

Finally, all experienced researchers know that any research project is linked to the resources that are available. These may be access to a particular population in order to derive a sample (this is discussed primarily in Volume 4 but also in Volume 5) or to have the time and the person power to collect and analyse the data.

Chapter Summary

- This chapter is an introduction to some of the foundational concepts in quantitative research. It begins with the idea that all research is problem solving, but these problems have to be turned into research questions that can be answered. Research questions are formulated in the context of theory – What is a theory and what kinds of theories are there?
- The chapter then goes on to describe what is meant by a 'variable' and how these relate to 'cases'. Having formulated a research question(s), this needs to be operationalised through identified variables into research hypotheses.
- Finally, the chapter discusses the role of research in description and explanation and also how quantitative research is able to identify causes and the limits to this.

Further Reading

This chapter is about the journey from theory to planning your research. There are many excellent books in this area. Here, we suggest three of the clearest guides to this process.

Frankfort-Nachmias, C., & Nachmias, D. (1996). *Research methods in the social sciences.* Arnold.

Now something of a classic, this book is a general introduction to research methods, but because of this it describes the journey from theory to carrying out research within that broader context.

Litwin, M. (1995). *How to measure survey reliability and validity*. Sage.

This is a short book; nevertheless, it provides a very clear non-technical understanding of reliability and different types of validity and how they may be achieved.

Byrne, D. (2002). *Interpreting quantitative data*. Sage.

This book takes a realist approach to quantitative research and emphasises the importance of causality as complex and takes a sophisticated approach to measurement.

Notes

i Many psychologists believe that happiness can be measured. See, for example, https://worlddatabaseofhappiness.eur.nl/hap_quer/introtext_measures3.pdf.

https://doi.org/10.4135/9781529682809