

Chapter 3

Quantitative Research Principles and Methods for Human-Focused Research in Engineering Design

Mark A. Robinson

Abstract Engineering design is increasingly recognised as a complex socio-technical process where the human and social aspects of the system require alignment with those focusing on technical product development. Social science research methods are therefore essential to conduct effective and holistic research into such processes. Accordingly, this chapter provides a grounding in the principles and methods of quantitative social science research. First, the measurement of variables in a reliable and valid manner is considered. Second, scientific principles and the nature of variable relationships are examined, including main effects, mediation effects, and moderation effects. Third, experimental and correlational research designs for exploring the relationships between variables are discussed. Fourth, an overview of statistical methods for analysing quantitative data is provided. Finally, participant sampling, ethical issues, and specialist methods are considered.

Keywords Quantitative research · Measurement · Research design

3.1 Introduction

Historically, it was generally thought that engineering design was concerned solely with technical work, much of it solitary, grounded in sciences such as physics, mathematics, and chemistry (Pahl and Beitz 1984). While these foundations are critical to product design, there has been a recognition in recent years

M.A. Robinson (✉)

Leeds University Business School, University of Leeds, Leeds, UK
e-mail: m.robinson@lubs.leeds.ac.uk

that the working processes and organisational systems in which such products are developed are quintessential examples of complex socio-technical systems (Baxter and Sommerville 2011). Accordingly, the human and social aspects require alignment with the technical product development processes (Crowder et al. 2012; Davis et al. 2014). Engineering design, as Bucciarelli (1988) noted, is a “social process”, one involving “distributed cognition” (Busby 2001) in team-working environments (Dong 2005). It is also a work domain involving complex problem solving (Goldschmidt and Smolkov 2006), creativity (Howard et al. 2008), and complex cognitive visualisation (Demian and Fruchter 2009). It is therefore an ideal domain in which to study human behaviour and cognition. Through my own research, for instance, I have found engineering design to involve socially interactive work some 40 % of the time (Robinson 2012), in complex team environments (Crowder et al. 2012), where the generation, processing, and transfer of information are key (Robinson 2010), and where a range of technical and non-technical competencies underpin effective performance (Robinson et al. 2005).

In recognition of this change of perception, engineering design research is increasingly focusing on the human aspects of work in this field alongside its traditional focus on product development. Much of this research has been conducted by researchers with engineering backgrounds, such as that exploring expertise and task performance (Ahmed et al. 2003), creativity (Howard et al. 2008), problem-solving activities (Cash et al. 2014), information seeking (Aurisicchio et al. 2010), and the evolution of social knowledge networks (Štorga et al. 2013). Other research in this area has been conducted by researchers with social science backgrounds, such as that exploring job design (Lauche 2005), competencies (Robinson et al. 2005), and the role of trust in innovation (Clegg et al. 2002). Part 2 of this book provides examples of the application of psychology, a discipline central to both social and biological sciences, to engineering design research.

However, despite many such examples of excellent, rigorous research, there remains a general lack of awareness of social science research principles in much of the work in this area. This is not due to any lack of ability—indeed, the quantitative methods used by engineering designers to develop and analyse their products are generally more advanced than social science research methods—rather, it is indicative of the lack of social science training in most formal engineering curricula. Thus, in this chapter, I aim to provide a solid grounding in key research principles and methods from social science for those with engineering design backgrounds conducting human-focused research in this area. To do so, I will draw on a hypothetical research study, gradually increasing the complexity of this example to illustrate key research principles. I will provide indicative supporting references for readers to consult, although these research methods are widely covered throughout the social science literature. Finally, I will also include examples from the engineering design literature of the application of such methods in previous research.

3.2 Measurement

A quantitative research study starts by identifying and defining the variables of interest, including how to measure them in a reliable and valid manner. In this section, we will discuss the systematic steps researchers should take to achieve these objectives.

3.2.1 Identifying, Defining, and Measuring Variables

Let us assume, for example, that we wish to study the effects of communication on team performance in an engineering design company (for related research, see Patrashkova-Volzdoska et al. 2003). As both are complex constructs, we must first decide which specific facets to focus on. For instance, communication may encompass frequency, media, recipients, and sources (Patrashkova-Volzdoska et al. 2003; Robinson 2012), while team performance may encompass time, cost, and quality (Atkinson 1999). Guided by the research literature and the nature of the practical problem we are addressing, we will focus here on the facets communication frequency and speed of team work (i.e. performing work in less time) as our research *variables*. Variables are so-called as they exhibit change, across both the unit of analysis (e.g. people, companies) and time, enabling research inferences to be made (Field 2013), as we discuss in Sect. 3.3.

Having established our specific focus, we must now decide how to measure each variable. To do so, we *operationally define* them by specifying the type of data we will use to represent and measure our variables in this research (Foster and Parker 1995). For *quantitative* research, we will be seeking numerical data, preferably of the type that enables us to determine which of the two measurements of a variable is higher (ordinal data), and also the exact distance between these two measurements (interval data), and also using a measurement scale with a true zero (ratio data) (Field 2013). Either such quantitative data can be collected directly by the researcher specifically for the research, so-called *primary data*, or the researcher can use existing data that have been collected for other purposes, so-called *secondary data* (Cowton 1998).

Within quantitative social science research, questionnaires are a popular and effective method for collecting primary data. These involve participants responding to a number of questions or statements (“items”) about focal variables, using standardised measurement scales, to indicate the level of a variable in a particular context or scenario (Hinkin 1998). For instance, Peeters et al. (2007) used a 55-item questionnaire to measure three types of design behaviour—creation, planning, and cooperation—in multidisciplinary teams, with a 5-point response scale ranging from “highly disagree” (coded 1) to “highly agree” (coded 5). We could use such an approach in our example, by choosing existing questionnaire items from the research literature. If we were unable to find suitable items to measure our variables, we could

develop our own, such as “How many times per week do you e-mail your team leader?” for the variable communication frequency, or “What percentage of your team’s projects are completed on schedule?” for the variable speed of team work.

A further option here would be to use existing secondary data available from the engineering design company to measure our variables. Although such data may not be readily available, they can often be more accurate, as we discuss below, and more efficient to use having already been collected. In our example here, a useful measure of communication frequency may be the number of e-mails that team members send to each other per week, recorded directly from the company’s computer systems, although there may be ethical issues with accessing such data, as we discuss later in Sect. 3.5.2. Indeed, such official e-mail records have previously been used in engineering design research investigating communication content and context (Loftus et al. 2013) and social knowledge networks (Štorga et al. 2013). For speed of team work, a useful measure could be calculated by comparing actual project duration to planned project duration for each team, with relevant dates obtained directly from official company records. Adopting a similar approach, previous research examining the work of electronics design teams used a company’s Gantt chart records to infer whether work was progressing on schedule (Jagodzinski et al. 2000).

3.2.2 Reliability

Having identified potential measures of our variables, communication frequency, and speed of team work, we must now consider their appropriateness and accuracy further before deciding which to use in our example research study. Within social science research, appropriateness and accuracy of measurement are usually jointly considered from the perspective of *reliability* and *validity*. Broadly, reliability refers to whether a measurement method yields consistent results, and we consider it first here because it is a prerequisite of validity (Cook 2009).

The two types of reliability most frequently encountered in social science research are *internal reliability* and *inter-rater reliability*. Internal reliability refers to whether the different components of a measure, where they exist, measure the variable consistently (Gregory 2007). It is most commonly examined in research using questionnaires, where multiple statements or questions are used to measure each variable, such as communication frequency here. To do so, a long-standing and widely used statistical coefficient called *Cronbach’s alpha* (α , Cronbach 1951) is calculated, using standard statistical software (see Sect. 3.4), to ascertain the consistency of participants’ numerical responses to each of the statements or questions measuring the same variable. The α statistic ranges from 0 to 1, with higher values indicating greater internal reliability, and a threshold of $\alpha \geq 0.70$ considered sound (Cortina 1993). For instance, Peeters et al. (2007) calculated the internal reliability of their 5 items measuring the variable “reflecting on the design” to be $\alpha = 0.80$ when first developed.

Inter-rater reliability refers to consistency between multiple participants rating the same variable (Gregory 2007). In our example, if all the members of each team rate the speed of their team's work, then there would have to be agreement or consistency between the ratings of each team member for there to be inter-rater reliability. There are several statistical coefficients that can be calculated using standard statistical software (see Sect. 3.4), of which the *intra-class correlation coefficient* (ICC, Shrout and Fleiss 1979) is one prominent example. Ranging from 0 to 1, a value of $ICC \geq 0.60$ would generally indicate acceptable inter-rater reliability (Shrout 1998), although there are several different versions of this statistic for different purposes (Shrout and Fleiss 1979). For instance, Oman et al. (2013) used this method to assess the inter-rater reliability of judges' ratings of the creativity of engineering design solutions, finding them to have acceptable average reliability of $ICC = 0.80$.

Most reliability measurements in social science are focused on the ratings of participants involved in studies collecting primary data. However, the principles of calculating reliability can still be applied to secondary data acquired from companies. For instance, multiple measurements of the same variable drawn from the same secondary data source, such as the e-mail frequency data or project durations we have considered here, could also be assessed for reliability using either of the above two methods.

3.2.3 Validity

Validity refers to whether the measure used measures what it claims to (Cook 2009). So, to be valid, the measure of communication frequency in our example would need to truly measure communication frequency rather than another variable. In social science, there are three main methods of establishing the validity of a measure, each linked to a specific type of validity: *content validity*, *criterion validity*, and *construct validity* (Cook 2009). There are two further types of validity that relate to research design rather than measurement—*internal validity* and *external validity* (Campbell 1986)—that we will also discuss in Sect. 3.3.

Content validity concerns whether all components of a variable, and those components alone, are measured (Moskal and Leydens 2000). Put simply, the measure should be both comprehensive and pure. So, to be comprehensive, our measure of communication frequency should address all potential communication modes, including face to face, e-mail, telephone, instant messenger, and other written media (Robinson 2012). Meanwhile, to be pure, our measure should not address work tasks irrelevant to communication frequency. The irrelevance of some tasks, such as travelling, will be obvious, but with other tasks, such as report writing, a judgement has to be made about whether this matches the operational definition of communication frequency used in the study. A common approach to establishing content validity is to consult experts in the domain being researched about the completeness and relevance of the measure, as Dooley et al. (2001) did with their

questionnaire measure of design software process maturity and Robinson (2012) did with his measurement categories for engineering design tasks.

Criterion validity refers to whether a measurement of a variable is highly related to the actual level of that variable (Gregory 2007). It is generally measured by a correlation coefficient (see Sect. 3.4), usually Pearson's r , ranging from -1.00 to $+1.00$, with positive values indicating a positive relationship; $r \geq +0.30$ indicates moderate validity and $r \geq +0.50$ high validity (Cohen 1988). It arose in the field of personnel recruitment and so is often conceptualised as the relationship between scores on recruitment tests and subsequent job performance (Cook 2009). Indeed, Shah et al. (2009) used this application in their validation of assessment tests for design skills, finding a correlation of $r = 0.60$ with performance in design contests. However, criterion validity is broader and essentially refers to the relationship between the measurement and an independent objective measurement of the same variable (Moskal and Leydens 2000). So, in our example, criterion validity could be calculated for primary data measures, such as the questionnaire items measuring communication frequency, with reference to equivalent secondary data from the company involved, such as the company's e-mail and telephone records indicating frequency.

Construct validity is most commonly determined by whether the measure is highly related to other measures of the same variable (Gregory 2007). Defined as such, it can overlap with criterion validity somewhat; however, the construct validity of a primary data measure is usually measured with reference to another primary data measure, rather than the objective secondary data that criterion validity is concerned with.

3.3 Research Design

Having identified our research variables and established how to measure them in a reliable and valid manner, we can now examine the relationships between these variables. To do so, we must draw on scientific research principles to collect data systematically using experimental or correlational research designs, as we will discuss in this section.

3.3.1 *Scientific Principles*

Debates continue about whether social science is truly a science (Winch 1990), and the lack of consensus can be partially attributed to the methodological diversity of its component disciplines. However, social science with a strong quantitative focus—such as most psychology research—is guided by the scientific method and can therefore lay the strongest claims to being a science (Dienes 2008). A key tenet of science is the principle of difference, which states that if two situations

are identical except for one difference, and the outcomes of the two situations are different, then the initial difference is the cause of the different outcomes (Hole 2012). This consequential relationship between inputs and outcomes is referred to as *cause and effect*, or *causality* (Field 2013). Researchers can have further confidence in this causality if the cause occurs before the effect, known as *temporal precedence* (Brewer and Crano 2014)—although asymptomatic causes can sometimes obscure this—and the effect either does not occur or is weakened by the absence of the cause (Hole 2012).

In quantitative social science research, a cause is referred to as the *independent variable* or *predictor*, and an effect is referred to as the *dependent variable* or *outcome* (Field 2013). Essentially, then, quantitative social science research examines whether changes in one or more independent variables—such as communication frequency in our example—cause changes in one or more dependent variables—such as speed of team work. Thus, such research is concerned with examining the relationship between two or more variables, and such relationships are often represented using *path diagrams* (Baron and Kenny 1986), such as those shown in Fig. 3.1. Here, variables are represented by boxes and the relationships between variables by connecting arrows.

A prerequisite for the scientific examination of relationships between variables is the reliable and valid measurement of those variables (Cook 2009), as we discussed in Sect. 3.2. Another key tenet of science is that such variable relationships are predicted before the research is conducted, or a priori, in the form of falsifiable statements known as *hypotheses* (Foster and Parker 1995). Hypotheses should be clear and testable, and specify the direction of the relationship, for example “communication frequency is positively related to speed of team work”.

When exploring a new research topic, quantitative social science research progresses systematically, building on previous research findings to increase the complexity of the variable relationships it examines (Petty 1997), as shown in Fig. 3.1. Some researchers have referred to this as establishing the *what*, *how*, and *when* of a research topic (Baron and Kenny 1986), and we shall use this framework here. The simplest relationship is between a single independent variable and a single dependent variable, or establishing *what* the *main effect* is (Baron and Kenny 1986). Here, in Fig. 3.1a, we have indicated a positive relationship between communication frequency and speed of team work: as the former increases, so too does the latter and vice versa for decreases. This could also be illustrated graphically, as shown in Fig. 3.2a.

This is an important finding in its own right and a useful starting point. However, in many cases, we may wish to know more detail about this main effect. So, next, we could explore the mechanism through which this effect occurs, or the *how*. It could be the case, for instance, that communication frequency causes speed of team work *indirectly*, by first causing a better common understanding between team members, or what psychologists call shared mental models (Mathieu et al. 2000), which then, in turn, causes speed of team work, as shown in Fig. 3.1b. Such an indirect effect is called *mediation*, and the intervening variable—shared mental models, here—is called a *mediator variable*

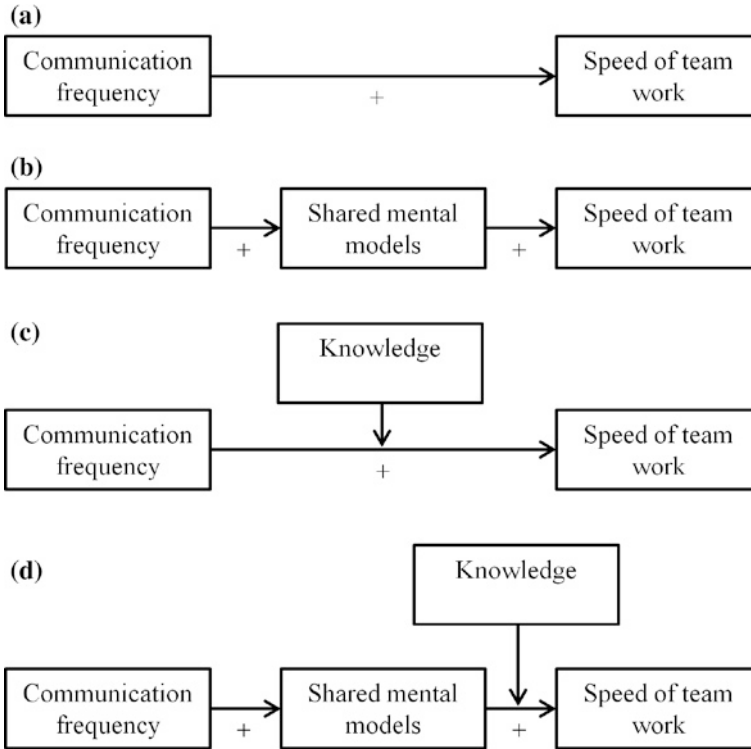


Fig. 3.1 Building a theoretical model by establishing the main effect, mediation effect, and moderation effect (Baron and Kenny 1986) of a research topic. **a** Main effect/“What?”. **b** Mediation effect/“How?”. **c** Moderation effect/“When?”. **d** Theoretical model/“What?”, “How?”, and “When?”

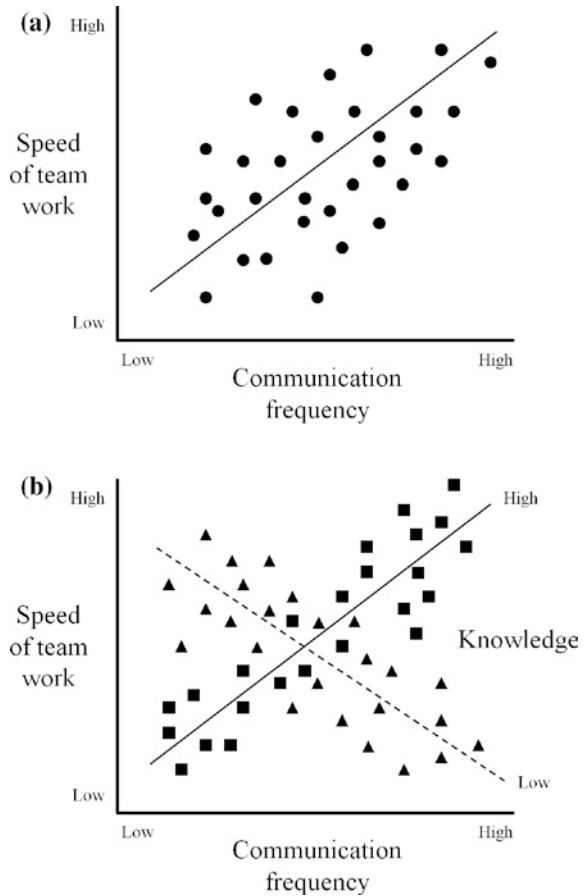
(or mediator) (Baron and Kenny 1986). For instance, Johnson and Filippini (2013) found that the positive relationship between integration activities and performance in new product development was an indirect one, mediated by integration capabilities; thus, activities led to capabilities, which in turn led to performance.

So, now we have further detail about this main effect and how it happens indirectly via a mediator variable. However, we may wish to know even more detail, so now we could explore the conditions under which the effect is present or strongest, or the *when*. Communication frequency is only likely to increase the speed of team work if that communication is useful in some way, so perhaps this effect only occurs when the knowledge level of those communicating is high (Cross and Sproull 2004), as shown in Fig. 3.1c. If we found this to be the case in our research, then there would be an *interaction* or *moderation effect* occurring, and knowledge would be called a *moderator variable* (or moderator) (Baron and Kenny 1986). For instance, Robinson et al. (2005) found an interaction between

engineering designers' ratings of the importance of creativity and innovation to their present and future job roles; in this instance, time (i.e. present or future job) was the moderator variable.

A graphical representation can help clarify the nature of a moderation effect, and Fig. 3.2b provides one such example. Here, there is a positive relationship between communication frequency and speed of team work when knowledge is high (i.e. the solid line and square data points), but the relationship actually becomes negative when knowledge is low (i.e. the dotted line and triangular data points), indicating that non-knowledgeable communication is actually counterproductive. This is an extreme example, with the lines for the different levels of the moderator variable, knowledge, facing in opposite directions to form a cross. In reality, most moderation effects are less dramatic and they are identifiable from converging lines with slightly different gradients.

Fig. 3.2 Graphical representations of a main effect and a moderation effect. **a** Main effect/"What?". **b** Moderation effect/"When?"



In summary then, to understand *what* is happening we must first establish that one variable affects another variable (Fig. 3.1a: a main effect). Then, to understand this main effect in more detail, we can examine *how* it occurs (Fig. 3.1b: a mediation effect), or *when* it occurs (Fig. 3.1c: a moderation effect). These last two questions can be addressed in either order, and their results combined (Fig. 3.1d). By following this systematic research approach, and extending it, it is possible to develop highly complex and nuanced models of causal effects to test, and this is how academic theories are developed in social science (Petty 1997). Part 4 of this book addresses theory and model development specifically in an engineering design context.

3.3.2 Experimental Research Designs

Once we have operationally defined our variables, selected reliable and valid measures, and decided which variable relationships we are examining, we can now design our research study. The purest implementation of the scientific method is the *experiment*. Here, the researcher has full *control* over the independent variables and is able to actively *manipulate* their levels systematically, using different *experimental conditions*, to accurately examine their effect on the dependent variables (Foster and Parker 1995). Often, the dependent variables are measured before and after the administration of the independent variable, known as *pre-measures* and *post-measures*, to gauge the change caused by the independent variable (Liu et al. 2009). Researchers can also include a *control condition* where the independent variable is not administered, and/or a *placebo condition* where the independent variable is administered in the same structure but with inert content (Williams et al. 2002). Structurally, these experimental methods are identical to those used in clinical pharmaceutical trials (Reginster et al. 2001), but applied to human behaviour, cognition, and organisational processes, rather than health.

Figure 3.3 shows the hypothetical results of two experimental research designs. The first, Fig. 3.3a, shows the results of an experiment with three conditions with pre-measures and post-measures of the dependent variable. Here, the control condition shows no change, while the two experimental conditions demonstrate the positive effects of communication frequency, the independent variable, on speed of team work, the dependent variable, with the latter increasing in each case. The second, Fig. 3.3b, shows the results of a quasi-field experiment (see below), in a company for instance. Here, communication frequency has been operationally defined more narrowly as the presence or absence of weekly meetings, as it would be impossible to control all other communication outside of the laboratory. Furthermore, the company wishes to implement weekly meetings throughout the company, so there is no true control condition here. However, to address this, the implementation of weekly meetings could be conducted in two phases (e.g. with a one-month gap between different departments) to effectively create a control condition as shown. Again, the positive effect of weekly meetings, the independent

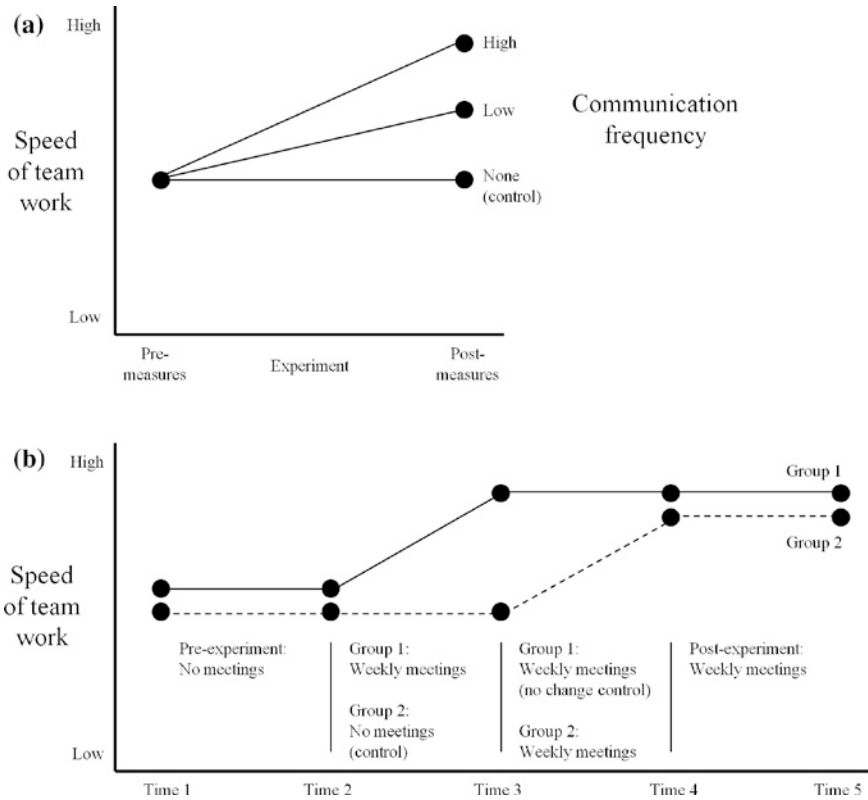


Fig. 3.3 Hypothetical results of two experimental research designs. **a** Hypothetical results of an experiment with pre-measures and post-measures. **b** Hypothetical results of a phased quasi-field experiment to create control groups

variable, on speed of team work, the dependent variable, is demonstrated by the increases in the latter following their implementation.

Participation in experiments occurs in one of two ways. First, different *groups* of participants can be randomly allocated to different conditions, a design known as *between-participants* or *independent measures* (Field 2013). Here, the random allocation of participants helps to randomly distribute their personal differences (e.g. gender, age) between groups, somewhat controlling for them. Second, all participants can be allocated to each of the experimental conditions in turn, a design known as *within-participants* or *repeated measures* (Field 2013). Although this places greater demands on participants, it offers the benefit of ensuring there are no personal differences between participants in different conditions, as they are the same people. However, *order effects*, such as practice or fatigue, must be controlled for by *counterbalancing* the conditions so that equal numbers of participants undertake the conditions in different orders (Reese 1997).

Finally, the researcher also has full control over the experimental environment—very often a laboratory—and so is able to strictly control (i.e. eliminate or reduce) the effects of any other variables unrelated to those the experiment is designed to examine. Some of these *extraneous variables* are randomly distributed and merely reduce the sensitivity of the experiment to detect effects, but others vary systematically with the dependent variable—so-called *confounding variables*—and can substantially bias the experiment unless controlled (Foster and Parker 1995). In experimental research, it is best to control such variables methodologically, by designing them out. Where this is not possible, as in much applied research including correlational designs (see Sect. 3.3.3), such variables can be statistically controlled for (Field 2013).

Having full control over all variables in this way ensures that the relationships between independent variables and dependent variables can be isolated. This gives us confidence that any changes observed in the dependent variables are due solely to changes in the independent variables, which would indicate high *internal validity* (Campbell 1986). Granting the researcher full control of the experiment in these ways is the method's greatest strength. However, this control comes at a price as it also necessitates experiments being conducted in artificial controllable environments, rather than realistic applied settings, making the experiment low in *external validity* or generalisability (Campbell 1986).

We could apply such an experimental approach to our example study. Participants could undertake a standard engineering design task in small teams of four, with time to completion converted to speed (i.e. task per time) as a measure of the dependent variable, speed of team work. For simplicity, we will create two levels of our independent variable, communication frequency, represented by two conditions. In the first condition, high communication frequency, participants are permitted to exchange ten written notes, of ten words or fewer, with the other three team members. In the second condition, low communication frequency, participants are only permitted to exchange two such written notes. If we adopted a within-participants design, we would need two equivalent engineering design tasks of equal difficulty, to ensure that participants encountered a new task each time, presented in a counterbalanced order. We could then run this experiment to see which condition resulted in the fastest speed of team work.

Having established this main effect, we could then introduce the moderator variable, knowledge, into a follow-up experiment. Here, we could manipulate the level of knowledge available in each condition by providing different levels of information. For the high-knowledge condition, we could provide the group with ten recommendations about the engineering design task, and for the low-knowledge condition, we could provide just two recommendations. We could then systematically integrate the independent variable and moderator variable conditions to yield the following four experimental conditions: (1) low communication frequency, low knowledge; (2) low communication frequency, high knowledge; (3) high communication frequency, low knowledge; and (4) high communication frequency, high knowledge. We could then run this second, more complex, experiment to see which conditions resulted in the fastest speed of team work and whether a moderation effect or interaction exists.

Cash et al. (2012) undertook a similar experiment to examine the effect of design information—the independent variable—on the number, originality, and effectiveness of design ideas—the dependent variables. The experiment used a between-participants design with five teams of three participants, each undertaking a standard two-hour design task to develop a new environmentally friendly refrigerator. Each team received a different type of information, representing the five experimental conditions, ranging from no information at all in the control condition through to data pages and videos in the condition with most information. Given the between-participants design, the researchers also sought to control for team role personality types to ensure an equivalent composition for each team. The results indicated that the provision of information was generally positively related to performance in terms of design ideas.

So far, we have discussed pure experiments in artificial environments. However, in many cases, researchers may wish to examine such issues in a more realistic applied setting, such as a company. Sometimes, it is still possible for researchers to retain full control of the independent variables, although it will not be possible to fully eliminate extraneous variables (e.g. background office distractions), so the sensitivity of the experiment to detect effects will be reduced. Such experiments are known as *field experiments* (Dvir et al. 2002) and what they gain in external validity, they lose in internal validity (Campbell 1986). In some such instances, though, it will not be possible to randomly allocate participants to experimental conditions, as the company will have their own strategy for administering the independent variable for business reasons. Experiments without such random allocation are referred to as *quasi-experiments* (Grant and Wall 2009). For instance, Davis (2011) used a quasi-experiment to examine the effects of a change in physical office layouts on communication in an engineering company. However, the company involved was implementing the office changes one department at a time, so it was not possible to randomly allocate participants to conditions. As most field experiments and quasi-experiments are conducted in applied real-world settings, they tend to be longer in duration than laboratory-based experiments, often lasting weeks or months rather than hours.

3.3.3 Correlational Research Designs

In experimental research designs, the researcher actively manipulates the independent variables to examine their effect on the dependent variables (Foster and Parker 1995). However, outside of a controlled laboratory environment, it may not be possible or even desirable to do so. So, in our example, it would essentially be impossible to manipulate the frequency with which engineering designers communicate with each other in a real-world company environment. Furthermore, to increase external validity (Campbell 1986), it would actually be desirable to study realistic levels of communication frequency. So, in such circumstances, as with much applied social science, the research will examine *naturally occurring* levels

of independent variables and dependent variables (Tokunaga 2015). Such research is referred to as *correlational research*, to distinguish it from experimental research (Mitchell 1985). Strictly, it is inaccurate to refer to independent variables and dependent variables in correlational research, as no experimental manipulation occurs, so the alternative terms predictor variables (or predictors) and outcome variables (or outcomes) are generally used, respectively (Field 2013). However, these terms are still often used interchangeably, such as in SPSS statistical analysis software (see Sect. 3.4).

As predictor and outcome variables are naturally occurring, and the former are not manipulated in controlled conditions, correlational research has lower internal validity, so the causality of variable relationships is less clear (Campbell 1986). For instance, it may be unclear whether A causes B, B causes A, or both have another cause. Indeed, variants of the phrase “correlation is not causation” are frequently found in the methodological literature (Bleske-Rechek et al. 2015). Nevertheless, well-conducted correlational research does incorporate several key features of experimental research to improve causal inferences, albeit with a lower level of confidence than experimental research. First, researchers still control for extraneous variables (Foster and Parker 1995) where possible, but typically do so statistically rather than methodologically as is done in experiments (Carlson and Wu 2012). Second, correlational research should also be guided in advance by a sound theoretical rationale drawn from the existing research literature and then designed to test hypotheses (Foster and Parker 1995). Third, predictors should be measured earlier in time than outcomes, so that there is temporal precedence (Brewer and Crano 2014). This feature, or its absence, gives rise to two distinct types of correlational research: (1) *longitudinal research*, where predictors are measured earlier than outcomes, and (2) *cross-sectional research*, where predictors and outcomes are measured at the same time (Rindfleisch et al. 2008). Although methodologically superior, longitudinal research is more difficult to conduct due to the practical difficulties of collecting data from the same people repeatedly (e.g. participants may leave the company after the first round of data collection). For this reason, much social science research is of a cross-sectional nature. Fourth, whenever possible, measurements of predictors and outcomes should be collected using different methods to ensure common method bias does not artificially inflate the relationship between them (Podsakoff et al. 2003). This applies equally to experimental research, although it is unusual not to use different measures in these contexts as the experimental tasks usually necessitate it.

So, returning to our example, we will now consider how we could undertake a correlational study. Our predictor variable communication frequency could be measured with a questionnaire, using either existing items, or our own such as “How many times per week do you e-mail your team leader?”, as discussed earlier. We could measure our outcome variable speed of team work with reference to official company records about actual project durations and planned project durations. By acquiring predictor and outcome measures from different sources in this way, we could guard against common method bias (Podsakoff et al. 2003). Measuring both variables simultaneously would yield a cross-sectional study, but it would be

advantageous to measure communication frequency several months earlier than speed of team work to yield temporal precedence with two *time points* and greater confidence in causality (Brewer and Crano 2014). The questionnaire could be extended to measure shared mental models and knowledge, our respective mediator and moderator variables (Baron and Kenny 1986). For the mediation effect, it would be advantageous to introduce a third time point, between the measurement of predictor and outcome variables, so that there is temporal precedence (Brewer and Crano 2014) for both sequential relationships comprising the mediation effect (see Fig. 3.1b).

One published example of such a longitudinal study was undertaken by Kazanjian and Rao (1999) to examine the development of engineering capability in recently established high-technology firms. First, using a questionnaire, they measured the predictor variables CEO's background, presence of a head of engineering, management team size, and the formality and centrality of decision making. Then, using a second questionnaire 18 months later, they measured the outcome variable engineering capability. Statistical analyses indicated that the presence of a head of engineering and management team size were both significant predictors of subsequent engineering capability, with the former a positive predictor and the latter negative.

Table 3.1 provides a summary of the discussions in Sect. 3.3 concerning the features, advantages, and disadvantages of experimental and correlational research designs.

3.4 Statistical Data Analysis

The statistical analysis of quantitative social science data is a highly specialised field in its own right with accompanying computer software such as Statistical Package for the Social Sciences (SPSS). It is therefore beyond the remit of this chapter to provide detailed guidance in this area; however, we will briefly examine some of the key principles and methods and provide examples of their use in the engineering design literature. Readers seeking detailed guidance should consult some of the excellent books available about conducting statistical analyses using SPSS software, such as Field (2013) or Gray and Kinnear (2012). All of the statistical analysis techniques discussed below can be quickly calculated using SPSS and similar software.

There are two broad types of statistical analyses—*descriptive statistics* and *inferential statistics*—and we shall address each in turn here. Descriptive statistics, as the name implies, are concerned with describing the data collected about a particular variable in terms of its *central tendency* or *average* value and its *variability* or *range* (Foster and Parker 1995). There are three measures of average, namely the *mode*, which is the most frequently occurring value, the *median*, which is the centrally ranked value, and the *mean*, which is calculated by summing all data values and dividing by the number of data values (Field 2013). To examine the

Table 3.1 Comparison of the features, advantages, and disadvantages of experimental and correlational research designs

Methodological criteria	Experimental research designs		Correlational research designs	
	Experiments	Field experiments	Longitudinal	Cross-sectional
Internal validity (i.e. scientific approach)	Very high	High	Moderate	Low
a. Researcher manipulation of independent variables/predictors	Very high	High	None	None
b. Researcher control over extraneous variables	Very high (mainly methodological)	High (mainly methodological)	Moderate (mainly statistical)	Moderate (mainly statistical)
c. Temporal precedence (i.e. independent variable/predictor measured before dependent variable/outcome)	Yes	Yes	Yes	No
d. Random allocation of participants to conditions	Yes (no for quasi-experiments)	Yes (no for quasi-field experiments)	Not applicable	Not applicable
External validity (i.e. generalisable to the real world)	Low	High	Very high	Moderate
a. Realism of environment	Low (often in laboratories)	High (often in companies)	Very high (often in companies with natural data)	Very high (often in companies with natural data)
b. Representativeness of participants	Moderate (often student samples)	High (often company employees)	High (often company employees)	High (often company employees)
c. Realism of study duration	Low (often hours)	High (often months)	Very high (often months or years)	Very low (single time point only)

variability of these data values, we can calculate either the *range* between the lowest and highest values, or the *standard deviation* which is essentially the absolute mean difference between the mean and each data value (Foster and Parker 1995). The mean and standard deviation are the most frequently used of these statistics and the two are usually presented together as measurements of each variable. In many cases, such descriptive statistics are useful in their own right. For instance, Robinson (2012) found in his electronic work sampling study that engineering designers spent a mean of 24.96 % of their time engaged in socially interactive technical work and that the accompanying standard deviation was 9.77 %.

While descriptive statistics provide measurements of each variable, inferential statistics enable us to examine the relationships between variables, to test hypotheses, and to generalise beyond the immediate research (Foster and Parker 1995). A useful although simplistic way of understanding inferential statistics is that

they help us test *differences* or *associations* between two or more variables (Gray and Kinnear 2012). Returning to our example, let us assume in our earlier experiment that we wish to test the difference between the speed of team work of those teams in the low communication frequency and high communication frequency experimental conditions. One simple option would be to examine the mean speed of team work in each experimental condition to see which was higher. However, when comparing any data values, there are always variations that occur solely by chance, so we use inferential statistics to establish whether any differences are due to the independent variable rather than chance (Foster and Parker 1995).

By using the relevant inferential statistical test, we can compare the mean values of our dependent variable, speed of team work, in the two experimental conditions to obtain the *probability level* or *p-value* of the difference to determine whether it is *statistically significant* and therefore supports the hypothesis (Gray and Kinnear 2012). *P-values* range from 0 to 1, with a value of $p \leq 0.05$ considered the key threshold for supporting the hypothesis, indicating that there is less than a 5 % probability that the difference was due to chance (Foster and Parker 1995). Although very widely used, several social scientists and statisticians have recently cautioned against complete reliance on *p-values* and suggest calculating *effect sizes* also (Wright 2003; Cohen 1988).

There are many inferential statistical tests covering a wide range of research scenarios, including parametric and nonparametric, and univariate and multivariate (Gray and Kinnear 2012). However, given space constraints, we shall only discuss four of the most frequently used statistical tests briefly here, namely the *t test*, *analysis of variance (ANOVA)*, *correlation*, and *regression*. *T tests* examine the difference in mean values between two sets of data, either from the same source (e.g. participants, companies) in different scenarios or from different sources (Field 2013). For instance, Robinson et al. (2005) used within-participants *t tests* to compare participants' ratings of the present and future importance of various competencies for engineering design roles. The *t tests* indicated that some of the competencies, such as commercial awareness and innovation, had statistically significantly higher mean importance ratings for the future than the present.

ANOVAs are similar to *t tests*, in that they also measure differences in mean values between sets of data from the same or different sources; however, they extend this capability to multiple sets of data, including interactions between two independent variables (Gray and Kinnear 2012). A key point to be aware of is that *t tests* and ANOVAs both test for differences in dependent variables caused by different *categories* of independent variable (Field 2013). For instance, Robinson (2012) used a two-way within-participants ANOVA to examine the time engineering designers spent engaged in different categories of work, finding that they spent significantly more time in (a) technical than non-technical work and (b) non-socially interactive work than socially interactive work. However, there was no significant interaction between the time spent in these types of work. Given their analysis of data arising from categorical independent variables, both *t tests* and ANOVAs are frequently used to analyse the results of experimental research designs (Gray and Kinnear 2012), although not exclusively so. The ANOVA approach has also been

extended into a method called analysis of covariance (ANCOVA) which also enables researchers to control statistically for extraneous variables (Field 2013).

Correlation is a statistical method for examining the association or correlation between two variables, to determine whether it is positive or negative (Gray and Kinnear 2012). With positive correlations, both variables change together in the same direction; so, as one increases, so does the other and vice versa for decreases (e.g. the square data points in Fig. 3.2b and the accompanying solid line). With negative correlations, both variables change together in opposite directions; so, as one increases, the other decreases and vice versa (e.g. the triangular data points in Fig. 3.2b and the accompanying dotted line). Pearson's r (see Sect. 3.2.3 also) is by far the most common statistical correlation coefficient, ranging from -1.00 to $+1.00$, with the valence indicating whether the correlation is positive or negative (Field 2013). The closer the absolute correlation coefficient is to 1, in either direction, the stronger the correlation is, with absolute values of $r \geq |0.30|$ considered medium in size and those of $r \geq |0.50|$ considered high (Cohen 1988). Correlations can be calculated between any two variables, although usually they examine the association between a predictor variable and an outcome variable, despite the earlier caveats we discussed about causality in correlational research (see Sect. 3.3.3). For instance, Birdi et al. (2014) found a correlation of $r = 0.42$ between creativity skills and the implementation of ideas in their study of innovation in an engineering design and manufacturing company.

Regression extends correlation to identify a "line of best fit" through the cloud of plotted data points (e.g. Fig. 3.2a), minimising the overall distances or *residuals* between this line and all the data points in the cloud (Field 2013). Regression coefficients are then calculated for each predictor variable, indicating the gradient of the line, together with where it intercepts the y-axis, from which a regression equation can be generated to predict outcome values from particular values of predictor variables (Gray and Kinnear 2012). Regression analysis also allows researchers to determine the percentage of variance in the outcome variable that is explained by the predictor variables, both for single predictors and for multiple predictors combined (Tabachnick and Fidell 2013). This is essentially an indication of the predictive accuracy of the identified regression result. For instance, Ng et al. (2010) used regression analysis in their research examining performance in a company manufacturing semiconductors. They found that 54 % of the variance in the outcome engineering performance was jointly accounted for by the predictors total quality management, concurrent engineering, and knowledge management. Finally, more complex forms of regression also enable researchers to examine mediation and moderation effects (Baron and Kenny 1986; Fig. 3.1) and to control for extraneous variables (Foster and Parker 1995).

3.5 Further Considerations in Quantitative Research

In this section, we address three further topics of importance to quantitative research. As each is a specialist topic in its own right, only a brief overview is provided here together with references for interested readers to consult for further information.

3.5.1 Participant Sampling

A key contributor to the external validity (Campbell 1986) of a research study is the profile of participants selected by the researchers. Participants represent a smaller *sample* of a larger *population* of people that researchers wish to generalise their results to and should therefore be *representative* of the wider population from which they are drawn (Fife-Schaw 2000). Ideally, to achieve this, we would randomly select participants from the wider population, to obtain a true *random sample* that is unbiased and therefore representative (Field 2013). Where the population are distributed among various categories of importance to the research—such as age groups or departments of a company—we can also choose to randomly sample participants from within these categories (or “strata”) by using *stratified random sampling* to ensure accurate proportionality (Foster and Parker 1995). In applied research, however, practical constraints often prevent truly random sampling, in which case simple (i.e. non-random) stratified sampling can help mitigate any resultant biases and lack of representativeness.

Alongside representativeness, *sample size* is a key consideration for ensuring external validity (Campbell 1986) with larger samples generally preferable (Fife-Schaw 2000) for two main reasons. First, larger sample sizes provide more statistical power to detect significant effects (Cohen 1988), and some multivariate statistical methods also require large participant-to-variable ratios (Tabachnick and Fidell 2013). Second, to generalise research results to a population, it is necessary to sample a certain proportion of that population, although this proportion decreases as the population size increases (Bartlett et al. 2001). Many useful sample size calculators are readily available to help researchers calculate the number of participants required in various circumstances (NSS 2015).

3.5.2 Research Ethics

Unlike some technical engineering design research, social science research usually involves *human participants*. Any research with people involves a careful consideration of *ethical issues* to ensure their well-being. Most universities and research institutions have their own formal ethical review procedures that have to be followed to gain clearance for data collection. A number of professional social science organisations—such as the American Psychological Association (APA 2010) and the UK’s Economic and Social Research Council (ESRC 2015)—also have their own *ethical research guidelines* that their members must adhere to. All such guidelines have the following key principles in common. First, participation in the research must be *voluntary*, with *informed consent* and the *right to withdraw* at any time. Second, participants’ mental and physical *well-being* is paramount, and if the study conceals information from participants—as some experiments do for methodological reasons—then they must be fully *debriefed* afterwards. Third,

unless participants agree otherwise, data collected in the research should remain *secure* and *confidential* to the researchers and should only be presented in an *anonymous* manner.

3.5.3 Specialist Quantitative Methods

Finally, there are a number of specialist quantitative research methods based on the principles outlined in this chapter that social scientists are now increasingly using, including longitudinal diary studies (Bolger et al. 2003), the analysis of multilevel, hierarchical, “nested” data (Osborne 2000), social network analysis (Hanneman and Riddle 2005), agent-based simulation (Hughes et al. 2012), and the analysis of “big data” (McAfee et al. 2012). Although coverage of these specialist methods is beyond the remit of this chapter, interested readers should consult these references for further information. Part 3 of this book also addresses social network analysis and agent-based simulation in an engineering design context.

3.6 Conclusion

In this chapter, I have sought to provide engineering design researchers with a grounding in the principles and methods of quantitative social science research. First, we considered how to define variables and measure them in a reliable and valid manner. Second, we considered scientific principles and how to examine the relationships between variables, starting with main effects and progressing to mediation and moderation effects. Third, we discussed experimental and correlational research designs and the trade-off between internal and external validity these entail. Fourth, we considered the statistical methods used to analyse the quantitative data collected. Finally, we considered participant sampling, ethical issues, and specialist quantitative methods. Throughout the chapter, I have illustrated these principles and methods using an example research study together with further examples from the engineering design literature. It is my hope that this chapter will be of use to engineering design researchers, without formal social science training, who wish to undertake research examining the human, social, and organisational aspects of engineering design work.

References

- Ahmed S, Wallace KM, Blessing LT (2003) Understanding the differences between how novice and experienced designers approach design tasks. *Res Eng Des* 14(1):1–11
- APA (2010) American Psychological Association’s ethical principles of psychologists and code of conduct. <http://www.apa.org/ethics/code/>

- Atkinson R (1999) Project management: cost, time and quality, two best guesses and a phenomenon, it's time to accept other success criteria. *Int J Proj Manag* 17(6):337–342
- Auriscchio M, Bracewell R, Wallace K (2010) Understanding how the information requests of aerospace engineering designers influence information-seeking behaviour. *J Eng Design* 21(6):707–730
- Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51(6):1173–1182
- Bartlett JE, Kotrlík JW, Higgins CC (2001) Organizational research: determining appropriate sample size in survey research. *Inf Technol Learn Perform J* 19(1):43–50
- Baxter G, Sommerville I (2011) Socio-technical systems: from design methods to systems engineering. *Interact Comput* 23(1):4–17
- Birdi K, Leach D, Magadley W (2014) The relationship of individual capabilities and environmental support with different facets of designers' innovative behavior. *J Prod Innovat Manag* (in press)
- Bleske-Rechek A, Morrison KM, Heidtke LD (2015) Causal inference from descriptions of experimental and non-experimental research: public understanding of correlation-versus-causation. *J Gen Psychol* 142(1):48–70
- Bolger N, Davis A, Rafaeli E (2003) Diary methods: capturing life as it is lived. *Annu Rev Psychol* 54:579–616
- Brewer MB, Crano WD (2014) Research design and issues of validity. In: Reis HT, Judd CM (eds) *Handbook of research methods in social and personality psychology*, 2nd edn. University of Cambridge Press, Cambridge, pp 11–26
- Bucciarelli LL (1988) An ethnographic perspective on engineering design. *Des Stud* 9(3):159–168
- Busby JS (2001) Error and distributed cognition in design. *Des Stud* 22(3):233–254
- Campbell DT (1986) Re-labelling internal and external validity for applied social scientists. In: Trochim WMK (ed) *Advances in quasi-experimental design and analysis*. Jossey-Bass, San Francisco, pp 67–77
- Carlson KD, Wu J (2012) The illusion of statistical control: control variable practice in management research. *Organ Res Methods* 15(3):413–435
- Cash P, Elias E, Dekoninck E, Culley S (2012) Methodological insights from a rigorous small scale design experiment. *Des Stud* 33(2):208–235
- Cash P, Stanković T, Štorga M (2014) Using visual information analysis to explore complex patterns in the activity of designers. *Des Stud* 35(1):1–28
- Clegg C, Unsworth K, Epitropaki O, Parker G (2002) Implicating trust in the innovation process. *J Occup Organ Psychol* 75(4):409–422
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Erlbaum, Hillsdale
- Cook M (2009) *Personnel selection: adding value through people*, 5th edn. John Wiley & Sons, Chichester
- Cortina JM (1993) What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 78(1):98–104
- Cowton CJ (1998) The use of secondary data in business ethics research. *J Bus Ethics* 17(4):423–434
- Cronbach LJ (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3):297–334
- Cross R, Sproull L (2004) More than an answer: information relationships for actionable knowledge. *Organ Sci* 15(4):446–462
- Crowder RM, Robinson MA, Hughes HPN, Sim YW (2012) The development of an agent-based modeling framework for simulating engineering team work. *IEEE Trans Syst Man Cybern A* 42(6):1425–1439
- Davis MC. (2011) Psychological and organizational issues in the design of buildings and workplaces. Doctoral thesis, University of Leeds

- Davis MC, Challenger R, Jayewardene DN, Clegg CW (2014) Advancing socio-technical systems thinking: a call for bravery. *Appl Ergon* 45(2):171–180
- Demian P, Fruchter R (2009) Effective visualisation of design versions: visual storytelling for design reuse. *Res Eng Des* 19(4):193–204
- Dienes Z (2008) Understanding psychology as a science: an introduction to scientific and statistical inference. Palgrave Macmillan, New York
- Dong A (2005) The latent semantic approach to studying design team communication. *Des Stud* 26(5):445–461
- Dooley K, Subra A, Anderson J (2001) Maturity and its impact on new product development project performance. *Res Eng Des* 13(1):23–29
- Dvir T, Eden D, Avolio BJ, Shamir B (2002) Impact of transformational leadership on follower development and performance: a field experiment. *Acad Manage J* 45(4):735–744
- ESRC (2015) Economic and Social Research Council's framework for research ethics. <http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015/>
- Field AP (2013) Discovering statistics using IBM SPSS statistics, 4th edn. Sage, London
- Fife-Schaw C (2000) Surveys and sampling issues. In: Breakwell GM, Hammond S, Fife-Schaw C (eds) *Research methods in psychology*, 2nd edn. Sage, London, pp 88–104
- Foster JJ, Parker I (1995) Carrying out investigations in psychology. BPS Books, Leicester
- Goldschmidt G, Smolkov M (2006) Variances in the impact of visual stimuli on design problem solving performance. *Des Stud* 27(5):549–569
- Grant AM, Wall TD (2009) The neglected science and art of quasi-experimentation: Why-to, when-to, and how-to advice for organizational researchers. *Organ Res Methods* 12(4):653–686
- Gray CD, Kinnear PR (2012) IBM SPSS Statistics 19 made simple. Psychology Press, Hove
- Gregory RJ (2007) Psychological testing: history, principles, and applications. Pearson Education Inc, London
- Hanneman RA, Riddle M (2005) Introduction to social network methods. University of California, Riverside. <http://faculty.ucr.edu/~hanneman/nettext/>
- Hinkin TR (1998) A brief tutorial on the development of measures for use in survey questionnaires. *Organ Res Methods* 1(1):104–121
- Hole G (2012) Experimental design. In: Breakwell GM, Smith JA, Wright DB (eds) *Research methods in psychology*, 4th edn. Sage, London, pp 39–74
- Howard TJ, Culley SJ, Dekoninck E (2008) Describing the creative design process by the integration of engineering design and cognitive psychology literature. *Des Stud* 29(2):160–180
- Hughes HPN, Clegg CW, Robinson MA, Crowder RM (2012) Agent-based modelling and simulation: The potential contribution to organizational psychology. *J Occup Organ Psychol* 85(3):487–502
- Jagodzinski P, Reid FJM, Culverhouse P, Parsons R, Phillips I (2000) A study of electronics engineering design teams. *Des Stud* 21(4):375–402
- Johnson WH, Filippini R (2013) Integration capabilities as mediator of product development practices–performance. *J Eng Technol Manage* 30(1):95–111
- Kazanjian RK, Rao H (1999) Research note: the creation of capabilities in new ventures—a longitudinal study. *Organ Stud* 20(1):125–142
- Lauche K (2005) Job design for good design practice. *Des Stud* 26(2):191–213
- Liu WC, Wang CJ, Tan OS, Koh C, Ee J (2009) A self-determination approach to understanding students' motivation in project work. *Learn Individ Differ* 19(1):139–145
- Loftus C, McMahon C, Hicks B (2013) An empirical study on improving the understanding of email records by augmenting with information on context. In: Proceedings of the 19th international conference on engineering design (ICED13), Seoul, South Korea, 19–22 Aug 2013, pp 169–178
- Mathieu JE, Heffner TS, Goodwin GF, Salas E, Cannon-Bowers JA (2000) The influence of shared mental models on team process and performance. *J Appl Psychol* 85(2):273–283

- McAfee A, Brynjolfsson E, Davenport TH, Patil DJ, Barton D (2012) Big data: the management revolution. *Harvard Bus Rev* 90(10):61–67
- Mitchell TR (1985) An evaluation of the validity of correlational research conducted in organizations. *Acad Manage Rev* 10(2):192–205
- Moskal BM, Leydens JA (2000) Scoring rubric development: Validity and reliability. *Pract Assess Res Eval* 7(10):71–81
- Ng PK, Goh GGG, Eze UC (2010) The influence of total quality management, concurrent engineering and knowledge management in a semiconductor manufacturing firm. In: Proceedings of the IEEE international conference on industrial engineering and engineering management (IEEM), Macao, China, 7–10 Dec 2010, pp 240–244
- NSS (2015) Australian National Statistical Service's sample size calculator. <http://www.nss.gov.au/nss/home.nsf/pages/Sample+size+calculator>
- Oman SK, Tumer IY, Wood K, Seepersad C (2013) A comparison of creativity and innovation metrics and sample validation through in-class design projects. *Res Eng Des* 24(1):65–92
- Osborne JW (2000) Advantages of hierarchical linear modeling. *Pract Assess Res Eval* 7(1):1–3
- Pahl G, Beitz W (1984) In: Wallace K (ed) *Engineering design*. The Design Council, London
- Patrashkova-Volzdoska RR, McComb S, Green SG, Compton WD (2003) Examining a curvilinear relationship between communication frequency and team performance in cross-functional project teams. *IEEE Trans Eng Manage* 50(3):262–269
- Peeters MAG, van Tuijl HFJM, Reymen IMMJ, Rutte CG (2007) The development of a design behaviour questionnaire for multidisciplinary teams. *Des Stud* 28(6):623–643
- Petty RE (1997) The evolution of theory and research in social psychology: from single to multiple effect and process models of persuasion. In: McGarty C, Haslam SA (eds) *The message of social psychology*. Blackwell Publishers, Oxford, pp 268–290
- Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* 88(5):879–903
- Reese HW (1997) Counterbalancing and other uses of repeated-measures Latin-square designs: analyses and interpretations. *J Exp Child Psychol* 64(1):137–158
- Reginster JY, Deroisy R, Rovati LC, Lee RL, Lejeune E, Bruyere O et al (2001) Long-term effects of glucosamine sulphate on osteoarthritis progression: a randomised, placebo-controlled clinical trial. *Lancet* 357(9252):251–256
- Rindfleisch A, Malter AJ, Ganesan S, Moorman C (2008) Cross-sectional versus longitudinal survey research: concepts, findings, and guidelines. *J Marketing Res* 45(3):261–279
- Robinson MA (2010) An empirical analysis of engineers' information behaviors. *J Am Soc Inf Sci Technol* 61(4):640–658
- Robinson MA (2012) How design engineers spend their time: job content and task satisfaction. *Des Stud* 33(4):391–425
- Robinson MA, Sparrow PR, Clegg C, Birdi K (2005) Design engineering competencies: Future requirements and predicted changes in the forthcoming decade. *Des Stud* 26(2):123–153
- Shah J, Smith S, Woodward J (2009) Development of standardized tests for design skills. In: Proceedings of the 17th international conference on engineering design (ICED09), Palo Alto, USA, 24–27 Aug 2009, pp 269–280
- Shrout PE (1998) Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 7:301–317
- Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bul* 86(2):420–428
- Štorga M, Mostashari A, Stanković T (2013) Visualisation of the organisation knowledge structure evolution. *J Know Manage* 17(5):724–740
- Tabachnick BG, Fidell LS (2013) *Using multivariate statistics*, 6th edn. Pearson, London
- Tokunaga HT (2015) *Fundamental statistics for the social and behavioral sciences*. Sage, London
- Williams AM, Ward P, Knowles JM, Smeeton NJ (2002) Anticipation skill in a real-world task: measurement, training, and transfer in tennis. *J Exp Psychol Appl* 8(4):259–270

- Winch P (1990) *The idea of a social science and its relation to philosophy*, 2nd edn. Routledge, London
- Wright DB (2003) Making friends with your data: improving how statistics are conducted and reported. *Brit J Educ Psychol* 73(1):123–136