Understanding Social Work Research

The Philosophy of Social Research

Contributors: Hugh McLaughlin
Book Title: Understanding Social Work Research
Chapter Title: "The Philosophy of Social Research"
Pub. Date: 2012
Access Date: June 04, 2015
Publishing Company: SAGE Publications Ltd
City: London
Print ISBN: 9780857028723
Online ISBN: 9781473913844
DOI: http://dx.doi.org/10.4135/9781473913844.n3
Print pages: 24-47

http://dx.doi.org/10.4135/9781473913844.n3
[p. 24 ↓]

Chapter 3: The Philosophy of Social Research

Introduction

This chapter explores ontology, epistemology and the major paradigms or ways of making knowledge claims about the social world. The major paradigms explored consist of positivism and interpretivism with an analysis of their associated research methods, strengths and weaknesses. The chapter also introduces the development of pragmatism and participatory research approaches.

Ontology

Ontology is the science or study of being. Snape and Spencer (2003) identify three key questions to examine the nature of existence and what there is to know about the world. In particular:

- Whether or not social reality exists independently of human conceptions and interpretations.
- Whether there is a common, shared, social reality or just multiple context specific realities.
- Whether or not social behaviour is governed by ‘laws’ that can be seen as immutable or generalisable. (Snape and Spencer, 2003: 11)

These points identify key ontological debates, including whether there is a captive social reality and how it should be constructed. In response to these questions there are two major positions, realism and idealism. Realism is based on the perspective that there

is an external reality that exists independently of our views or understanding about it. Macquarrie (1973: 57) observes that:

> If there were no human beings, there might still be galaxies, trees, rocks and so on – and doubtless there were in those long stretches of time before the evolution of *Homo sapiens* or any other human species that may have existed on earth. (Italics in original)

[p. 25 ↓] Idealism, on the other hand asserts that reality is only knowable through the human mind and through socially constructed meanings. What is real is only 'ideas' that are confined to what is in the mind. Crotty (2003) comments on Macquarrie's observation that although it is conceivable to consider the existence of a world that is independent of our consciousness that this is not the same as saying that meaning exists independently of consciousness. Or as Crotty (2003: 10–11) explains it:

> The existence of a world without a mind is conceivable. Meaning without mind is not.

Whilst realism and idealism represent the two extreme positions, there have been a number of attempts to modify these, including critical realism (Bhaskar, 1979) and subtle realism (Hammersley, 1992) as variants of the realist position but influenced by idealism. In these positions it is accepted that there is an external reality but that reality is only knowable through the human mind and socially constructed representations. There are similar representations within idealism such as subtle idealism and relativism. Subtle realism recognises that reality is only knowable socially constructed meanings, but that these meanings are shared allowing the creation of a collective meaning. Relativism similarly claims that reality is only knowable through socially constructed meanings but that there is no single reality, only multiple realities (Crotty, 2003).

Reflexive Questions

Where do you stand in this debate – are you a realist, idealist or one of the variants in between? How do you think that this might impact on how you might view the world and

the different types of knowledge claims that you might make and the different types of methods that you might want to use?

Having looked at ontological versions of 'what is' the structure of reality, we now move on to epistemology or the 'nature of knowledge' (Crotty, 2003). Epistemology refers to how we philosophically ground what we can know about the world, the types of knowledge claims that we can make about the world and how we can then assure the credibility of such claims. In particular we will focus on positivism and interpretivism and then on different types of research methods and highlight pragmatism and participatory approaches.

Positivism

Positivism has a long intellectual history dating back at least to Bacon (1551–1626) and Descartes (1595–1650). Auguste Comte (1798–1857) is often considered the first **[p. 26 ↓]** self-conscious voice proclaiming positivism. He argued that society, or the social world, could be studied using the same logic of enquiry as that employed in the natural sciences. Phenomena in both the natural and the social world were subject to invariant laws. The differences between them occurred because of their respective subject matters, which were little more than irritants to overcome by developing appropriate research techniques and methods. Importantly, such a view suggests a deterministic conception of the human race and society by effectively underplaying those factors regarded as uniquely human: free will, choice, morality, emotions and the like. In this approach the pursuit of knowledge is achieved through the process of deduction that leads to experimentation, verification, explanation and finally to prediction.

Giddens (1977: 28–9) identified four major claims made by positivists:

- Reality consists of what is available to the senses.
- Science is the primary discipline.
- The natural and social sciences share a common unity of method.
- There is a fundamental distinction between fact and value.

The fourth of these factors suggests that facts, being the product of science, are superior, whilst values represent an entirely different and inferior order of phenomena.

Associated with the positivist's view of classification is that of measurement. Accordingly, there have been great efforts to scale all kinds of variables in order to achieve an exactness and precision characteristic of natural sciences. It is possible to see this in many forms of psychology and sociology with their focus on stress, attitudinal scales or criminal behaviour.

Positivists believe that the basis of science lies in the theoretically neutral observation language, which is both ontologically and epistemologically primary. As Delanty and Strydom (2003) note, neither of theses terms can be seen as unequivocal. It is generally accepted, though, that ontology relates to what is the essence of things that make up the world or the theory of the nature of the world. Epistemology, on the other hand, relates to what is the character of our knowledge of the world and what is to count as facts. Both concepts are linked in that claims about what exists in the world almost inevitably lead to questions about how what exists in the world and how it is made known (Delanty and Strydom, 2003).

Statements made in the theoretically neutral observation language are directly verifiable as true or false by looking at the ‘facts’ of the world. This represents a correspondence theory of truth, in which the truth of a statement is confirmed by the correspondence with the facts. If it corresponds with the facts, it is true, if not, it is false. Language and the facts ‘would speak for themselves’. This form of empiricism can best be seen in the popular TV series *CSI* in which Grisholm, a crime scene investigator examines all the evidence, in minute detail, to ascertain the facts of the death and to discover whether it was murder and if so who the **[p. 27 ↓]** murderer was. Grisholm famously encouraged his team to let the ‘facts speak for themselves’.

Similarly, positivist statements would be directly verifiable as true or false by their correspondence with the facts. The beliefs we hold or the values we subscribe to are as factually ‘brute’ as atoms, velocities or simple harmonic motion. If social scientists or social workers would only use carefully constructed apparatus – questionnaires, Likert scales and the like – inner mental states could, in principle, be researched empirically. Standardised lists could be developed; all social phenomena could be

classified, correlated and measured. Hypotheses could be formulated and tested with a view to proving whether they were true or false. As such, the world would become predictable and with predictability comes the potential for control. This would provide the social worker with the opportunity to intervene in social situations knowing that their interventions would result in positive outcomes; children would be safeguarded, people with disabilities would be empowered and the elderly supported to live their lives independently.

Falsification and Popper

Karl Popper (1902–94), a philosopher, was a key figure in the development of positivism. Popper was particularly concerned not with how to verify a theory, which he believed was impossible, but how to refute one. Popper grudgingly admitted that metaphysical ideas may have helped with the development of science but that the primary purpose of empirical science is to draw a line between the empirical and the metaphysical, science and pseudoscience. He developed the idea of deductivism or hypothetico-deductivism. In it he contends that a scientific theory can never be accorded more than provisional acceptance (Popper, 1980). Popper gave the example of white swans to demonstrate his point. If our concept of a swan includes the notion of them being white, then we only need to see one black swan to falsify our previously held theory of what it is to be a swan. Popper was able to show that although deductivism was not able to prove a universal statement, all statements were in principle refutable:

> There can be no ultimate statements in science: there can be no statements that cannot be tested, and therefore none that cannot in principle be refuted, by falsifying some of the conclusions that can be deduced from them. (Popper, 1980: 47, underlining in original)

Popper thus revised the orthodox positivist conception of science; no longer was the object of science to infer from specific instances to generalisations but to search for ways of refuting what he called ‘conjectural hypotheses’. Science thus becomes not a body of accumulated and accumulating true theories but a series **[p. 28 ↓]** of series of conjectures or hypotheses that are yet to be refuted (Hughes, 1990). The best theories

will provide very precise predictions across a range of spheres, allowing for empirical testing and opportunities for refutations. As such this suggests an almost evolutionary survival of the fittest.

Popper's contribution of post-positivism emphasises that theory acceptance must always be tentative can be challenged on both theoretical and practical grounds. On theoretical grounds, if theory acceptance is always tentative, then how can theory rejection be decisive when observation statements are by their very nature theory dependent and fallible? On the practical level, Popper suggests that social scientists should set to disprove their theories, to find the ‘black swan’. How realistic this is, is open to debate. From my own experience of researchers, they often start with the opposite viewpoint and set out to prove what they already believe to be true. May (2001) also makes the point that if our empirical evidence falsifies a theory, is this sufficient reason for rejecting it? We may just have found a deviant case or a new result that is yet to be built into the overall theory.

Reflexive Questions

At this point you should write down what you consider to be the key tenets of positivism and Popper's development of positivism. Can you identify in what ways positivist assumptions are used in your practice?

The Interpretivist Reaction

Like the positivist tradition, the interpretive tradition has its roots in the seventeenth century with Vico (1668–1744), who stressed that you could not study humans and society in the same way that you studied inanimate nature. The former implied subjective understanding and thus required a wholly different method of inquiry to that of the natural sciences. Society – a product of the human mind – was not only intellectually different, but also subjective and emotional, requiring different models of explanation.

Thus developed an alternative evolving framework stressing ‘humanistic’ or ‘interpretive’ approaches. These alternative approaches rejected the view that the scientific method could be applied to the study of social life and instead emphasised the importance of interpretation and understanding as the only legitimate ways of gaining understanding. The term ‘interpretivism’ is being used here as an umbrella for a range of approaches, including ethnomethodology (Garfinkel, 1967) phenomenology (Schutz, 1973), symbolic interactionism (Goffman, 1961), interpretive interactionism (Denzin, 1989), new paradigm inquiry (Reason and Rowan, 1981), and social constructionism (Berger and Luckman, 1979). In this section it is impossible to cover all of the nuances of the different **[p. 29 ↓]** approaches, nor would all of the proponents of the different approaches necessarily agree with this articulation of their position.

The interpretivist tradition does not set out to gather facts or to measure the frequency of occurrences. In fact, one of their major criticisms of the positivist is that they analysed out, or reduced to a set of statistics, those unique features that make social life a distinctive human experience. What exactly is left out is open to debate but may include: choice, moral and political concerns, emotions, values or the self.

Schutz (1978), an important interpretivist, expresses the difference between the natural and social sciences like this:

> The worlds of nature, as explored by the natural scientist, does not ‘mean’ anything to the molecules atoms and electrons therein. The observational field of the social scientist, however, namely a social reality, has a specific meaning and relevance structure for the human beings living, acting and thinking therein. (Schutz, 1978: 31)

Thus one of the key differences between positivism and interpretivism can be seen to be the ability of the subjects of social life to create their own commonsense structures and to be able to interpret their own experiences. In other words, unlike atoms or molecules, social actors can talk about, explain to others or justify their actions. Knowledge is then not something ‘out there’ to be discovered, but something derived and created from the experiences of the social actors.

This highlighting of meaning and the describing of actions leads the researcher to impute motives of one sort or another. The analytic force of motives and reasons lie not so much in their being 'internal' but in their being tantamount to rules for governing behaviour. This behaviour though can often be explained in a variety of ways. Thus it is possible for a social worker with a 'nod' to communicate to a service user that they wish them to continue with what they are saying, to indicate that they agree with what is being said, to indicate that they disagree with what is being said or to indicate to a colleague that they want to leave. The same movement can in various circumstances and with various intentions constitute any of these actions; it could also be argued that as a result of this it constitutes none of them. The observer cannot see directly into the minds of the observed to inspect their motives. Nonetheless if certain particulars of the context are supplied, an interpretation, of greater or lesser value, can be attempted.

Positivists developed their version of social reality by drawing a distinction between identifiable acts, structures and institutions as 'brute' facts, on the one hand, and beliefs, values and attitudes, on the other. These two orders of reality were then correlated in order to derive generalisations or regularities that then become the substance of social life. The 'brute' facts are considered as objective whilst the values, beliefs and attitudes are considered as a subjective reality, an inferior status. The elements of meaning were thus relegated to secondary **[p. 30 ↓]** versions of reality. For the interpretivist reality cannot be identified apart from the language in which it is embedded. Social realities are constructed, reconstructed, negotiated and renegotiated in and through meanings. Meaning is thus not only about grammatical rules, but also about social interaction. Language and the importance of language is critical in this tradition. The reality of the natural or social sciences cannot be known independently of the concepts available in language. Second, meanings are not totally idiosyncratic, otherwise it would be impossible to communicate. This is not to deny that there are differences in the way that black and white people, women and men, the disabled and non-disabled or children and adults all experience the world. Meanings are therefore not finitely specific, but achieve meaning from their background, context and the interpretations of the language speakers and receivers. Finally, disputes about meaning do not necessarily stem from deficiencies or inadequacies of natural language but may represent inherent features of social reality.

Interpretivism and Relativism

The interpretivist cannot help but be continuously engaged with his/her own subject matter. This poses the eternal problem for the social researcher in how they can maintain that their reality or interpretation is more accurate and valid than that of the subjects of the study. This is particularly the case when we remember that both use the layperson's world as their reference point and share the same resources.

Hammersley (1995) identifies relativism as the key stumbling block for inter-pretivists. Whilst accepting that there is no universally agreed definition of relativism, what appears to be central to the notion is a view of knowledge as paradigm dependent. Paradigms are discussed later, but it is useful to think of them here as a worldview containing a set of assumptions that go beyond rational explanation. When someone makes a knowledge claim, we can ask by what criteria she or he judges her or his knowledge claim to be true. Then, when the criteria are presented we can ask on what basis this set of criteria is believed to be valid. When that basis is identified we can then ask why that is believed to be true. This argument becomes increasingly circular, with no definitive end-point. The interpretivist becomes hoisted by their own petard, not only is a relativist position false – when viewed from other positions – but also when viewed within the framework of interpretivist assumptions.

In the end we need to recognise that absolute certainty is not available, and that attempts to produce absolute knowledge by sense data are bound for failure. Such a view is in danger of sinking into total relativism wherein anything can be true in some framework, if not our own. To avoid this position, Hammersley **[p. 31 ↓]** (1995) invokes Pierce's ‘commonsensism’, where assumptions are relied on until subject to genuine doubt. This potentially produces more questions than it answers, as there is often nothing common about commonsense. One group's taken-for-granted assumptions do not necessarily transfer to any other group. In the 1990s, Oliver (1990, 1993) argued that people with disabilities should not take part in research unless it was informed by the social model of disability as their experience of able-bodied researchers' commonsense understandings were experienced as oppressive and discriminatory. We next move onto research methods or the tools that researchers use to help them make sense of their data.

Reflexive Questions

Having read this section on interpretivism, you should write down what you consider to be the key tenets of interpretivism and its critique. Can you identify in what ways interpretivist assumptions are used in your practice?

Having looked at both the interpretivist and positivist approaches, which do you favour, and why?

Research Tools

Having read the preceding sections, you will have noted that both positivism and interpretivism make different assumptions about how the world can be known. Before we look at some of the well-known research methods, it is important to note that new research methods are continually under development and that there are a wide range of both qualitative and quantitative methods that social work researchers could use, including: documentary analysis, discourse analysis, narratives, biographical methods, increasingly complex statistical quantitative methods as well as making use of the opportunities provided by the digital media and computers. These methods are beyond this textbook and you are now invited to consider which of the identified methods are associated with which philosophical approach.

Reflexive Questions

Can you identify which of the following you would associate with a positivist position?

- 1. Participant observation.
- 2. Questionnaires.
- 3. Surveys.
- 4. Interviews.

What reasons did you give for your view?

[p. 32 ↓] Traditionally, research methods are split between quantitative and qualitative approaches. Positivist researchers are often associated with quantitative methods, including random control trials, surveys and questionnaires. It is important first to clarify what we mean by quantitative:

> Quantitative methods (normally using deductive logic) seek regularities in human lives, by separating the social world into empirical components called variables which can be represented numerically as frequencies or rates, whose associations with each other can be explored by statistical techniques, and accessed through researcher-introduced stimuli and systematic measurement. (Payne and Payne, 2004: 180)

'Quantitative methods' is an umbrella term and covers a wide range of methods that have been informed by positivist assumptions. At it's simplest it involves counting how frequently things happen (e.g. the number of GCSEs that looked-after children attain as compared with children in the wider population) and the presentation of these frequencies in tables and graphs. This is then often extended to look at whether two or more factors are associated, related or can even be seen to be causal. In the example given above, Jackson (1987, 1994) and Heath et al., (1994) have both demonstrated that being accommodated is associated with poor educational attainment that is also associated with poor life chances.

Payne and Payne (2004) identify the common features shared by almost all forms of quantitative research.

- The core concern is to describe and account for regularities in social behaviour.
- Patterns of behaviour can be separated into variables, and represented by numbers.
- Explanations are expressed as *associations (usually statistical) between variables*, ideally in a form that enables prediction of outcomes from known regularities.
- Social phenomena is explored through *systematic, repeated and controlled measurements*.

- They are based on the assumption that *social processes exist outside of individual actor's comprehension*, constraining individual actions, and accessible to researchers by virtue of their prior theoretical and empirical knowledge. (Payne and Payne, 2004: 181–2, italics in original)

In seeking to explain quantitative research methods I would like to examine random controlled trials, surveys and questionnaires.

Random Controlled Trials: 'The Gold Standard'

The random controlled trial (RCT) is most closely associated with medical research, but it is possible to undertake RCTs in other knowledge domains. Sheldon (2000) has consistently advocated that social work needs more RCTs.

[p. 33 ↓] The RCT is often cited as the 'gold standard' method of assessing the efficacy of treatment methods (Reynolds, 2000). The central feature of an RCT is the random allocation of potential participants to an experimental or control group. The intention is to eliminate bias by ensuring that all conditions are randomised, including the preferences and expectations of patients and doctors. Randomisation does not necessarily ensure that groups are identical but what it does is ensure that any differences that arise do so as the result of chance. In some of these trials, the treatment group will be further divided with one half of the group receiving the treatment and the other half a placebo. Ideally this is done as a 'double blind', where neither the patient nor the doctor are aware of who is receiving the treatment and who the placebo. As MacDonald and Popay (2010) acknowledge, knowing a respondent's allocation may influence the observations made and judgements inferred. For similar reasons, whoever is evaluating the clinical outcomes should not be aware of who has and who has not been treated.

RCTs then follow up their patients in order to identify the relative outcomes for the control and experimental group with a view to establishing whether the results are clinically important or not. Clinical importance is often used to replace statistical

significance. Clinical importance may refer to the speed of recovery of the patient, whether the patient has suffered side-effects, required re-admission to hospital or even whether the patient survived the treatment or not. Clinical importance though may be identified differently by the anaesthetist: the patient survived the operation; the surgeon: the patient was discharged from hospital; the GP: the patient was still alive and survived six months after the operation; and the patient whose quality of life had improved.

Surveys: 'Would You Mind If I Asked You a Couple of Questions?'

All of us will have experienced being stopped in the street by someone holding a clipboard who then asked us a number of questions about which adverts we watched, drinks we drank or foods we ate. During this process, the researcher would be ticking boxes in response to our answers. While this is market research, the use of surveys is also a central part of social research. Surveys are one of the most frequently employed methods in social research. Governments, academic researchers and campaigning organisations alike, use surveys (May and Williams, 2001).

Ackroyd and Hughes (1981) have characterised surveys under four different headings:

- The factual or social survey aimed at eliciting general facts, rather than opinions or attitudes, about the conditions and the organisation of whole societies.
- Attitude surveying which focuses on the attitudes people hold as a means to seek to explain and potentially predict their behaviour. This represents a move away from factual surveying to public opinions.
- Social psychological surveying which is more explanation and theory oriented and seeks to investigate personality via various types of attitude measurement techniques.
- Explanatory surveys designed to test some theoretical explanation.

The first two of these were designed to achieve practical as opposed to theoretical ends, although it could be claimed that they all seek to provide or contribute some degree of explanation.

Payne and Payne (2004: 219) note that surveys have typically three types of characteristics:

> They *collect data in a standardised way* from a *sample* of respondents, enabling the data to be *codified, normally into a quantitative form*. (Italics in original)

There are an increasing range of surveys and these include face-to-face interviewing, telephone, postal, email, Web-based or texted surveys. As can be seen, some of these survey types require the interviewer to be present and others are self-completion. Where the interviewer is not present it is very important for the questions to be clear and unambiguous. Preferably a survey will be piloted to pre-test the questions for their ease and clarity of operation, to check whether they are addressing the research question(s) and whether they provide a means of differentiating between different respondent groups.

The survey questions are usually contained in a questionnaire, which starts with a set of classificatory questions contained in the 'personal' section of the questionnaire; these may include questions about age, gender, ethnicity, occupation, salary and so on. The researcher must also make a judgement about whether to use open or closed questions. Open questions allow the interviewee greater freedom to respond to the question in a way that suits their interpretation. Closed questions limit the number of possible answers, allowing for easier and cheaper analysis.

Surveys very rarely report a 100 per cent response rate, for example people move home, go on holiday, cannot be traced or die. Generally, 70 per cent is seen as an adequate response rate for face-to-face interviews. In self-completion and postal surveys, 33 per cent is seen as more typical.

Easterby-Smith and Thorpe (1996) describe the benefits of quantitative methods in general and of surveys in particular as:

> They can provide wide coverage of the range of situations; they can be fast and economical; and, particularly when statistics are aggregated from large samples, they may **[p. 35 ↓]** be of considerable relevance to policy decisions. On the debit side, these methods tend to be rather inflexible and artificial; they are not very effective in understanding the processes or the significances that people attach to actions, they are not very helpful in generalising theories. (Easterby-Smith and Thorpe, 1996: 32)

Qualitative Methods

The preceding two research tools methods were primarily associated with quantitative methods. The next three are primarily associated with qualitative. Qualitative methods are linked to the interpretivist perspective of philosophy. Silverman (1993: 170) has stated that qualitative methods are 'especially interested in how ordinary people observe and describe their lives'. As such, Payne and Payne (2004: 175) describe qualitative methods in the following way:

> Qualitative methods produce detailed and non-quantitative accounts of small groups, seeking to interpret the meanings people make of their lives in natural settings, on the assumption that social interactions form an integrated set of relationships best understood by inductive processes.

Qualitative, like quantitative, is an umbrella term. Qualitative methods refer to a set of approaches that share common features. Qualitative methods:

- Focus on seeking out and interpreting the meanings that people ascribe to their own actions.
- Actions are seen as contextualised, holistic and part of a social process.
- Seek to encounter social phenomena as they naturally occur.
- They work with smaller samples looking for depth and detail of meaning with a less general and abstracted level of explanation.

- They use inductive as opposed to deductive logic allowing ideas to emerge as they explore the data. (Payne and Payne, 2004)

Thus the qualitative methods focus on individuals, their interactions, emphasising interpretation and meaning and the ways in which mutual understandings are negotiated. This alternative research direction requires a different set of methods than that of the quantitative researcher.

Qualitative methods have the ability to look at change processes over time, to understand people's meanings, to adjust to new issues and ideas as they emerge. They also contribute to the evolution of new theories and are seen as ‘natural’ rather than as ‘artificial’. In particular we are going to focus on participant observation, ethnography and interviewing.

[p. 36 ↓]

Participant Observation

Participant observation requires the researcher to immerse themselves in the lives of those being studied. Participant observation thus requires the researcher to engage in a number of activities, including looking, listening, enquiring and recording. May (2001) comments that the apparent naturalness of this approach leads those new to social research to assume that they can undertake this approach with ease. On the contrary, it is probably one of the most personally demanding and analytically difficult methods of social research to undertake. Depending on the aims of the study, the researcher may be required to spend a great deal of time in unfamiliar surroundings, building and maintaining relationships with people whom they have little personal affinity, making copious notes on what to others would appear to be mundane happenings, putting themselves at a degree of personal risk and then spending months analysing the data after the fieldwork has been completed.

There are two main types of participant observer, the researcher as ‘complete participant’ and the researcher as ‘participant-as-observer’. In the ‘complete participant’ role, the researcher becomes a fullfledged member of the group under study, the

purpose of the research being concealed from this group. There have been a number of such studies in the past, including one where a social work lecturer caused a major stir by living in a council estate in Glasgow whilst she observed and wrote down her views on the resident's lifestyle. Following publication of her views, some of the residents were very disparaging about her comments and felt let down that she had written such comments about them. In this type of approach, the residents were not aware that they were being observed nor of the value judgements that were being made about them and the way they lived.

The more common approach is of ‘participant-as-observer’, wherein both the researched and the researcher are aware that theirs is a fieldwork relationship. Since the research relationship is not concealed the researcher is able to use other methods, for example interviewing to complement their research. This type of research has been used with teenage gangs, drug users and in residential settings.

Participant observers may work in teams, but more often they work alone. In the process of observing, they witness how their research subjects interact with their social environment, continually interpreting and applying new knowledge. The researcher or ethnographer is the instrument of data collection. The ethnographer enters the research subject's social universe and using a range of techniques that could include: ‘living among people’, interviews or life histories. Researchers accept that they will ‘contaminate’ the situation although in doing ethnography this engagement is used to an advantage. In this process, ethnographers draw upon their own experience and biography to help understand the research process. This use of the researcher's own cultural equipment is used reflexively to make sense of social **[p. 37 ↓]** action in context. Reflexivity thus implies that ‘knowledge is made rather than revealed’ (Taylor and White, 2000: 199). Ethnography thus requires us to consider how power is exercised in the research process and the implications of this for what does and does not constitute knowledge.

Focus Groups

In recent years, there has been an explosion in the use of focus groups for all sorts of activities, including the market testing of new products, voting intentions and within

a wide range of social care activities where service user perceptions are wanted to help inform service provision. Focus groups allow us to collect a large amount of research data for relatively little financial cost or just as importantly the researcher's time (Hardwick and Worsley, 2011). Cronin (2001: 165) defines a focus group as a group interview:

> A focus group consists of a small group of individuals, usually numbering between six and ten people, who met together to express views about a particular topic defined by the researchers. A **facilitator**, or **moderator**, leads the group and guides the discussion between the participants. In general, focus groups last one-and-a-half to two hours and are tape-recorded. (Bold type in original)

Focus groups are thus managed discussions that are organised to explore a specific set of issues that involve some kind of group activity. Focus groups are particularly useful when the researcher wants to explore people's experiences, opinions and concerns. What sets them apart is that focus group participants, unlike individual interviews or questionnaires, engage in discussion with each other, creating an interactive and dynamic process led by a moderator or facilitator.

There are at least four advantages of using focus groups:

- They provide an opportunity to observe and collect a large amount of data and interaction over a short period of time.
- Discussions should provide rich data as participants present and defend their own views whilst challenging the views of others.
- This very process may help participants clarify their own views but also open them up to alternative views that they would not have considered.
- Focus groups encourage theorisation and elaboration.

On the negative side, the data that focus groups generate is limited in its generalisability. This is because focus groups do not generally consist of randomly selected individuals and those who are willing to participate in a focus group may be different from those who are not. There is also an issue of interviewer bias whereby the **[p. 38 ↓]** researcher is required to be both reflective and reflexive in order to minimise the

impact of their role. Due to these difficulties, focus groups are often included as part of a research strategy although they can be used and justified in their own terms.

Research Techniques Can Be Both Quantitative and Qualitative

In this section I would like to look at a major research tool that can be both quantitative and qualitative. The importance of this method is that it is closely related to the key social work skill of interviewing. Traditionally the worlds of the social worker and the researcher have been seen as mutually exclusive and esoteric activities. One sat in an ivory tower theorising about the world whilst the other had to grapple with the wicked problems and messy business of the world. Both, though, may have more in common than is usually acknowledged. I would like to suggest that interviewing is a central technique for both. Kadushin (1972: 8), in a classic text on interviewing, described the essential features of an interview as:

> a specialised form of communication. A communication interchange in the interview involves two people, each of whom possesses a receiving system, a processing system and a transmitting system. The receiving system consists of the five senses, the receptors. Communication in the interview involves primarily the use of two sense receptors – the eyes and the ears. Having received the incoming signal, one processes it; this involves making sense of the received message, giving it meaning. The processing activity consists of recalling stored information, relating related information to the message, thinking about the message, evaluating the message, translating it so the message is coherent with the receiver's frame of reference. As receivers we select certain items from the incoming message, ignore others, and rearrange what we hear into interpretable patterns. We then typically formulate a response.

Interviewing aims at establishing a framework for future evaluation and enquiry in both spheres. Both probe and must be continually aware of the conditions under which information is generated. The difference occurs at a later stage when social

workers use the data from the interview to assess a situation and to decide whether an intervention is required, and if so, what type of intervention. This does not mean that social workers eschew theoretical considerations or give up their analytical powers but that their priorities are different. It should also be remembered that researchers do not necessarily stop at the theoretical, but may wish to see their work established as current practice or result in a change of policy.

At this point it could be suggested that interviews could be seen as a qualitative tool, given their negotiative and social constructivist agenda. As the reader will already be aware, such distinctions are not always as clear-cut as we would first **[p. 39 ↓]** believe. It is possible for an interview to be quantitative, qualitative or both. What decides whether it is one or the other is the nature of the questions being asked. When we discussed surveys, we noted that the questions could be open or closed. Quantitative interviews would use closed questions and qualitative interviews open questions. One would be asking the respondent to identify a pre-designed response whilst the other would use a much more discursive approach, possibly probing answers to elicit deeper and more meaningful understandings.

An interview is neither inherently qualitative nor quantitative, although we normally associate it with qualitative approaches. Interviews also regularly use schedules to help direct the researcher to the question areas in which they want a response. Within the social sciences, semi-structured interviews are very common which include both closed (quantitative) and open (qualitative) questions. The closed questions may be seeking to gain comparative views of a particular policy or implementation strategy and the open questioning trying to construct or understand the reasoning behind that view. The closed questions may focus on outcomes, but without the open questions they suffer the severe limitation that they do not tell us what these outcomes were outcomes of.

In the last two sections we looked at some examples of qualitative and quantitative research methods. At one level it has been possible to associate different methods with different research epistemologies. However, a commitment to qualitative approaches does not necessarily imply innumeracy as qualitative methods often include statements about sample proportions and analysis of field notes may include content counting. It has also been possible to show that research methods can be both quantitative and

qualitative, or even both at the same time. At this point it would be useful to explore the notion of a paradigm and how this impacts on how the world can be known.

Paradigms: As Worldviews

It is often presented that the positivist and interpretive perspectives on the world represent incommensurate paradigms of how we understand the world. Kuhn (1970: 175), a historian, discussed how natural scientists engage in debates about the phenomena that they study and how they move from one major theory to another and in so doing defined paradigms as:

> The entire constellation of beliefs, values, techniques and so on shared by members of a given community.

As such, paradigms are the worldviews or theories, which define legitimate areas for research questions, methods and solutions for a scientific community. While **[p. 40 ↓]** Kuhn's views were developed in response to the natural sciences, they were also viewed as very pertinent to the social sciences. As Gilbert (2001) acknowledges, social research – and we can also include social work research in this – is situated within a 'paradigm', a scientific tradition. Any new research project is linked into what has gone before. The problems that many social work researchers tackle are derived from previous social work research, they will have been discussed in the relevant journals and the methods used will have been honed by previous researchers. Evidence of the linking between new research to previous ideas and concepts is an important function of the literature review and references sprinkled throughout the article or book. These references not only acknowledge previous work, but also borrow their authority and that of the author to legitimate their own work.

Gilbert (2001) suggests that knowledge is constructed and linked to a particular professional knowledge community. This links back to Kuhn (1970), who suggested that during periods of 'normal science' the focus is on: first, clarifying the facts of the within the paradigm; second, comparing the facts with predictions from the paradigm; and third, articulating and fleshing out the paradigm. During this period of normal science, if results do not fit the theory then they are overlooked, suppressed or explained away as

a novelty. With time these novelties come to be seen as more than just another puzzle requiring a solution. These novelties become too great to be ignored and there begins a transition to crisis and abnormal science when the very foundations of the science are challenged.

This creates the conditions for a change in paradigm choice. Kuhn (1970) claims that the choice of a paradigm is not a rational process and can never be settled unequivocally just by logic or experimental evidence. The new paradigm is not a special case of the old one, but is incompatible and incommensurate with what has gone on before.

Paradigms and Incommensurability

A critical concept in relation to paradigms is the notion of incommensurability. The traditional view of incommensurability is that rival positions, or paradigms, are irreconcilable. It is not possible to believe in both positions at the same time, to believe in one is to deny the other. The second aspect of incommensurability relates to the relationship between paradigm and method. As previously noted in this chapter, both positivism and interpretivism help to inform particular social research methods. Research incorporating quantitative methods presupposes certain beliefs about what the world is and how it can be known. Likewise, research with qualitative methods makes different competing and antithetical assumptions. Paley (2000) questions whether this relationship is either logical or **[p. 41 ↓]** preordained. In order to make his point Paley uses maps as a metaphor. An Ordnance Survey map is more detailed and complete than a motorway map. But, for the motorist the fact that a motorway map does not include every bridle path or footpath does not mean that it is incomplete. Indeed, for the motorist it is the fact that the map is not to scale that is a benefit, otherwise it would be impossible to see the roads! This he suggests is a matter of function, and methods like maps are dictated not by inherent philosophical positions but by particular functions. Both maps are tools and are only suitable for particular purposes and tasks. What then should drive a choice of method is not a philosophical standpoint position but a question as to what is the most appropriate tool for the task in hand. This leads us into the issue of mixed methods approaches and pragmatism.

Mixed Methods Approaches

Since writing the first edition of this book the increasing acceptance and development of pragmatism has been a key change within the social work research landscape. This is not to mean that mixed methods were not used before, but when they were they were often viewed as 'less than perfect' research. More recently we have seen the growth of publications in this field and the acceptance of mixed methods and pragmatism as a paradigm in its own right. Cresswell (2003) notes that pragmatism has its roots in the works of John Dewey (1859–1952) and Margaret Mead (1901–1978) and its key features are:

- Pragmatism is not committed to any one system of philosophy.
- Individual members have a freedom of choice to choose which methods, techniques or procedures best meet their needs.
- Pragmatists do not see the world as an absolute entity.
- Truth is what works at the time.
- Pragmatist researchers look to the 'what' and 'how' of research based on its intended consequences. Mixed methods researchers need to establish a purpose for their 'mixing'.
- Pragmatist researchers agree that research always occurs in social, hierarchal, political and other contexts.
- Pragmatists believe that we need to stop asking questions about reality and the laws of nature. (Creswell, 2003: 12)

Certainly the viewpoint of this book is that it is more important to ensure a suitable fit between the research question and the method(s) being adopted than to achieve a form of epistemological and methodological purity. This also opens up the possibility of being able to use different methods in the same study – the motorway map of the survey approach to gain a broad brush understanding of a **[p. 42 ↓]** particular issue, for example the characteristics of children in residential care including their age on admission, reason for admission, length of time in care, number of placements and so on. This could then be followed up with an ethnographic study to examine the meaning of being looked after as experienced by young people, their carers and their parents. As Ritchie (2003: 43) observes:

> When using qualitative and quantitative research in harness, it is important to recognise that each offers a different way of knowing about the world. Although they may well be addressing the same research issue, they will provide a different 'reading' or form of calibration on that issue. As a consequence, it should not be expected that the evidence generated from the two approaches will replicate each other. Instead the purpose of interlocking qualitative and quantitative data is to achieve an extended understanding that neither method alone can offer.

Thus when adopting a mixed methods approach, it is important to recognise that not only does each method have a different way of knowing about the world but that they will also approach the question differently and provide a different answer. As a consequence, it is imperative for researchers who use mixed methods approaches to explain why the data and their 'meanings' are different and to avoid the situation where one approach becomes the dominant one and conflicts between the data become hidden.

Triangulation is potentially important here; triangulation originated from quantitative research and the 'multiple operationalism' of Campbell and Fiske (1959), whereby multiple measures are used to ensure that the variance reflected is that of the trait or treatment and not associated with any other measure. Triangulation has also come to mean 'convergence' between researchers and convergence amongst theories. To achieve this it is generally accepted that the researcher should pick triangulation sources that have different biases and different strengths so that they can complement one another. Triangulation thus involves the use of different methods and sources to check the data's integrity and/or to extend the inferences that can be drawn from the data.

We should not just assume, though, that by using multiple methods, or by using triangulation, that our research will automatically result in sounder conclusions. Shaw (2003: 110) approvingly quotes Trend's classic account of an evaluation of a USA programme evaluating the effectiveness of direct payment of housing allowances to low income families. Trend concluded:

> The complementarity is not always apparent. Simply using different perspectives, with the expectation that they will validate each other, does not tell us what to do if the pieces do not fit. (Trend, 1979: 110)

Using multiple methods and triangulation does not remove the responsibility from the researcher in order to ensure that these methods work together in such **[p. 43 ↓]** a way that they provide additionality and address the research question. Using triangulation does not remove the responsibility for ensuring that generated data is analysed rigorously and methodically, identifying both areas of correspondence and dispute.

MacDonald (1999: 98) argues that if there is to be a rapprochement of quantitative and qualitative approaches then four principles will need to be conserved:

- As far as possible, researchers should make explicit the assumptions and values underpinning the questions they ask and the methods they deploy.
- The methods should both be internally and externally robust and valid.
- The methods employed should neither be oppressive to the researcher or the researched.
- That the research should be oriented towards knowledge that can be used by users themselves.

Whilst the first two principles would appear to be supportive of Paley's position, the second two are of a different order. The last two principles could be seen as particularly important to ethically driven professions such as social work, with a value base and commitment to service user empowerment.

Paley's perspective is not universally accepted and D'Cruz and Jones (2004) argue that you cannot ignore that methods have a relationship to the philosophical positions in which they have their roots. They do argue, though, for a more hardheaded approach to method choice to ensure that powerful groups will engage with the results and not just seek to attack the methodology. In so doing, they note that discussions about the relationship between paradigms and methods has become more relevant as knowledge and social research has challenged the previously unseen positivistic, scientific, white, Western male perspectives. Feminist, postmodern, poststructural and post-colonial

perspectives have all served to reinforce the positioning of the knower as essential in the creation of knowledge and given marginal groups a voice. This leads us onto participatory knowledge claims.

Participatory Knowledge Claims

Both positivism and interpretivist approaches can be criticised for their acceptance of the status quo and by implication their conservative stances. Positivist perspectives impose structural laws and theories that reduce minority or marginalised groups' experience to numbers whilst failing to address issues of social justice. Interpretivist approaches are criticised for not going far enough to support marginalised groups (Cresswell, 2003). In response to these criticisms there **[p. 44 ↓]** has developed an alternative advocacy or participatory approach influenced by the works of Marx, Friere and Habermas. Cresswell (2003: 11) identifies four key tenets of this approach:

- Participatory action is recursive or dialectical and is focused on bringing about change in practice.
- It is focused on helping individuals free themselves from constraints found in the media, in language, in work procedures and in relations of power in educational settings.
- It is emancipatory in that it helps unshackle people from the constraints of irrational and unjust structures that limit self-development and self-determination.
- It is practical and collaborative because it is inquiry completed ‘with’ others rather than ‘on’ or ‘to’ others. In this spirit advocacy/participatory authors engage the participants as active collaborators in their inquiries.

From the characteristics identified above it is easily seen why such approaches are often favoured by social work researchers and are discussed further in Chapters 5 and 8. Research within this approach is not just a process but also contains an inherent agenda to address social injustice, challenge oppression, promote empowerment and challenge inequality at individual group or societal levels. As Glasby and Beresford (2006) assert, research only becomes justified when it is accompanied by change. The participatory approach has a history of being developed from feminist perspectives,

racialised discourses, critical theory, queer theory and disability inquiry that have all tended to favour social models of explanation and sought to challenge accepted and dominant discourses.

This research approach has also suggested that the choice of research methods is not just an issue of the appropriate tool for the task, but also a statement concerning how the researcher positions themselves and how they consider the world can be known. This does not mean that we should adopt an incommensurate or different paradigm position, but demands that we make clear the intellectual, ethical and methodological assumptions with which we are working. D'Cruz and Jones, (2004: 57) express their position well:

> as social workers, we must be aware of the political and ethical processes of knowledge construction. Social work research is just one of these aspects. If we are to achieve social change as social workers, we are immediately positioned both politically and ethically in relation to social issues and social problems. Therefore we cannot escape our personal or professional assumptions or goals.

At this point it is important to understand where the writer locates himself. I am a white, middle-class (at least by occupation, if not by birth) male, married with two children, brought up in Northern Ireland during 'The Troubles', but who has not lived there since finishing his first degree. Politically, I adopt a **[p. 45 ↓]** critical perspective and see the world as an unequal place both within the four nations and between the 'developed' and 'majority' worlds. As a social worker, or social work manager, for most of my career I have struggled with the contradictions and ambiguities of social work practice as I strived to steer a course in which service users could be empowered, or at worst, be no worse off after than before my contact. As I have predominately worked in childcare, this has created difficulties with the sometimes conflicting needs of children and parents. On occasions, this has also included differences of interests between brothers and sisters and mothers and fathers, I believe that social work can make a difference, but is in danger of becoming overly managerial and bureaucratised, losing the human connection, creativity and 'subversive' nature that gives it its critical edge.

Ontologically, I tend to favour the ‘subtle realism’ of Hammersley (1992), whereby I accept that there is a world independent of individual subjective understanding, but that that world is only accessible through the understandings of research respondents, including that of the researcher. Individual research respondent's interpretations are thus critical; different vantage points will yield different explanations and experiences, but these do not necessarily contradict the notion of an external reality. I hope that this will help the reader to begin to identify some of my biases and be able to read this book with the knowledge of where the writer is located and ask themselves how would a writer located in a different standpoint respond to the issues identified in this book.

Reflexive Questions

How would you locate yourself?

How does this impact on how you see the world?

How does this relate to why you want to be a social worker?

How does this impact on the area of social work that you want to follow?

Summary

This chapter has raised a number of important philosophical issues that you may feel the need to re-read in order to begin to understand them fully. Often social workers and researchers find philosophy difficult, preferring to be ‘doers’ rather than ‘thinkers’. But, as I hope, you now appreciate, both concepts are linked and how you think impacts on what you do and vice versa. In particular the chapter has highlighted two of the major paradigms of social research – positivism and interpretivism, identifying some of their key assumptions and claims to knowledge and identifying how both approaches represent different versions of how the world can be understood. Different research approaches were identified and it was shown how they relate to some of the key social research methods and **[p. 46 ↓]** different paradigms. This was followed by a discussion on the nature of paradigms, their incommensurability and whether this

incommensurability also relates to method. In response to this, we noted the growth of pragmatism and mixed methods and their view that it is the nature of the research question that should drive the chosen research method. Alongside these differing perspectives, it was also suggested that the researcher needs to reflect on how their own standpoint will affect what methods are chosen and how they are combined.

Suggested Reading

Cresswell, J.W. (2009) Research Design: Qualitative, Quantitative and Mixed Methods Approaches , 3rd edn. London: Sage.

Particularly useful for its explanation and justifcation of mixed methods.

D'Cruz, H. and Jones, M. (2004) Social Work Research: Ethical and Political Contexts . London: Sage.

This is a well-written easy-to-read book that was also identified in Chapter 1.

Hughes, J. (1990) The Philosophy of Social Research , 2nd edn. London: Longman.

A challenging but engaging introduction to the philosophy of social research.

http://dx.doi.org/10.4135/9781473913844.n3

Qualitative Methods in Business Research

2 Research Philosophy

Contributors: Päivi Eriksson & Anne Kovalainen
Book Title: Qualitative Methods in Business Research
Chapter Title: "2 Research Philosophy"
Pub. Date: 2008
Access Date: June 05, 2015
Publishing Company: SAGE Publications Ltd
City: London
Print ISBN: 9781412903172
Online ISBN: 9780857028044
DOI: http://dx.doi.org/10.4135/9780857028044.d8
Print pages: 11-25

http://dx.doi.org/10.4135/9780857028044.d8

2 Research Philosophy

This chapter will give you information on:

- the relevance of philosophical issues in business research;
- the key concepts and positions of the philosophy of science;
- the logic of how theory and empirical analysis are related;
- the role of reflexivity in qualitative research.

The relevance of philosophical issues in business research

There are several issues that you need to take into consideration when starting a research project. Some of these are more practical by nature, relating to the research design and process, the planning of the practicalities concerning data acquisition, access to the research site, gathering materials and analysing them. Most of these issues we will valorize throughout the book, and the entire planning process that relates in general to research design we will discuss in Chapter 3.

Among the issues that you need to consider in the beginning of your research project are philosophical aspects and questions that lurk behind every research method and methodological approach. We think that, in order to gain a good understanding of what you can do with methods in your research, you should be at least somewhat familiar with the basic philosophical **concepts,** positions and traditions. All research methods are in close connection to research philosophy and to the ways it is possible to bring forward new knowledge through research.

[p. 11 ↓]

Knowledge about basic issues of philosophy promotes better informed business research

Many practically oriented business researchers do not explicitly state the philosophical viewpoints of their research. This often implies that they either find philosophical questions as non-relevant in their research settings, or take their own philosophical position as self-evident and known. The latter case is particularly prevalent in the mainstream quantitative research.

In fact, it often happens in business research seminars that philosophical discussions arise, especially in relation to qualitative business research settings. There can be several reasons for this. Qualitative business research offers an alternative for the mainstream quantitative business research approach, and it also often discusses the taken-for-granted philosophical assumptions of quantitative research.

It is possible to do qualitative business research without much knowledge of the basic concepts in the philosophy of social sciences that concern various ways of doing research. We think, however, that it is helpful for you to have knowledge of the basic philosophical concepts and ideas for research in order to be able to design a solid piece of study that delivers what it promises (Box 2.1). Statements about what constitutes your research phenomena have implications for the ways in which it is possible to gain knowledge of it.

Therefore, at the minimum, the exploration of philosophical concepts assists you in specifying your overall research design and strategy. These will in turn set the directions for your research, how to proceed from your research questions to the conclusions. You will need to make decisions about the type of empirical data that you will collect, how you will analyse it, rules about how to interpret the analyses, and ideas of how to present your conclusions. The exploration of philosophical concepts will also help you in making decisions about the issues that all have effect in your

Box 2.1 The relevance of philosophical issues

Understanding philosophical issues provides you with a better understanding about:

- #why researchers are interested in different topics and research questions;
- #why qualitative research can be performed in many ways;
- #why qualitative data can be collected and analysed in many differing ways;
- #why different methods are being used in analysing the data.

research design: what kinds of questions do you ask in your research, and in what ways can you answer those questions with your research.

Qualitative approaches attach to philosophical questions in different ways

It is particularly relevant for a novice business researcher to learn that qualitative research approaches can be related to the philosophy of social sciences in differing ways. Most research approaches are not related to one specific tradition of the philosophy of science. This means that methods can be used within several philosophical traditions; therefore, you do not necessarily need to be an expert in the philosophy of science to get it right. For example, **case study** (see Chapter 9) and **focus group** research (see Chapter 12) can be adopted and used in research in several philosophical traditions.

On the other hand, some qualitative research approaches do sign up for a specific philosophical thinking through their theoretical ideas and attachments. Good examples of these approaches are **critical research,** which draws to a varying degree on **critical theory** (see Chapter 17), and **feminist research,** which draws on feminist theories (see Chapter 16). When you choose to follow these approaches, it is very advisable to learn more about the philosophical questions of that particular approach.

Reflexivity in qualitative research

With the constant growth of information and research knowledge, it becomes ever more important to assess the 'production process' of that knowledge and the knowledge itself. In practice, this usually means that you reflect on how you produce knowledge as a researcher, what kind of knowledge it is, and how you can relate this new knowledge to other knowledge you might already have. This everyday reflection is a way to think through your research project throughout the entire process.

In scientific thinking, reflection bears the same idea as in everyday life. It means careful reconsideration of knowledge: how it is produced, described, and justified. Reflexivity is especially important when you think of your epistemological assumptions and commitments in research. On what basis knowledge is argued for and claimed is a question that is not solely related to qualitative research, but to all research knowledge. However, it is most often discussed and brought forward in a qualitative research setting, to add to the transparency of knowledge claims and sources of information, and to open up the relationships that exist between the knower, i.e. you are a researcher, and your subject of interests, i.e. the economy, society, organizations, companies, firms and people working there.

Key concepts of the philosophy of science

Ontology, epistemology, methodology, methods and **paradigm** are key concepts in the philosophy of social sciences. For many researchers, ontology, **[p. 13 ↓]** epistemology and methodology together can be related to each other as **framework,** or even one unified view that some researchers call a paradigm (Burrell and Morgan, 1979; Guba and Lincoln, 1994). All these concepts relate to each other, but in various ways, depending on the more general philosophical position of your research. In this chapter we will introduce some key philosophical concepts and positions, and discuss their relations to the qualitative research approaches that we illustrate in this book (Box 2.2).

Ontology

Ontology concerns the ideas about the existence of and relationship between people, society and the world in general. Ontological assumptions embrace all theories and methodological positions. Several of the qualitative approaches that we will introduce in this book are based on the ontological assumption in which reality is understood as *subjective.* This means that it is based upon perceptions and experiences that may be different for each person, and change over time and **context.**

In comparison, in quantitative research it is usual to assume that the social world exists as a distinctive and separate, i.e. objective, reality. The division between **objectivism** and **subjectivism** is one aspect of ontology in philosophy, which refers to the study of conceptions of reality. Instead of subjectivism, the term **constructionism** is often used to describe the social nature of reality.

Objectivism as an ontological starting point assumes that the social world has existence independently of people and their actions and activities. The objectivist view on ontology assumes that social reality has an independent existence outside the knower, i.e. you as a researcher. Constructionism, on the other hand, assumes that

Box 2.2 Some key philosophical concepts

Questions that the key philosophical concepts aim to answer: *Ontology.* What is there in the world?

- #*Epistemology.* What is knowledge and what are the sources and limits of knowledge?
- #*Methodology.* How can knowledge about a given issue or problem be produced?
- #*Methods.* What are the specific ways of data collection and analysis that can be used?
- #*Paradigm.* What are the conceptual and/or methodological models that relate to a scientific discipline during a particular period of time?

[p. 14 ↓]

social actors produce social reality through social interaction. This means that they also can change their views and understandings of social reality through interaction.

In other words, the subjectivist view on ontology (i.e. constructionism) assumes that the reality for a knower, such as you, is an output of social and cognitive processes. Therefore, two realities alike cannot exist. A focal point in the social constructionist view is that reality does not exist outside individuals; 'reality' is always about individuals' and groups' interpretations (Blaikie, 1993: 94).

When considering the ontological perspective of your research, you should think about what you see as fundamental properties in the social world that are worth studying. This may sound rather difficult, because ontological assumptions are usually more or less taken for granted. Let us take an example. If you are interested in studying what managers do and why, you must first decide whether you believe that they act, for instance, on the basis of biologically determined personalities, cognitively adopted attitudes, or socially constructed identities. All three can indicate very different world views in terms of what is considered as essential in existence and being, what should be studied, and how it can be studied.

Epistemology

In addition to ontology, which focuses on the question 'What is there in the world?', it is helpful to understand what *epistemology* in research means. Ontological claims in research are closely related to epistemological claims, and they usually are discussed together. Epistemology is concerned with the questions 'What is knowledge and what are the sources and limits of knowledge?'.

At large, epistemology defines how knowledge can be produced and argued for. Epistemology defines the criteria by which knowledge is possible. In scientific research, epistemology defines and gives structures to what kind of scientific knowledge is available, what are the limits for that knowledge. In addition, epistemology offers us an answer to the question of what constitutes scientific practice and process. Hence,

there is not just one way of defining the answer for these questions, but several different epistemological commitments and directions exist.

Also, in epistemology there is an objectivist and a subjectivist view. According to the objective view in epistemology, it is possible that there exists a world that is external and theory neutral. According to the subjective epistemological view, no access to the external world beyond our own observations and interpretations is possible.

In addition, there are several directions through which epistemology can be defined (Box 2.3). These directions do not conflate to the qualitativequantitative divide, but are based on the ways through which knowledge claims are made. We will take these up as an initial orientation to epistemic questions of research, and for showing you the complexity that lies behind each method. In case you are interested in finding out more about the differences between epistemological directions, you can read specific literature or attend a course in the philosophy of sciences and social sciences.

[p. 15 ↓]

Box 2.3 Foundation for different epistemological directions

The key division between different epistemological directions lies in three basic assumptions.

- Epistemologies differ in terms of whether they can be considered as being ‘foundationalist’ epistemologies or not. Foundationalist epistemologies seek permanent and reliable criteria for knowledge that is produced in research.
- Epistemologies differ in terms of what is the role that they give to the researcher. Is the researcher an autonomous and detached knower, or are they part of the knowledge production process, and if so, to what extent?
- Epistemologies differ in terms of how they establish relationship between idea and object, or concept and observation. The concept can be separate,

> it can be closely related and even corresponding, or it can be contextually differing from observation (Schwandt, 2001: 71).

Understanding the above-mentioned aspects, of what can be known that exist, makes it easier for any researcher to realize that whatever knowledge we produce in research, that knowledge is seldom based on one unified idea of science and research. Instead, different and equally legitimate philosophically embedded views exist regarding how and in what ways we can know the world. This is the basic assumption concerning the methods and knowledge on methodological tools.

Within epistemology, there are several directions that are associated with the main philosophical positions in social science, which we will introduce later in this chapter. First, there is **empiricism,** in which reality is constituted of observable material things. Empiricism is associated with the philosophical position called **positivism.** Second, there is subjectivism, which views reality as being socially constructed. This means that knowledge is available only through social actors. This epistemological view is associated with the position called interpretivism. Third, there is substantialism, which takes reality as material, but acknowledges that people interpret it differently in different times and contexts. This epistemological view is most often associated with the position called **critical realism.**

Methodologies and method

Methodologies are concerned with how we come to know of the world, but they are more practical in nature than epistemologies. Epistemology and methodology are closely related: the former involves the philosophy of how we come to know **[p. 16 ↓]** the world, whereas the latter involves the same from a practical point of view. Methodology refers to organizing principles, which provide the procedure for guiding the research process and research design that you will learn about in Chapter 3.

Sometimes methodology is called the philosophy of methods. The focal point of methodology is to describe how a given issue or problem can be studied. David Silverman (2005: 4) writes that methodologies can be defined broadly and schematically

(e.g. quantitative and qualitative methodologies), or narrowly and precisely (e.g. grounded theory, case study, ethnography).

Methodology is focused on the specific ways (the methods) that we can use in research when trying to understand our world better. Methods are often divided into methods of data collection (e.g. interviews, **observation)** and methods of data analysis (e.g. **thematic analysis, narrative analysis).** In the method chapters of Part II (Chapters 917), you will learn more about these (see also Chapters 7 and 8 in Part I). Although some methods are better suited to some methodologies (e.g. observation with **ethnography,** or in broader sense with qualitative methodology), they are not rigidly bound to each other in a way that certain methodologies would rely on a very restricted body of methods.

Paradigm

Very often in research method books, and even in textbooks, you will find the term ‘paradigm‘. The historian of science Thomas Kuhn gave the term paradigm the meaning that has become common and much used sometimes even misused within the research context. Kuhn referred to the set of practices that define a scientific discipline during a particular period of time. As a natural scientist, Kuhn did not consider the concept of paradigm to be appropriate for describing the development in the social sciences.

Kuhn (1970), in his book *The Structure of Scientific Revolutions,* writes that he developed the concept of paradigm to be able to distinguish the social sciences’ development from the development within the natural sciences. He had observed that the researchers in social science were never in agreement on theories, concepts or methodologies. Therefore, he concluded that there cannot be any paradigms in the social sciences, but that social sciences are in a pre-paradigmatic phase in the development of scientific knowledge. For a paradigm, researchers need to share not only theories, but also a shared basis for **theory** choice (Kuhn, 1977: 322).

Despite this, the concept of paradigm is widely used in social sciences and in business research. In this discussion, paradigm as a term has shifted away from the original remarks made by Kuhn, and can be defined as a world view or a belief system that

guides a researcher in their work (Guba and Lincoln, 1994). Burrell and Morgan's (1979) model of “four sociological paradigms for organizational analysis”, see also Berg (2004) (functional, interpretative, radical humanist, and radical structuralist paradigms) gained wide popularity among business researchers in the 1980s and 1990s.

[p. 17 ↓]

Even if Kuhn did not argue that paradigms would be invariable, they sometimes were mistakenly taken as such in the discussions. Also, the description of competing paradigms of inquiry that was introduced in 1994 by Guba and Lincoln is often referred to. They identify positivism, **postpositivism,** critical theory, and **constructivism** as the major paradigms that frame social science research. They also claim that these paradigms compete for acceptance as the paradigm of choice in qualitative research. More recently, management and business researchers have been more concerned about epistemological discussions than debating on paradigms (Gill and Johnson, 1997; Willmott, 1997; Johnson and Duberley, 2000).

The main philosophical positions

As is the case with many other terms and concepts, the social scientists use the main philosophical concepts in somewhat differing ways. The issues that we discuss here under the title ‘philosophical positions’ can be found in other methodology books under such titles as ‘paradigms‘, ‘epistemologies’ or ‘research philosophies’ and ‘research traditions‘.

Also, the divisions made between philosophical positions vary in literature. Several textbooks label all qualitative research as being interpretative, but some prefer to make distinctions between various philosophical positions that inform qualitative research, including, among other things, **postpositivism,** critical realism, constructionism, and **postmodernism.** This is because some qualitative research is more inclined to follow the natural science model with hypothesis testing, etc. This is due to the differences in the epistemological and, more generally, philosophical positions of research settings. In the following, we will briefly describe the most common philosophical positions that business researchers rely on and describe how they direct research interests (Box 2.4).

Positivism and postpositivism

Management and organization researchers Johnson and Duberley (2000: 38) suggest that positivism, also known as logical positivism, is the mainstream philosophical position of management studies. They propose that while management as science is fragmented and diffuse, positivism has been one programme to unify management research. Another, additional explanation for the dominance of positivism lies in the nature of management and business knowledge. This knowledge is often functional by nature, and there is a desire for universal truth that would hold across industries, businesses, cultures and countries. Often, managerial implications in research are seen as important value added. These aspects can call for positivist approach in research settings.

Positivism, a term coined by Auguste Comte (18981857), refers to an assumption that the only legitimate knowledge can be found from experience.

[p. 18 ↓]

Box 2.4 The main philosophical positions

- #*Positivism:* knowledge of the world is obtained through applying the scientific methods to experiences and to empirical world. *Postpositivism*: a reformed version of positivism, also includes critique towards the basic assumptions of positivism. *Critical realism* : combines some of the ideas in positivist and constructionist thinking; concerned with the identification of the structures of the world.
- #*Interpretivism and constructionism*: background in hermeneutics and phenomenology; concerned with subjective and shared meanings.
- #*Hermeneutics*: refers to the necessary condition of interpretation and understanding as part of the research process.
- #*Postmodernism*: rejects the positivist, rational and generalizable basis for scientific research, which would explain the world from an objective standpoint.

- #*Poststructuralism*: stands for the most extreme rejection of positivism within postmodernism.

According to the basic claim of positivism, research produces facts and accounts that correspond to an independent reality, is value free and prioritizes observation. Positivists believe in empiricism: the idea that observation and measurement are the essence of scientific endeavour. The key approach of the scientific method is the experiment in which the operationalization of issues that are studied is the prevailing idea: only things that are measurable can be dealt with.

There are also several directions within positivism. At first, positivism relied on empiricism as the foundationalist epistemology, but later positivism has moved towards a non-foundationalist epistemology. Despite its several directions, positivisms in plural share some basic ideas of knowledge production, according to which the aim of research should be in finding causal explanations and regularities. Various versions of positivism find methodological unity in natural and social sciences, and sign for value-free science ideals. These elements of positivism can be found in most positivist research as the core orientation, and within qualitative and quantitative business research.

Positivism does have relevance in business research, but it is more closely related to the logic of and ways for doing quantitative research. Also, qualitative research can subscribe to some version of positivism, when hypothesis (or theory) testing are at the forefront in research. Also, older versions of the grounded theory approach sign for some ideas of positivism in research (for more, see **[p. 19 ↓]** Chapter 11). Other approaches have more relevance for qualitative research than positivism.

Postpositivism (also known as postempiricism) developed through the main criticism of positivism. It argues that the knower and known cannot be separated (as positivism claims). It also questions the idea of a shared, single reality. It suggests that, although human beings cannot perfectly understand reality, researchers can approach it with rigorous data collection and analysis. Postpositivist thinking is influential within qualitative research, as it covers philosophical positions and methodologies as different as scientific and critical realism, grounded theory (Glaser and Strauss, 1967), and symbolic interactionism (Blumer, 1969). Furthermore, the detailed ways of analysing

qualitative data devised by Miles and Huberman (1994), for example, are often called postpositivist.

Critical realism

Critical realism agrees with positivists that there is an observable world independent of human consciousness. At the same time, it suggests that knowledge about the world is socially constructed. Critical realism is closely related to the works of Roy Bhaskar and Rom Harré (Danermark et al., 2002). Critical realists use the word ‘critical’ in a particular way; this is discussed more in Chapter 17, where critical research as a research method will be outlined.

In business research, Johnson and Duberley (2000), Reed (2005) and Contu and Willmott (2005) suggest that critical realism provides one alternative for those studying management and organization. Johnson and Duberley (2000) suggest that critical realism allows you to use multi-methodological approaches, which, in turn, enhance more detailed and accurate analyses when looking for causalities in companies’ development, for example. Adoption of critical realism in business research is not, however, an easy or uncontested approach.

Interpretivism and constructionism

There are many forms of interpretivism and constructionism, but common to all of these is a concern with subjective and shared meanings. These philosophical positions are interested in how people, as individuals or as a group, interpret and understand social events and settings. As much of the qualitative research focuses on human action and understanding, interpretation is an important part of any analysis of qualitative materials (e.g. Alvesson and Willmott, 2003). The philosophical base of interpretative and constructionist research is in hermeneutics and in **phenomenology,** which have an influence on the ideas of social construction of reality (Berger and Luckmann, 1967).

Interpretive and constructionist researchers start out with the assumption that access to shared dynamic and changing and individually constructed reality is only through

social constructions such as **language** and shared meanings. This is why **[p. 20 ↓]** interpretative and constructionist research does not only focus on the contents of empirical data, but also on how these contents are produced through language practices. Furthermore, research done from these philosophical positions does not predefine dependent and independent variables, but focuses on the full complexity of human sense making as the situations emerge. It is also assumed that there are many possible interpretations of the same data, all of which are potentially meaningful.

Perhaps the dominant form of current interpretive research is social constructionism, which was introduced by Berger and Luckmann in their book published in 1967. Social constructionism seeks to understand how the seemingly 'objective' features, such as industries, organizations and technologies, are constituted by subjective meanings of individuals and intersubjective processes such as **discourses.**

Vivien Burr (1995) identifies four basic assumptions of the social constructionist philosophical position. First, it takes a critical stance towards taken-for-granted knowledge, trying to open it up for discussion. It is assumed that the world does not present itself objectively to the observer, but is known through human experience, which is mediated by language. Second, the categories in language that are used to classify things around us are produced through social interaction within a group of individuals at a particular time and in a particular place. Third, knowledge is sustained by social processes and conventions of communication. Fourth, knowledge and social action go together.

In this view, reality is socially constructed by interconnected patterns of communication. Therefore, reality is not defined by individual acts, but by complex and organized patterns of ongoing actions. Schwandt (2001: 32) remarks that there are both weak and strong versions of social constructionism that differ in their views regarding the social construction of everything. For qualitative research, constructionist views on knowledge production are useful, as they emphasize the close relationship between researcher and researched field, interaction and understanding as basic tenets of research. **Reflexivity** is one key part of constructionism.

Hermeneutics

Hermeneutics is a term that originates from Friedrich Schleiermacher (17681834). Also, the term ‘interpretivism’ is often used for hermeneutics. Hermeneutics refers to the necessary condition of interpretation and understanding as part of the research process: inescapable action of interpretation (verstehen) taking place in all research. It is ontological by nature, but still, hermeneutics and interpretivism have given resonance to later epistemological developments in asserting that there is a fundamental difference between natural science's and social science's subject matters: human intentions are crucially moulding and changing the reality. Because of this, understanding of human intentions is needed, and it is the understanding of the human actions that is the foundation for all knowledge in social sciences. Much of the qualitative research focuses on human actions and understanding; therefore, **[p. 21 ↓]** interpretation is indeed an important part in any qualitative research (e.g. Alvesson and Willmott, 2003).

Postmodernism and poststructuralism

Since the 1980s, postmodernism has attracted considerable interest among qualitative researchers. In terms of the philosophy of social sciences, postmodernism is a non-foundationalist epistemology. It rejects the positivist, rational and generalizable basis for scientific research, which would explain the world from an objective standpoint. The most extreme rejection of positivism within postmodernism is **poststructuralism.** It is derived from the idea of deconstruction (Derrida, 1978), which holds that there are no grounds for truth outside the text. Postmodernism in philosophy signs for epistemology with a small ‘e‘. This means that it rejects any common or shared ground for knowing. When doing this, it also rejects the ‘knower’ as an authority of any knowledge, e.g. in management; see Alvesson and Willmott (1996, 2003).

Postmodernism has had a strong appeal, especially within organization studies, culturally oriented marketing studies, and strategic management research. Knights and Morgan (1991) presented a classic analysis of corporate strategy as discourse, and Hassard and Parker (1993) edited a comprehensive collection of the postmodern

research in management studies. On the other hand, Alvesson and Willmott (2003) edited a book that took critical distance from postmodern research in management studies. It now seems that the influence of postmodernism is on the decline in business research, but it has had some influence on the way research is done, not only through the emphasis on language and its role, but also by emphasizing the heterogeneity of researched issues. A good overall account of the influence of postmodernism in management studies can be found in Johnson and Duberley (2000).

Two basic aspects of inquiry

How to bring forward knowledge about the world in research? There are two basic models of social science research, called **deduction** and **induction,** that to a large extent cover the different general models of science according to which scientific knowledge is achieved (Box 2.5). In addition, some researchers prefer to describe their study following the **abduction** logic. Inductive reasoning draws from observed cases more general statements or general claims about most cases of the same kind. Deductive reasoning is concerned with the formulation of hypotheses and theories from which particular phenomena can then be explained.

Even though a particular study may look like it is purely deductive (e.g. an experiment designed to test the hypothesized effects of some treatment on some outcome), most social research involves both inductive and deductive reasoning processes at some point in the same project. Therefore, it is good to keep in mind that labelling your research as deductive, inductive, or abductive does not, by itself,

[p. 22 ↓]

Box 2.5 Two basic models of research

- Deduction: theory is the first source of knowledge; research proceeds from theory, through hypothesis, to empirical analysis. A strict form is not necessarily suitable for qualitative research.

- Induction: theories are outcomes of empirical research; research proceeds from empirical research to theoretical results. Pure induction is rare, or even impossible.

tell the reader exactly how your research process proceeded. Nor does it justify your methodological choices.

If you want to use these terms in explaining the logic of your research, you should be careful to describe in more detail how you have applied inductive or deductive reasoning in the course of your research.

Deduction

Despite divergent business disciplines, the idea of *deduction* in research has been by far the strongest way to build up the theoretical knowledge base. Deduction rests on the idea that theory is the first source of knowledge. On the basis of what is known about a phenomenon theoretically, the researcher is able to deduce one or more hypotheses. The hypotheses are then subjected to empirical study. The process of deduction is linear, following the logic of proceeding from theory to empirical research. The certainty in theory development is gained through the hypothesis testing in empirical scrutiny. As multi-discipline approaches and differing ideas of the role and nature of theories in research have emerged, the strict deductive model of research is not considered suitable for most qualitative business research.

Induction

Much of the (quantitative) business research follows the logic of theory testing through hypothesis scrutiny in empirical world. However, many business study researchers find this model lacking because they see theories as outcomes of empirical research, not prior to it. Theories can also be seen as corrective mode concerning findings or even publications that come forward during the research process (Johnson and Duberley, 2000). Therefore, induction in research has gained a firm foothold. When you take the relationship between theory and empirical research as inductive, you follow the logic

of proceeding from empirical research to theoretical results. **[p. 23 ↓]** In other words, the research process develops, starting from empirical materials, not from theoretical propositions.

Abduction

These two 'ideal types' of research logic or traditions, deduction and induction, seldom exist as clear-cut alternatives. Many researchers use both induction and deduction in different phases of their study, which means that you move iteratively between these two during a research process. Some research methods books offer abduction as a way to combine deduction and induction in one research project. Abduction refers to the process of moving from the everyday descriptions and meanings given by people, to categories and concepts that create the basis of an understanding or an explanation to the phenomenon described.

Abduction, as defined by philosopher Charles Sanders Peirce, can be considered as the logic of exploratory data analysis. For Peirce, abduction referred to the process of generating new ideas or hypotheses. According to his idea, deduction can be used to evaluate the hypotheses and induction for justifying them with empirical data (Staat, 1993; Schwandt, 2001). In practice, abduction is difficult to dissect from the iterative work taking place in all empirical research. Some researchers also talk about the *hermeneutic circle* in much the same meaning and relate abduction closer to interpretivism. In general, no single model of scientific research is used, as the whole research process most often consists of various forms of reasoning.

This chapter has briefly illustrated the complexity of the conceptual grid of philosophy that embraces the knowledge of and about the different research methods. You do not always need to explicate your philosophical position and commitments in great detail, nor do you need to know the most advanced philosophical conceptual discussions when you write a thesis on business-related issues. However, it is most often useful to be knowledgeable of the key concepts and background assumptions of each method.

Finally, research methods and their use change and develop over time. So does the philosophical knowledge concerning the ways we can know about the social world

with various methods. It is important to keep in mind that, if the qualitative research approach that you choose is based on and shows an interest in philosophical questions, there is no excuse for you to be ignorant about them.

Key points of the chapter

- #
  Key concepts, their contents and relationships between ontology, epistemology, methodology and method are not fixed but there are different ways of understanding and relating the philosophical and methodological commitments in knowledge production to each other
- Main philosophical traditions and positions that are relevant for qualitative business research are positivism and its different forms, critical realism, interpretivism and constructionism, hermeneutics, postmodernism and poststructuralism.
- The concepts of deduction, induction and abduction are clarifying the way and directions for arguments and knowledge claims. However, they seldom can be found purely presented, but almost all qualitative research uses most of the logic.
- Reflexivity is increasingly important part of any research design and can be related to the basic premises of knowledge production, theories and methods used and to the results of the research. It valorizes the modes of knowledge production we sign for in our research.

Further reading

The book *Understanding Management Research* by Johnson and Duberley (2000) gives a fresh and balanced overview of the different paradigmatic views and research traditions within contemporary management research.

Exercise 1 Analysing the philosophical and methodological choices of qualitative research

The purpose of this exercise is to learn to understand the relevance of philosophical commitments in qualitative research and their connections to the overall logic of the research.

Go to the library of your university and choose one recently published doctoral thesis that is based on qualitative research. Read through the thesis by focusing on its philosophical background and the logic of the research in particular. Answer the following questions:

- What is the philosophical background of this research? How clearly does the researcher explicate their position?
- What is the logic of the research? How clearly does the researcher explicate the logic of her research? Does it follow inductive or deductive reasoning or both? In what ways?
- Do you find the philosophical background of the study compatible with its overall logic? Why? Why not?

http://dx.doi.org/10.4135/9780857028044.d8

Mixed Methods Research: A Research Paradigm Whose Time Has Come

by R. Burke Johnson and Anthony J. Onwuegbuzie

The purposes of this article are to position mixed methods research (*mixed research* is a synonym) as the natural complement to traditional qualitative and quantitative research, to present pragmatism as offering an attractive philosophical partner for mixed methods research, and to provide a framework for designing and conducting mixed methods research. In doing this, we briefly review the paradigm "wars" and incompatibility thesis, we show some commonalities between quantitative and qualitative research, we explain the tenets of pragmatism, we explain the fundamental principle of mixed research and how to apply it, we provide specific sets of designs for the two major types of mixed methods research (*mixed-model* designs and *mixed-method* designs), and, finally, we explain mixed methods research as following (recursively) an eight-step process. A key feature of mixed methods research is its methodological pluralism or eclecticism, which frequently results in superior research (compared to monomethod research). Mixed methods research will be successful as more investigators study and help advance its concepts and as they regularly practice it.

For more than a century, the advocates of quantitative and qualitative research paradigms have engaged in ardent dispute.[1] From these debates, purists have emerged on both sides (cf. Campbell & Stanley, 1963; Lincoln & Guba, 1985).[2] Quantitative purists (Ayer, 1959; Maxwell & Delaney, 2004; Popper, 1959; Schrag, 1992) articulate assumptions that are consistent with what is commonly called a positivist philosophy.[3, 4] That is, quantitative purists believe that social observations should be treated as entities in much the same way that physical scientists treat physical phenomena. Further, they contend that the observer is separate from the entities that are subject to observation. Quantitative purists maintain that social science inquiry should be objective. That is, time- and context-free generalizations (Nagel, 1986) are desirable and possible, and real causes of social scientific outcomes can be determined reliably and validly. According to this school of thought, educational researchers should eliminate their biases, remain emotionally detached and uninvolved with the objects of study, and test or empirically justify their stated hypotheses. These researchers have traditionally called for rhetorical neutrality, involving a formal writing style using the impersonal passive voice and technical terminology, in which establishing and describing social laws is the major focus (Tashakkori & Teddlie, 1998).

Qualitative purists (also called *constructivists* and *interpretivists*) reject what they call positivism. They argue for the superiority of constructivism, idealism, relativism, humanism, hermeneutics, and, sometimes, postmodernism (Guba & Lincoln, 1989; Lincoln & Guba, 2000; Schwandt, 2000; Smith, 1983, 1984). These purists contend that multiple-constructed realities abound, that time- and context-free generalizations are neither desirable nor possible, that research is value-bound, that it is impossible to differentiate fully causes and effects, that logic flows from specific to general (e.g., explanations are generated inductively from the data), and that knower and known cannot be separated because the subjective knower is the only source of reality (Guba, 1990). Qualitative purists also are characterized by a dislike of a detached and passive style of writing, preferring, instead, detailed, rich, and thick (empathic) description, written directly and somewhat informally.

Both sets of purists view their paradigms as the ideal for research, and, implicitly if not explicitly, they advocate the *incompatibility thesis* (Howe, 1988), which posits that qualitative and quantitative research paradigms, including their associated methods, cannot and should not be mixed. The quantitative versus qualitative debate has been so divisive that some graduate students who graduate from educational institutions with an aspiration to gain employment in the world of academia or research are left with the impression that they have to pledge allegiance to one research school of thought or the other. Guba (a leading qualitative purist) clearly represented the purist position when he contended that "accommodation between paradigms is impossible . . . we are led to vastly diverse, disparate, and totally antithetical ends" (Guba, 1990, p. 81). A disturbing feature of the paradigm wars has been the relentless focus on the differences between the two orientations. Indeed, the two dominant research paradigms have resulted in two research cultures, "one professing the superiority of 'deep, rich observational data' and the other the virtues of 'hard, generalizable' . . . data" (Sieber, 1973, p. 1335).

Our purpose in writing this article is to present mixed methods research as the third research paradigm in educational research.[5] We hope the field will move beyond quantitative *versus* qualitative research arguments because, as recognized by mixed methods research, *both* quantitative and qualitative research are important and useful. The goal of mixed methods research is not to replace either of these approaches but rather to draw from the

Educational Researcher, Vol. 33, No. 7, pp. 14–26

strengths and minimize the weaknesses of both in single research studies and across studies. If you visualize a continuum with qualitative research anchored at one pole and quantitative research anchored at the other, mixed methods research covers the large set of points in the middle area. If one prefers to think categorically, mixed methods research sits in a new third chair, with qualitative research sitting on the left side and quantitative research sitting on the right side.

Mixed methods research offers great promise for practicing researchers who would like to see methodologists describe and develop techniques that are closer to what researchers actually use in practice. Mixed methods research as the third research paradigm can also help bridge the schism between quantitative and qualitative research (Onwuegbuzie & Leech, 2004a). Methodological work on the mixed methods research paradigm can be seen in several recent books (Brewer & Hunter, 1989; Creswell, 2003; Greene, Caracelli, & Graham, 1989; Johnson & Christensen, 2004; Newman & Benz, 1998; Reichardt & Rallis, 1994; Tashakkori & Teddlie, 1998, 2003). Much work remains to be undertaken in the area of mixed methods research regarding its philosophical positions, designs, data analysis, validity strategies, mixing and integration procedures, and rationales, among other things. We will try to clarify the most important issues in the remainder of this article.

Commonalities Among the Traditional Paradigms

Although there are many important paradigmatic differences between qualitative and quantitative research (which have been frequently written about in the *Educational Researcher* and other places), there are some similarities between the various approaches that are sometimes overlooked. For example, both quantitative and qualitative researchers use empirical observations to address research questions. Sechrest and Sidani (1995, p. 78) point out that both methodologies "describe their data, construct explanatory arguments from their data, and speculate about why the outcomes they observed happened as they did." Additionally, both sets of researchers incorporate safeguards into their inquiries in order to minimize confirmation bias and other sources of invalidity (or lack of trustworthiness) that have the potential to exist in every research study (Sandelowski, 1986).

Regardless of paradigmatic orientation, all research in the social sciences represents an attempt to provide warranted assertions about human beings (or specific groups of human beings) and the environments in which they live and evolve (Biesta & Burbules, 2003). In the social and behavioral sciences, this goal of understanding leads to the examination of many different phenomena, including holistic phenomena such as intentions, experiences, attitudes, and culture, as well as more reductive phenomena such as macromolecules, nerve cells, micro-level homunculi, and biochemical computational systems (de Jong, 2003). There is room in ontology for mental and social reality as well as the more micro and more clearly material reality. Although certain methodologies tend to be associated with one particular research tradition, Dzurec and Abraham (1993, p. 75) suggest that "the objectives, scope, and nature of inquiry are consistent across methods and across paradigms." We contend that researchers and research methodologists need to be asking when each research approach is most helpful and when and how they should be mixed or combined in their research studies.

We contend that epistemological and methodological pluralism should be promoted in educational research so that researchers are informed about epistemological and methodological possibilities and, ultimately, so that we are able to conduct more effective research. Today's research world is becoming increasingly interdisciplinary, complex, and dynamic; therefore, many researchers need to complement one method with another, and all researchers need a solid understanding of multiple methods used by other scholars to facilitate communication, to promote collaboration, and to provide superior research. Taking a non-purist or compatibilist or mixed position allows researchers to mix and match design components that offer the best chance of answering their specific research questions. Although many research procedures or methods typically have been linked to certain paradigms, this linkage between research paradigm and research methods is neither sacrosanct nor necessary (Howe, 1988, 1992). For example, qualitative researchers should be free to use quantitative methods, and quantitative researchers should be free to use qualitative methods. Also, research in a content domain that is dominated by one method often can be better informed by the use of multiple methods (e.g., to give a read on methods-induced bias, for corroboration, for complimentarity, for expansion; see Greene et al., 1989). We contend that epistemological and paradigmatic ecumenicalism is within reach in the research paradigm of mixed methods research.

Philosophical Issues Debates

As noted by Onwuegbuzie and Teddlie (2003), some individuals who engage in the qualitative versus quantitative paradigm debate appear to confuse the *logic of justification* with research methods. That is, there is a tendency among some researchers to treat epistemology and method as being synonymous (Bryman, 1984; Howe, 1992). This is far from being the case because the logic of justification (an important aspect of epistemology) does not dictate what specific data collection and data analytical methods researchers must use. There is rarely entailment from epistemology to methodology (Johnson, Meeker, Loomis, & Onwuegbuzie, 2004; Phillips, 2004). For example, differences in epistemological beliefs (such as a difference in beliefs about the appropriate logic of justification) should not prevent a qualitative researcher from utilizing data collection methods more typically associated with quantitative research, and vice versa.

There are several interesting myths that appear to be held by some purists. For example, on the "positivist" side of the fence, the barriers that quantitative educational researchers have built arise from a narrow definition of the concept of "science." [6] As noted by Onwuegbuzie (2002), modern day "positivists" claim that science involves confirmation and falsification, and that these methods and procedures are to be carried out objectively. However, they disregard the fact that many human (i.e., subjective) decisions are made throughout the research process and that researchers are members of various social groups. A few examples of subjectivism and intersubjectivism in quantitative research include deciding what to study (i.e., what are the important problems?), developing instruments that are believed to measure what the researcher views as being the target construct, choosing the

specific tests and items for measurement, making score interpretations, selecting alpha levels (e.g., .05), drawing conclusions and interpretations based on the collected data, deciding what elements of the data to emphasize or publish, and deciding what findings are practically significant. Obviously, the conduct of fully objective and value-free research is a myth, even though the regulatory ideal of objectivity can be a useful one.

Qualitative researchers also are not immune from constructive criticism. Some qualitative purists (e.g., Guba, 1990) openly admit that they adopt an unqualified or strong relativism, which is logically self-refuting and (in its strong form) hinders the development and use of systematic standards for judging research quality (when it comes to research quality, it is not the case that anyone's opinion about quality is just as good as the next person's, because some people have no training or expertise or even interest in research). We suspect that most researchers are *soft* relativists (e.g., respecting the opinions and views of different people and different groups). When dealing with human research, soft relativism simply refers to a respect and interest in understanding and depicting individual and social group differences (i.e., their different perspectives) and a respect for democratic approaches to group opinion and value selection. Again, however, a strong relativism or strong constructivism runs into problems; for example, it is not a matter of opinion (or individual reality) that one should or can drive on the left-hand side of the road in Great Britain—if one chooses to drive on the right side, he or she will likely have a head-on collision, at some point, and end up in the hospital intensive care unit, or worse (this is a case where subjective and objective realities directly meet and clash). The strong ontological relativistic or constructivist claim in qualitative research that multiple, contradictory, but equally valid accounts of the same phenomenon are multiple *realities* also poses some potential problems. Generally speaking, subjective states (i.e., created and experienced realities) that vary from person to person and that are sometimes called "realities" should probably be called (for the purposes of clarity and greater precision) multiple *perspectives or opinions or beliefs* (depending on the specific phenomenon being described) rather than multiple realities (Phillips & Burbules, 2000). If a qualitative researcher insists on using the word reality for subjective states, then for clarity we would recommend that the word subjective be placed in front of the word reality (i.e., as in subjective reality or in many cases intersubjective reality) to direct the reader to the focus of the statement. We agree with qualitative researchers that value stances are often needed in research; however, it also is important that research is more than simply one researcher's highly idiosyncratic opinions written into a report. Fortunately, many strategies are recognized and regularly used in qualitative research (such as member checking, triangulation, negative case sampling, pattern matching, external audits) to help overcome this potential problem and produce high-quality and rigorous qualitative research. Finally, qualitative researchers sometimes do not pay due attention to providing an adequate rationale for interpretations of their data (Onwuegbuzie, 2000), and qualitative methods of analyses too "often remain private and unavailable for public inspection" (Constas, 1992, p. 254). Without public inspection and adequate standards, how is one to decide whether what is claimed is trustworthy or defensible?

Fortunately, many (or most?) qualitative researchers and quantitative researchers (i.e., postpositivists) have now reached basic agreement on several major points of earlier philosophical disagreement (e.g., Phillips & Burbules, 2000; Reichardt & Cook, 1979; Reichardt & Rallis, 1994). Basic agreement has been reached on each of the following issues: (a) the relativity of the "light of reason" (i.e., what appears reasonable can vary across persons); (b) theory-laden perception or the theory-ladenness of facts (i.e., what we notice and observe is affected by our background knowledge, theories, and experiences; in short, observation is not a perfect and direct window into "reality"); (c) underdetermination of theory by evidence (i.e., it is possible for more than one theory to fit a single set of empirical data); (d) the Duhem-Quine thesis or idea of auxiliary assumptions (i.e., a hypothesis cannot be fully tested in isolation because to make the test we also must make various assumptions; the hypothesis is embedded in a holistic network of beliefs; and alternative explanations will continue to exist); (e) the problem of induction (i.e., the recognition that we only obtain probabilistic evidence, not final proof in empirical research; in short, we agree that the future may not resemble the past); (f) the social nature of the research enterprise (i.e., researchers are embedded in communities and they clearly have and are affected by their attitudes, values, and beliefs); and (g) the value-ladenness of inquiry (this is similar to the last point but specifically points out that human beings can never be completely value free, and that values affect what we choose to investigate, what we see, and how we interpret what we see).

Pragmatism as the Philosophical Partner for Mixed Methods Research

We do not aim to solve the metaphysical, epistemological, axiological (e.g., ethical, normative), and methodological differences between the purist positions. And we do not believe that mixed methods research is currently in a position to provide perfect solutions. Mixed methods research should, instead (at this time), use a method and philosophy that attempt to fit together the insights provided by qualitative and quantitative research into a workable solution. Along these lines, we advocate consideration of the pragmatic method of the classical pragmatists (e.g., Charles Sanders Peirce, William James, and John Dewey) as a way for researchers to think about the traditional dualisms that have been debated by the purists. Taking a pragmatic and balanced or pluralist position will help improve communication among researchers from different paradigms as they attempt to advance knowledge (Maxcy, 2003; Watson, 1990). Pragmatism also helps to shed light on how research approaches can be mixed fruitfully (Hoshmand, 2003); the bottom line is that research approaches should be mixed in ways that offer the best opportunities for answering important research questions.

The *pragmatic rule* or *maxim* or *method* states that the current meaning or instrumental or provisional truth value (which James [1995, 1907 original] would term "cash value") of an expression (e.g., "all reality has a material base" or "qualitative research is superior for uncovering humanistic research findings") is to be determined by the experiences or practical consequences of belief in or use of the expression in the world (Murphy, 1990). One can apply this sensible effects- or outcome-oriented rule through thinking (thinking about what will happen if you do X), practi-

cal experiences (observing what happens in your experience when you do X), or experiments (formally or informally trying a rule and observing the consequences or outcomes).

In the words of Charles Sanders Peirce (1878), the pragmatic method or maxim (which is used to determine the *meaning* of words, concepts, statements, ideas, beliefs) implies that we should "consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then our conception of these effects is the whole of our conception of the object" (this quote is found at the end of Section II in *How to Make Our Ideas Clear*). Building on Peirce's lead, James (1995, 1907 original) argued that "The pragmatic method is primarily a method of settling metaphysical disputes that otherwise might be interminable. . . . The pragmatic method in such cases is to try to interpret each notion by tracing its respective practical consequences" (p. 18). Extending the works of Peirce and James, Dewey spent his career applying pragmatic principles in developing his philosophy and in the practice of educating children (e.g., the Experimental School of Chicago). Dewey (1948, 1920 original) stated that "in order to discover the meaning of the idea [we must] ask for its consequences" (p. 132). In short, when judging ideas we should consider their empirical and practical consequences. Peirce, James, and Dewey were all interested in examining practical consequences and empirical findings to help in understanding the import of philosophical positions and, importantly, to help in deciding which action to take next as one attempts to better understand real-world phenomena (including psychological, social, and educational phenomena).

If two ontological positions about the mind/body problem (e.g., monism versus dualism), for example, do not make a difference in how we conduct our research then the distinction is, for practical purposes, not very meaningful. We suspect that some philosophical differences may lead to important practical consequences while many others may not.[7] The full sets of beliefs characterizing the qualitative and quantitative approaches or paradigms have resulted in different practices, and, based on our observation and study, we believe it is clear that both qualitative and quantitative research have many benefits and many costs. In some situations the qualitative approach will be more appropriate; in other situations the quantitative approach will be more appropriate. In many situations, researchers can put together insights and procedures from both approaches to produce a superior product (i.e., often mixed methods research provides a more workable solution and produces a superior product). We are advocating a needs-based or contingency approach to research method and concept selection.

Philosophical debates will not end as a result of pragmatism, and certainly they *should not* end. Nonetheless, we agree with others in the mixed methods research movement that consideration and discussion of pragmatism by research methodologists and empirical researchers will be productive because it offers an immediate and useful middle position philosophically and methodologically; it offers a practical and outcome-oriented method of inquiry that is based on action and leads, iteratively, to further action and the elimination of doubt; and it offers a method for selecting methodological mixes that can help researchers better answer many of their research questions. Pragmatically inclined philosophers and researchers also would suggest that we can reach some agreement about the importance of many (culturally derived) values and desired ends, such as, for example, preventing the dropping out of school by adolescents, reducing the use of illicit drugs by children and adolescents, finding effective teaching techniques for different kinds of students, educating children and adults (i.e., increasing their knowledge), helping to reduce discrimination in society, and attempting to eliminate or reduce mental, learning, and other disabilities. In other words, pragmatism takes an explicitly value-oriented approach to research.

We reject an incompatibilist, either/or approach to paradigm selection and we recommend a more pluralistic or compatibilist approach. Beyond the basic pragmatic method or maxim (i.e., translated in mixed methods research as "choose the combination or mixture of methods and procedures that works best for answering your research questions") there also is a full philosophical system of pragmatism which was systematically developed by the classical pragmatists (Peirce, James, Dewey) and has been refined in newer directions by latter-day neo-pragmatists (e.g., Davidson, Rescher, Rorty, Putnam) (see Menand, 1997; Murphy, 1990; Rescher, 2000; Rorty, 2000). To provide the reader with a better understanding of the full philosophy of pragmatism (for consideration), we have outlined, in Table 1, what we believe are classical pragmatism's most general and important characteristics.

Although we endorse pragmatism as a philosophy that can help to build bridges between conflicting philosophies, pragmatism, like all current philosophies, has some shortcomings. In Table 2 we present some of these. Researchers who are interested in applying pragmatism in their works should consider the shortcomings, which also need to be addressed by philosophically inclined methodologists as they work on the project of developing a fully working philosophy for mixed methods research. Practicing researchers should be reflexive and strategic in avoiding the potential consequences of these weaknesses in their works.

Comparing Qualitative, Quantitative, and Mixed Methods Research

Mixed methods research is formally defined here as *the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study*. Philosophically, it is the "third wave" or third research movement, a movement that moves past the paradigm wars by offering a logical and practical alternative. Philosophically, mixed research makes use of the pragmatic method and system of philosophy. Its logic of inquiry includes the use of induction (or discovery of patterns), deduction (testing of theories and hypotheses), and abduction (uncovering and relying on the best of a set of explanations for understanding one's results) (e.g., de Waal, 2001).

Mixed methods research also is an attempt to legitimate the use of multiple approaches in answering research questions, rather than restricting or constraining researchers' choices (i.e., it rejects dogmatism). It is an expansive and creative form of research, not a limiting form of research. It is inclusive, pluralistic, and complementary, and it suggests that researchers take an eclectic approach to method selection and the thinking about and conduct of research. What is most fundamental is the research question—research methods should *follow* research questions in a way that

Table 1
General Characteristics of Pragmatism

- The project of pragmatism has been to find a middle ground between philosophical dogmatisms and skepticism and to find a workable solution (sometimes including outright rejection) to many longstanding philosophical dualisms about which agreement has not been historically forthcoming.
- Rejects traditional dualisms (e.g., rationalism vs. empiricism, realism vs. antirealism, free will vs. determinism, Platonic appearance vs. reality, facts vs. values, subjectivism vs. objectivism) and generally prefers more moderate and commonsense versions of philosophical dualisms based on how well they work in solving problems.
- Recognizes the existence and importance of the natural or physical world as well as the emergent social and psychological world that includes language, culture, human institutions, and subjective thoughts.
- Places high regard for the reality of and influence of the inner world of human experience in action.
- Knowledge is viewed as being both constructed *and* based on the reality of the world we experience and live in.
- Replaces the historically popular epistemic distinction between subject and external object with the naturalistic and process-oriented organism-environment transaction.
- Endorses fallibilism (current beliefs and research conclusions are rarely, if ever, viewed as perfect, certain, or absolute).
- Justification comes in the form of what Dewey called "warranted assertability."
- According to Peirce, "reasoning should not form a chain which is no stronger than its weakest link, but a cable whose fibers may be ever so slender, provided they are sufficiently numerous and intimately connected" (1868, in Menand, 1997, pp. 5–6).
- Theories are viewed instrumentally (they become true and they are true to different degrees based on how well they currently work; workability is judged especially on the criteria of predictability and applicability).
- Endorses eclecticism and pluralism (e.g., different, even conflicting, theories and perspectives can be useful; observation, experience, and experiments are all useful ways to gain an understanding of people and the world).
- Human inquiry (i.e., what we do in our day-to-day lives as we interact with our environments) is viewed as being analogous to experimental and scientific inquiry. We all try out things to see what works, what solves problems, and what helps us to survive. We obtain warranted evidence that provides us with answers that are ultimately tentative (i.e., inquiry provides the best answers we can currently muster), but, in the long run, use of this "scientific" or evolutionary or practical epistemology moves us toward larger Truths.
- Endorses a strong and practical empiricism as the path to determine what works.
- Views current truth, meaning, and knowledge as tentative and as changing over time. What we obtain on a daily basis in research should be viewed as provisional truths.
- Capital "T" Truth (i.e., absolute Truth) is what will be the "final opinion" perhaps at the end of history. Lowercase "t" truths (i.e., the instrumental and provisional truths that we obtain and live by in the meantime) are given through experience and experimenting.
- Instrumental truths are a matter of degree (i.e., some estimates are more true than others). Instrumental truth is not "stagnant," and, therefore, James (1995: 1907) states that we must "be ready tomorrow to call it falsehood."
- Prefers action to philosophizing (pragmatism is, in a sense, an anti-philosophy).
- Takes an explicitly value-oriented approach to research that is derived from cultural values; specifically endorses shared values such as democracy, freedom, equality, and progress.
- Endorses practical theory (theory that informs effective practice; praxis).
- Organisms are constantly adapting to new situations and environments. Our thinking follows a dynamic homeostatic process of belief, doubt, inquiry, modified belief, new doubt, new inquiry, . . . , in an infinite loop, where the person or researcher (and research community) constantly tries to improve upon past understandings in a way that fits and works in the world in which he or she operates. The present is always a new starting point.
- Generally rejects reductionism (e.g., reducing culture, thoughts, and beliefs to nothing more than neurobiological processes).
- Offers the "pragmatic method" for solving traditional philosophical dualisms as well as for making methodological choices.

offers the best chance to obtain useful answers. Many research questions and combinations of questions are best and most fully answered through mixed research solutions.

In order to mix research in an effective manner, researchers first need to consider all of the relevant characteristics of quantitative and qualitative research. For example, the major characteristics of traditional *quantitative* research are a focus on deduction, confirmation, theory/hypothesis testing, explanation, prediction, standardized data collection, and statistical analysis (see Table 3 for a more complete list). The major characteristics of traditional *qualitative* research are induction, discovery, exploration, theory/hypothesis generation, the researcher as the primary "instrument" of data collection, and qualitative analysis (see Table 4 for a more complete list).

Gaining an understanding of the strengths and weaknesses of quantitative and qualitative research puts a researcher in a position to mix or combine strategies and to use what Johnson and Turner (2003) call the *fundamental principle of mixed research.* According to this principle, researchers should collect multiple data using different strategies, approaches, and methods in such a way that the resulting mixture or combination is likely to result in complementary strengths and nonoverlapping weaknesses (also see Brewer & Hunter, 1989). Effective use of this principle is a major source of justification for mixed methods research because the product will be superior to monomethod studies. For example, adding qualitative interviews to experiments as a manipulation check and perhaps as a way to discuss directly the issues under investigation and tap into participants' perspectives

Table 2
Some Weaknesses of Pragmatism

- Basic research may receive less attention than applied research because applied research may appear to produce more immediate and practical results.
- Pragmatism may promote incremental change rather than more fundamental, structural, or revolutionary change in society.
- Researchers from a transformative-emancipatory framework have suggested that pragmatic researchers sometimes fail to provide a satisfactory answer to the question "For whom is a pragmatic solution useful?" (Mertens, 2003).
- What is meant by usefulness or workability can be vague unless explicitly addressed by a researcher.
- Pragmatic theories of truth have difficulty dealing with the cases of *useful but non-true* beliefs or propositions and *non-useful but true* beliefs or propositions.
- Many come to pragmatism looking for a way to get around many traditional philosophical and ethical disputes (this includes the developers of pragmatism). Although pragmatism has worked moderately well, when put under the microscope, many current philosophers have rejected pragmatism because of its logical (as contrasted with practical) failing as a solution to many philosophical disputes.
- Some neo-pragmatists such as Rorty (and postmodernists) completely reject correspondence truth in any form, which troubles many philosophers.

Table 3
Strengths and Weaknesses of Quantitative Research

Strengths

- Testing and validating already constructed theories about how (and to a lesser degree, why) phenomena occur.
- Testing hypotheses that are constructed before the data are collected. Can generalize research findings when the data are based on random samples of sufficient size.
- Can generalize a research finding when it has been replicated on many different populations and subpopulations.
- Useful for obtaining data that allow quantitative predictions to be made.
- The researcher may construct a situation that eliminates the confounding influence of many variables, allowing one to more credibly assess *cause-and-effect* relationships.
- Data collection using some quantitative methods is relatively quick (e.g., telephone interviews).
- Provides precise, quantitative, numerical data.
- Data analysis is relatively less time consuming (using statistical software).
- The research results are relatively independent of the researcher (e.g., effect size, statistical significance).
- It may have higher credibility with many people in power (e.g., administrators, politicians, people who fund programs).
- It is useful for studying large numbers of people.

Weaknesses

- The researcher's categories that are used may not reflect local constituencies' understandings.
- The researcher's theories that are used may not reflect local constituencies' understandings.
- The researcher may miss out on phenomena occurring because of the focus on theory or hypothesis *testing* rather than on theory or hypothesis *generation* (called the *confirmation bias*).
- Knowledge produced may be too abstract and general for direct application to specific local situations, contexts, and individuals.

and meanings will help avoid some potential problems with the experimental method. As another example, in a qualitative research study the researcher might want to qualitatively observe and interview, but supplement this with a closed-ended instrument to systematically measure certain factors considered important in the relevant research literature. Both of these examples could be improved (if the research questions can be studied this way) by adding a component that surveys a randomly selected sample from the population of interest to improve generalizability. If findings are corroborated across different approaches then greater confidence can be held in the singular conclusion; if the findings conflict then the researcher has greater knowledge and can modify interpretations and conclusions accordingly. In many cases the goal of mixing is not to search for corroboration but rather to expand one's understanding (Onwuegbuzie & Leech, 2004b).

Tables 3 and 4 are specifically designed to aid in the construction of a combination of qualitative and quantitative research. After determining one's research question(s), one can decide whether mixed research offers the best potential for an answer; if this is the case, then one can use the tables as an aid to help in deciding on the combination of complementary strengths and nonoverlapping weaknesses that is appropriate for a particular study. Table 5 shows some of the strengths and weaknesses of mixed methods research, which should aid in the decision to use or not use a mixed methods research approach for a given research study.

Development of a Mixed Methods Research Typology

Our mixed methods research typologies (*mixed-model* designs and *mixed-method* designs) resulted from our consideration of many other typologies (especially Creswell, 1994; Morgan, 1998; Morse, 1991; Patton, 1990; and Tashakkori & Teddlie, 1998), as well as several dimensions which one should consider when planning to conduct a mixed research study. For example, it has been noted that one can construct mixed-model designs by mixing qualitative and quantitative approaches within and across the stages of research (in a simplified view, one can consider a single study as having three *stages*: stating the research objective, collecting the data, and analyzing/interpreting the data; see mixed-model designs in Johnson & Christensen, 2004; Tashakkori & Teddlie, 1998). According to Morgan (1998) and Morse (1991), one also may consider the dimension of paradigm emphasis (deciding whether to give the quantitative and qualitative components of a mixed study equal status or to give one paradigm the dominant status). Time ordering of the qualitative and quantitative phases is another important dimension, and the phases can be carried out sequentially or concurrently. Our mixed-method designs (discussed below) are based on the crossing of paradigm emphasis and time ordering of the quantitative and qualitative phases. Another dimension for viewing mixed methods re-

Table 4
Strengths and Weaknesses of Qualitative Research

Strengths

- The data are based on the participants' own categories of meaning.
- It is useful for studying a limited number of cases in depth.
- It is useful for describing complex phenomena.
- Provides individual case information.
- Can conduct cross-case comparisons and analysis.
- Provides understanding and description of people's personal experiences of phenomena (i.e., the "emic" or insider's viewpoint).
- Can describe, in rich detail, phenomena as they are situated and embedded in local contexts.
- The researcher identifies contextual and setting factors as they relate to the phenomenon of interest.
- The researcher can study dynamic processes (i.e., documenting sequential patterns and change).
- The researcher can use the primarily qualitative method of "grounded theory" to generate inductively a tentative but explanatory theory about a phenomenon.
- Can determine how participants interpret "constructs" (e.g., self-esteem, IQ).
- Data are usually collected in naturalistic settings in qualitative research.
- Qualitative approaches are responsive to local situations, conditions, and stakeholders' needs.
- Qualitative researchers are responsive to changes that occur during the conduct of a study (especially during extended fieldwork) and may shift the focus of their studies as a result.
- Qualitative data in the words and categories of participants lend themselves to exploring how and why phenomena occur.
- One can use an important case to demonstrate vividly a phenomenon to the readers of a report.
- Determine *idiographic* causation (i.e., determination of causes of a particular event).

Weaknesses

- Knowledge produced may not generalize to other people or other settings (i.e., findings may be unique to the relatively few people included in the research study).
- It is difficult to make quantitative predictions.
- It is more difficult to test hypotheses and theories.
- It may have lower credibility with some administrators and commissioners of programs.
- It generally takes more time to collect the data when compared to quantitative research.
- Data analysis is often time consuming.
- The results are more easily influenced by the researcher's personal biases and idiosyncrasies.

search is the degree of mixture, which would form a continuum from monomethod to fully mixed methods. Another dimension pertains to where mixing should occur (e.g., in the objective[s], methods of data collection, research methods, during data analysis, data interpretation). Yet another important dimension is whether one wants to take a critical theory/transformative-emancipatory (Mertens, 2003) approach or a less explicitly ideological approach to a study. Ultimately, the possible number of ways that studies can involve mixing is very large because of the many potential classification dimensions. It is a key point that mixed methods research truly opens up an exciting and almost unlimited potential for future research.

Toward a Parsimonious Typology of Mixed Research Methods

The majority of mixed methods research designs can be developed from the two major types of mixed methods research: *mixed-model* (mixing qualitative and quantitative approaches within or across the *stages* of the research process) and *mixed-method* (the inclusion of a quantitative *phase* and a qualitative *phase* in an overall research study). Six mixed-model designs are shown in Figure 1 (see Designs 2 through 7). These six designs are called *across-stage mixed-model designs* because the mixing takes place across the stages of the research process. An example of a *within-stage mixed-model design* would be the use of a questionnaire that includes a summated rating scale (quantitative data collection) and one or more open-ended questions (qualitative data collection).

Nine mixed-method designs are provided in Figure 2. The notation used (based on Morse, 1991) is explained at the bottom of the table. To construct a mixed-method design, the researcher must make two primary decisions: (a) whether one wants to operate largely within one dominant paradigm or not, and (b) whether one wants to conduct the phases concurrently or sequentially. In contrast to mixed-model designs, mixed-method designs are similar to conducting a quantitative mini-study and a qualitative mini-study in one overall research study. Nonetheless, to be considered a mixed-method design, the findings must be mixed or integrated at some point (e.g., a qualitative phase might be conducted to inform a quantitative phase, sequentially, or if the quantitative and qualitative phases are undertaken concurrently the findings must, at a minimum, be integrated during the interpretation of the findings).

It is important to understand that *one can easily create more user specific and more complex designs* than the ones shown in Figures 1 and 2. For example, one can develop a mixed-method design that has more stages (e.g., Qual → QUAN → Qual); one also can design a study that includes both mixed-model and mixed-method design features. The point is for the researcher to be creative and not be limited by the designs listed in this article. Furthermore, sometimes a design may emerge during a study in new ways, depending on the conditions and information that is obtained. A tenet of mixed methods research is that researchers should mindfully create designs that effectively answer their research questions; this stands in contrast to the common approach in traditional quantitative research where students are given a menu of designs from which to select.[8, 9] It also stands in stark contrast to the approach where one completely follows either the qualitative paradigm or the quantitative paradigm.

Table 5
Strengths and Weaknesses of Mixed Research

Strengths

- Words, pictures, and narrative can be used to add meaning to numbers.
- Numbers can be used to add precision to words, pictures, and narrative.
- Can provide quantitative and qualitative research strengths (i.e., see strengths listed in Tables 3 and 4).
- Researcher can generate and test a grounded theory.
- Can answer a broader and more complete range of research questions because the researcher is not confined to a single method or approach.
- The specific mixed *research designs* discussed in this article have specific strengths and weaknesses that should be considered (e.g., in a two-stage sequential design, the Stage 1 results can be used to develop and inform the purpose and design of the Stage 2 component).
- A researcher can use the strengths of an additional method to overcome the weaknesses in another method by using both in a research study.
- Can provide stronger evidence for a conclusion through convergence and corroboration of findings.
- Can add insights and understanding that might be missed when only a single method is used.
- Can be used to increase the generalizability of the results.
- Qualitative and quantitative research used together produce more complete knowledge necessary to inform theory and practice.

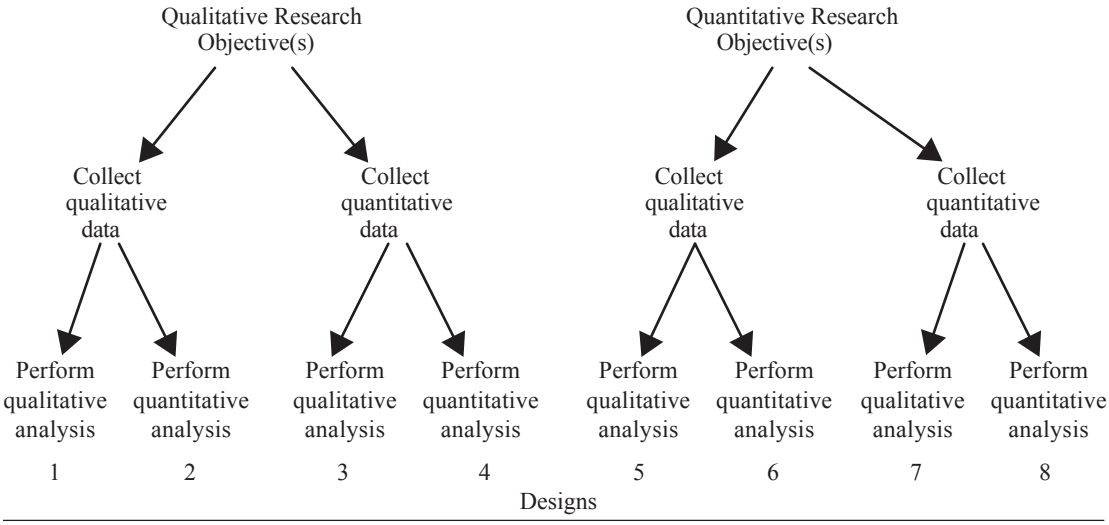
Weaknesses

- Can be difficult for a single researcher to carry out both qualitative and quantitative research, especially if two or more approaches are expected to be used concurrently; it may require a research team.
- Researcher has to learn about multiple methods and approaches and understand how to mix them appropriately.
- Methodological purists contend that one should always work within either a qualitative or a quantitative paradigm.
- More expensive.
- More time consuming.
- Some of the details of mixed research remain to be worked out fully by research methodologists (e.g., problems of paradigm mixing, how to qualitatively analyze quantitative data, how to interpret conflicting results).

A Mixed Methods Research Process Model

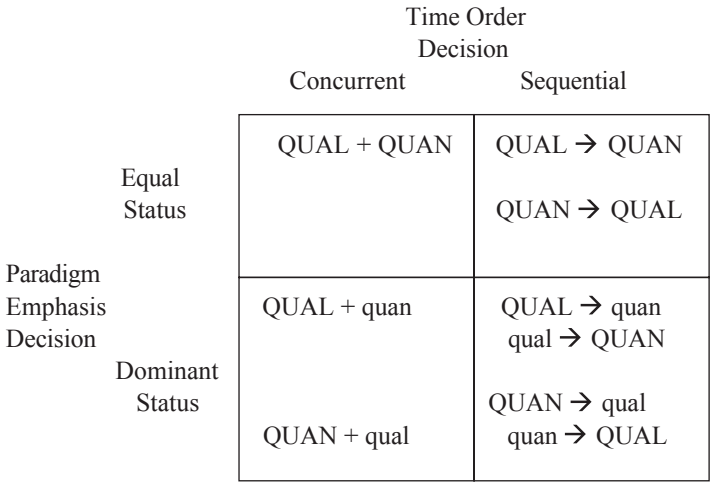
Our mixed methods research process model comprises eight distinct steps: (1) determine the research question; (2) determine whether a mixed design is appropriate; (3) select the mixed-method or mixed-model research design; (4) collect the data; (5) analyze the data; (6) interpret the data; (7) legitimate the data; and (8) draw conclusions (if warranted) and write the final report. These steps are displayed in Figure 3. Although mixed research starts with a purpose and one or more research questions, the rest of the steps can vary in order (i.e., they are not necessarily linear or unidirectional), and even the question and/or purpose can be revised when needed. Figure 3 shows several arrows leading from later steps to earlier steps indicating that mixed research involves a cyclical, recursive, and interactional process. Recursion can take place within a single study (especially an extended study); recursion can also take place across related studies by informing future research and leading to new or reformulated research purposes and questions.

Three steps in the mixed methods research process warrant some further discussion, especially purpose (Step 2), data analysis (Step 5), and legitimation (Step 7). As noted by Greene et al. (1989), there are five major purposes or rationales for conducting



Note. Designs 1 and 8 on the outer edges are the monomethod designs. The mixed-model designs are Designs 2, 3, 4, 5, 6, and 7.[10]

FIGURE 1. *Monomethod and mixed-model designs.*



Note. "qual" stands for qualitative, "quan" stands for quantitative, "+" stands for concurrent, "→" stands for sequential, capital letters denote high priority or weight, and lower case letters denote lower priority or weight.[11]

FIGURE 2. *Mixed-method design matrix with mixed-method research designs shown in the four cells.*

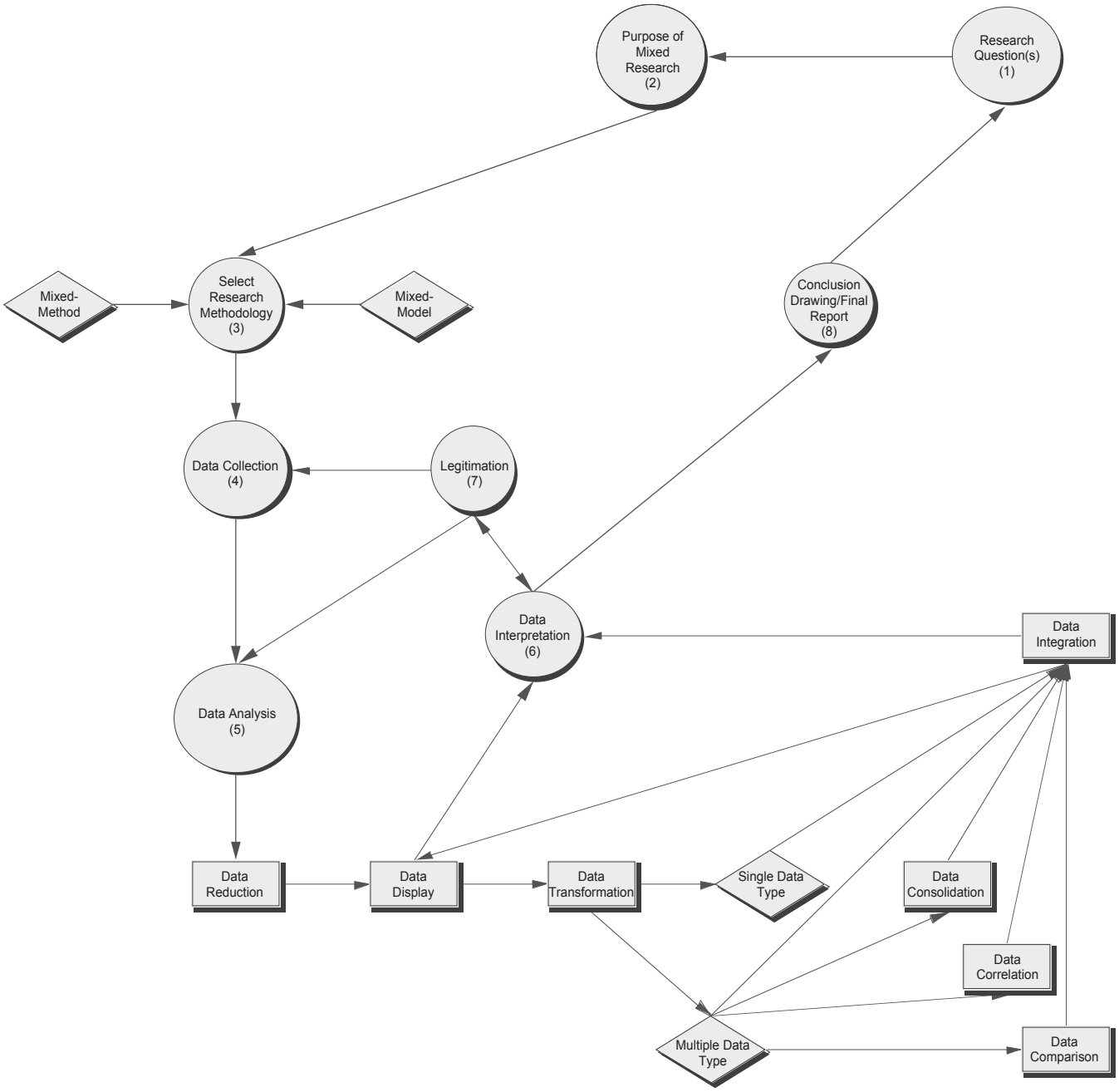
mixed methods research: (a) triangulation (i.e., seeking convergence and corroboration of results from different methods and designs studying the same phenomenon); (b) complementarity (i.e., seeking elaboration, enhancement, illustration, and clarification of the results from one method with results from the other method); (c) initiation (i.e., discovering paradoxes and contradictions that lead to a re-framing of the research question); (d) development (i.e., using the findings from one method to help inform the other method); and (e) expansion (i.e., seeking to expand the breadth and range of research by using different methods for different inquiry components).

The mixed methods research process model incorporates Onwuegbuzie and Teddlie's (2003) seven-stage conceptualization of the mixed methods data analysis process. According to these authors, the seven data analysis stages are as follows: (a) data reduction, (b) data display, (c) data transformation, (d) data correlation, (e) data consolidation, (f) data comparison, and (g) data integration. *Data reduction* involves reducing the dimensionality of the qualitative data (e.g., via exploratory thematic analysis, memoing) and quantitative data (e.g., via descriptive statistics, exploratory factor analysis, cluster analysis). *Data display*, involves describing pictorially the qualitative data (e.g., matrices, charts, graphs, networks, lists, rubrics, and Venn diagrams) and quantitative data (e.g., tables, graphs). This is followed (optionally) by the *data transformation* stage, wherein quantitative data are converted into narrative data that can be analyzed qualitatively (i.e., *qualitized*; Tashakkori & Teddlie, 1998) and/or qualitative data are converted into numerical codes that can be represented statistically (i.e., *quantitized*; Tashakkori & Teddlie, 1998). *Data correlation* involves the quantitative data being correlated with the qualitized data or the qualitative data being correlated with the quantitized data. This is followed by *data consolidation*, wherein both quantitative and qualitative data are combined to create new or consolidated variables or data sets. The next stage, *data comparison* involves comparing data from the qualitative and quantitative data sources. *Data integration* characterizes the final stage, whereby both quantitative and qualitative data are integrated into either a coherent whole or two separate sets (i.e., qualitative and quantitative) of coherent wholes.

The *legitimation* step involves assessing the trustworthiness of both the qualitative and quantitative data and subsequent interpretations. Frameworks such as the Quantitative Legitimation Model (Onwuegbuzie, 2003; which contains 50 sources of invalidity for the quantitative component of the mixed methods research at the data collection, data analysis, and data interpretation stages of the study) and the Qualitative Legitimation Model (Onwuegbuzie, 2000; Onwuegbuzie, Jiao, & Bostick, 2004; which contains 29 elements of legitimation for the qualitative component of the mixed methods research at the data collection, data analysis, and data interpretation stages of the study) can be used to assess the legitimacy of the qualitative and quantitative phases of the study, respectively. We have begun working on a validity or legitimation typology specifically for mixed research in Onwuegbuzie and Johnson (2004). It is important to note that the legitimation process might include additional data collection, data analysis, and/or data interpretation until as many rival explanations as possible have been reduced or eliminated.

The Future of Mixed Methods Research in Education

Mixed research actually has a long history in research practice because practicing researchers frequently ignore what is written by methodologists when they feel a mixed approach will best help them to answer their research questions. It is time that methodologists catch up with practicing researchers! It is now time that all researchers and research methodologists formally recognize the third research paradigm and begin systematically writing about it and using it. In general we recommend *contingency theory* for research approach selection, which accepts that quantitative, qualitative, and mixed research *are all superior under different cir-*



Note. Circles represent steps (1–8) in the mixed research process; rectangles represent steps in the mixed data analysis process; diamonds represent components.

FIGURE 3. *Mixed research process model.*

cumstances and it is the researcher's task to examine the specific contingencies and make the decision about which research approach, or which combination of approaches, should be used in a specific study. In this article we have outlined the philosophy of pragmatism, we have described mixed research and provided specific mixed-model and mixed-method designs, and we have discussed the fundamental principle of mixed research and provided tables of quantitative and qualitative research strengths and weaknesses to help apply the principle. Also, we have provided a mixed methods process model to help readers visualize the process. We hope we have made the case that mixed methods research is here to stay and that it should be widely recognized in education, as well as in our sister disciplines in the social and behavioral sciences, as the third major research paradigm.

As noted by Sechrest and Sidana (1995), growth in the mixed methods (i.e., pragmatist) movement has the potential to reduce some of the problems associated with singular methods. By utilizing quantitative and qualitative techniques within the same framework, mixed methods research can incorporate the strengths of both methodologies. Most importantly, investigators who conduct mixed methods research are more likely to select methods and approaches with respect to their underlying research questions, rather than with regard to some preconceived biases about which research paradigm should have hegemony in

social science research. By narrowing the divide between quantitative and qualitative researchers, mixed methods research has a great potential to promote a shared responsibility in the quest for attaining accountability for educational quality. The time has come for mixed methods research.

NOTES

[1] Thomas Kuhn (1962) popularized the idea of a paradigm. Later, when he was asked to explain more precisely what he meant by the term, he pointed out that it was a general concept and that it included a group of researchers having a common education and an agreement on "exemplars" of high quality research or thinking (Kuhn, 1977). In this article, by *research paradigm* we mean a set of beliefs, values, and assumptions that a community of researchers has in common regarding the nature and conduct of research. The beliefs include, but are not limited to, ontological beliefs, epistemological beliefs, axiological beliefs, aesthetic beliefs, and methodological beliefs. In short, as we use the term, a research paradigm refers to a research culture. We will be arguing that there is now a trilogy of major research paradigms: qualitative research, quantitative research, and mixed methods research.

[2] Campbell modified his view of qualitative research over time. For example, based on criticisms by qualitative and case study researchers of his term "one-shot case study" (which, unfortunately, is still used in several educational research books), Campbell changed this design name to the one-group posttest-only design; he made this change as part of his endorsement of case study research as an important research approach (e.g., see Campbell's introduction to Yin's case study research book: Yin, 1984).

[3] We do not mean to imply that there is anything inherently wrong with taking an extreme intellectual position. Most of the great thinkers in the history of philosophy and science (including social and behavioral science) were "extreme" for their times. Also, both qualitative and quantitative philosophies continue to be highly useful (i.e., both have many advantages when used in their pure forms).

[4] *Positivism* is a poor choice for labeling quantitative researchers today because positivism has long been replaced by newer philosophies of science (Yu, 2003). The term is more of a straw man (easily knocked down) for attack than standing for any actual practicing researchers. A term that better represents today's practicing quantitative researchers is *postpositivism* (Phillips & Burbules, 2000).

[5] Both of the authors of the current article prefer the label *mixed research* or *integrative research* rather than *mixed methods research*. The alternative labels are broader, more inclusive, and more clearly paradigmatic. We chose to use the term *mixed methods* in this article because of its current popularity.

[6] Here is a practical definition of science from an educational research textbook (Johnson & Christensen, 2004) that should be inclusive of quantitative and qualitative research: ". . . the root of the word science is the Latin *scientia*, which simply means 'knowledge.' *We define* science *in this book in a way that is inclusive of the different approaches to educational research*. We define science as an approach for the generation of knowledge that places high regard for empirical data and follows certain norms and practices that develop over time because of their usefulness. . . . The ultimate objective of most social, behavioral, and educational research is improvement of the world or social betterment."

[7] This is a very interesting empirical question that deserves more attention in the literature.

[8] Note that Shadish, Cook, and Campbell (2002) have attempted to move quantitative research away from this traditional "menu" approach. In this latest edition of Campbell and Stanley (1963), there is increased focus on understanding how to construct or create a research design that fits a particular situation.

[9] For additional mixed-method designs, see Creswell, Plano, Clark, Guttmann, and Hanson, 2003; Maxwell and Loomis, 2003.

[10] Here is the etiology of Figure 1: As far as we know, Patton (1990) first listed 6 of the mixed model designs (Designs 1, 2, 3, 5, 6, and 8). Then Tashakkori and Teddlie (1998) built on this by adding two designs (Designs 4 and 7) that were left out by Patton and they changed some labels to better fit their thinking (e.g., they introduced the term *mixed model*). Finally, in its present form, we first used (in an AERA conference paper) the full set of eight designs identified by Tashakkori and Teddlie (1998) while changing some labels to better fit our conceptualization. The term monomethods probably originated in Campbell and Fiske (1959).

[11] In developing Figure 2, we were probably most influenced by Morgan (1998), Morse (1991), and Tashakkori and Teddlie (1998). Several of the designs shown in the figure were introduced by Morse (1991).

REFERENCES

Ayer, A. J. (1959). *Logical positivism*. New York: The Free Press.

Biesta, G. J. J., & Burbules, N. C. (2003). *Pragmatism and educational research*. Lanham, MD: Rowman and Littlefield.

Brewer, J., & Hunter, A. (1989). *Multimethod research: A synthesis of styles*. Newbury Park, CA: Sage.

Bryman, A. (1984). The debate about quantitative and qualitative research: A question of method or epistemology? *British Journal of Sociology, 35*, 78–92.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.

Constas, M. A. (1992). Qualitative data analysis as a public event: The documentation of category development procedures. *American Educational Research Journal, 29,* 253–266.

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed approaches*. Thousand Oaks, CA: Sage.

Creswell, J. W, Plano Clark, V. L., Guttmann, M. L., & Hanson, E. E. (2003). Advanced mixed methods research design. In A.Tashakkori and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, CA: Sage.

de Jong, H.L. (2003). Causal and functional explanations. *Theory and Psychology, 13*(3), 291–317.

de Waal, C. (2001). *On Peirce*. Belmont, CA: Wadsworth.

Dewey, J. (1948, 1920). *Reconstruction in philosophy*. Boston, MA: Beacon Press.

Dzurec, L. C., & Abraham, J. L. (1993). The nature of inquiry: Linking quantitative and qualitative research. *Advances in Nursing Science, 16,* 73–79.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11,* 255–274.

Guba, E. G. (1990). The alternative paradigm dialog. In E. G. Guba (Ed.), *The paradigm dialog* (pp. 17–27). Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Hoshmand, L.T. (2003). Can lessons of history and logical analysis ensure progress in psychological science? *Theory and Psychology, 13*, 39–44.

Howe, K. R. (1988). Against the quantitative-qualitative incompatability thesis, or, Dogmas die hard. *Educational Researcher, 17*, 10–16.

Howe, K. R. (1992). Getting over the quantitative-qualitative debate. *American Journal of Education, 100*, 236–256.

James, W. (1995, 1907). *Pragmatism*. New York: Dover.
Johnson, R. B., & Christensen, L. B. (2004). *Educational research: Quantitative, qualitative, and mixed approaches*. Boston, MA: Allyn and Bacon.
Johnson, R. B., Meeker, K. M., Loomis, E. J., & Onwuegbuzie, A. J. (2004, April). *Development of the philosophical and methodological beliefs inventory*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
Johnson, R. B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A.Tashakkori, and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297–319). Thousand Oaks, CA: Sage.
Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change*. Chicago, IL: University of Chicago Press.
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin, and Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 163–188). Thousand Oaks, CA: Sage.
Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A.Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 51–89). Thousand Oaks, CA: Sage.
Maxwell, J. A., & Loomis, D. M. (2003). Mixed methods design: An alternative approach. In A.Tashakkori, and C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 241–272). Thousand Oaks, CA: Sage.
Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data*. Mahwah, NJ: Lawrence Erlbaum.
Menand, L. (1997). *Pragmatism: A reader*. New York: Vintage.
Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135–164). Thousand Oaks, CA: Sage.
Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 3*, 362–376.
Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*, 120–123.
Murphy, J. P. (1990). *Pragmatism: From Peirce to Davidson*. Boulder, CO: Westview.
Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.
Newman, I., & Benz, C. R. (1998). *Qualitative-quantitative research methodology: Exploring the interactive continuum*. Carbondale, IL: Southern Illinois University Press.
Onwuegbuzie, A. J. (2000, November). *Validity and qualitative research: An Oxymoron?* Paper presented at the annual meeting of the Association for the Advancement of Educational Research (AAER), Ponte Vedra, Florida.
Onwuegbuzie, A. J. (2002). Positivists, post-positivists, post-structuralists, and post-modernists: Why can't we all get along? Towards a framework for unifying research paradigms. *Education, 122*(3), 518–530.
Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in the Schools, 10*, 71–90.
Onwuegbuzie, A. J., Jiao, Q. G., & Bostick, S. L. (2004). *Library anxiety: Theory, research, and applications*. Lanham, MD: Scarecrow Press.
Onwuegbuzie, A. J., & Johnson, R. B. (2004, April). *Validity issues in mixed methods research*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
Onwuegbuzie, A. J., & Leech, N. L. (2004a). *On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies*. Manuscript submitted for publication.
Onwuegbuzie, A. J., & Leech, N. L. (2004b, February). *Enhancing the interpretation of significant findings: The role of mixed methods research*. Paper presented at the annual meeting of the Eastern Educational Research Association, Clearwater, FL.
Onwuegbuzie, A. J., & Leech, N. L. (in press). Taking the "Q" out of research: Teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity: International Journal of Methodology*.
Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351–383). Thousand Oaks, CA: Sage.
Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
Peirce, C. S. (1878). How to make our ideas clear. *Popular Science Monthly, 12*, 286–302.
Phillips, D. C. (2004, April). *Is mixed methods research an epistemological oxymoron?* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. New York: Rowman & Littlefield.
Popper, K. R. (1959). *The logic of scientific discovery*. New York: Routledge.
Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative versus quantitative methods. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research* (pp. 7–32). Newbury Park, CA: Sage.
Reichardt, S. S., & Rallis, S. F. (1994). Qualitative and quantitative inquiries are not incompatible: A call for a new partnership. In C. S. Reichardt & S. F. Rallis (Eds.), *The qualitative-quantitative debate: New perspectives* (pp. 85–91). San Francisco, CA: Jossey-Bass.
Rescher, N. (2000). Realistic pragmatism: An introduction to pragmatic philosophy. Albany, NY: State University of New York.
Rorty, R. (2000). Pragmatism. *International Journal of Psychoanalysis, 81*, 819–823.
Sandelowski, M. (1986). The problem of rigor in qualitative research. *Advances in Nursing Science, 8*(3), 27–37.
Schrag, F. (1992). In defense of positivist research paradigms, *Educational Researcher, 21*(5), 5–8.
Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin & Y. S. Lincoln, *Handbook of qualitative research* (pp. 189–213). Thousand Oaks, CA: Sage.
Sechrest, L., & Sidana, S. (1995). Quantitative and qualitative methods: Is there an alternative? *Evaluation and Program Planning, 18*, 77–87.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
Sieber, S. D. (1973). The integration of fieldwork and survey methods. *American Journal of Sociology, 73*, 1335–1359.
Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher, 12*, 6–13.
Smith, J. K. (1984). The problem of criteria for judging interpretive inquiry. *Educational Evaluation and Policy Analysis, 6*, 379–391.
Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Applied Social Research Methods Series (Vol. 46). Thousand Oaks, CA: Sage.
Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: Sage.

Watson, W. (1990). Types of pluralism. *The Monist, 73*(3), 350–367.
Yin, R. K. (1984). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
Yu, C. H. (2003). *Misconceived relationships between logical positivism and quantitative research.* Research Methods Forum [On-line]. Retrieved September 2, 2004 from http://www.aom.pace.edu/rmd/2002forum.html.

AUTHORS

R. BURKE JOHNSON is a Professor, University of South Alabama, College of Education, BSET, 3700 UCOM, Mobile, AL 36688; bjohnson@usouthal.edu. His area of specialization is research methodology.

ANTHONY J. ONWUEGBUZIE is an Associate Professor, Department of Educational Measurement and Research, University of South Florida, 4202 East Fowler Avenue, EDU 162, Tampa, FL 33620-7750; tonyonwuegbuzie@aol.com. His areas of specialization are disadvantaged and under-served populations (e.g., minorities and juvenile delinquents) and methodological topics in the areas of quantitative, qualitative, and mixed methods.

Manuscript received October 14, 2003
Revisions received March 1 and April 30, 2004
Accepted May 12, 2004

2

(POST-)POSITIVISM, SOCIAL CONSTRUCTIONISM, CRITICAL REALISM: THREE REFERENCE POINTS IN THE PHILOSOPHY OF SCIENCE

In this chapter we will discuss three overarching philosophies of science: positivism and post-positivism, social constructionism, and finally, critical realism. We take up the three orientations as a conceptual, terminologic, and thematic general background to the qualitative methodologies that follow. All three cut across the quantitative/qualitative dividing-line. Although the main thrust of positivism is quantitative, there have been cases of qualitative positivism, for instance in historiography. Conversely, social constructionism is mainly qualitative, but quantitative social constructionist studies do exist. Finally, critical realism bridges quantitative and qualitative studies – there is no tendency for critical realists to favour either of these type of studies.

During the twentieth century, *positivism* became, and remained for a long time, the dominating philosophy of science. Theory and data, induction and deduction, law-like statements, verification and falsification, were key words. In the second half of the century, positivism came under increasing attack from internal sources – the post-positivists – as well as external opponents; and in the last third of the century, philosophical positivism rapidly deflated. Positivism has some similarities to the data-oriented methods discussed in Chapter 3, especially grounded theory; what is perhaps less well known is that it has also been alluded to by Foucault, and has some paradoxical traits in common with postmodernism; important ideas in post-positivism have been influential to postmodernist thought (see Chapter 6).

Social constructionism has increasingly emerged as an important perspective within social science and has even become predominant in some areas. Generally it can be said that for social constructionism, in contrast to positivism, reality is precisely *socially constructed*. (What this means in more detail, we will return to.) The important thing for research therefore becomes to explore how these social constructions happen. This approach is not particularly theory-oriented; the focus is rather on the 'disclosure' of how social phenomena are socially constructed. As we shall see, social constructionism is very rich and multi-faceted, so what has been said thus far is only a first indication of direction. Social constructionism has quite often been associated with postmodernism, and this may be true at a more superficial plane, although their roots and basic tenets are different; social constructionism has also made an inroad in grounded theory, and has been linked to hermeneutics and

critical theory (sometimes called 'critical hermeneutics') as well. Feminism often emphasizes gender as a social construction.

According to *critical realism*, both positivism and social constructionism are too superficial, unrealistic and anthropocentric. For social constructionism, all knowledge is linked to our social constructions and should not rise – at least not too high – above these. For positivism, all knowledge comes to us as single sense-data, and theories are just human-made linkages between these single data. Critical realism, in contrast, asserts that there is a world independent of human beings, and also that there are deep structures in this world that can be represented by scientific theories; the latter therefore become central for this orientation. Critical realism has been presented as a possible successor to social constructionism, but if this will transpire remains to be seen. In its emphasis on underlying patterns, critical realism shares some tangential points with hermeneutics and critical theory; in its searching for some kind of scientific laws, and in its view of the commonality of social science and natural science research, it shares ground with positivism.

Rooted in other traditions, social constructionism and critical realism constitute two important alternatives to positivist and post-positivist conceptions of science. In particular social constructionism but also critical realism presently draw great and increasing attention. They are often used as contrasts and as points of departure for debate and criticism.

In what follows, we present the orientations in chronological order. Initially launched in the nineteenth century, positivism was first out; social constructionism was introduced in the late 1960s; and critical realism in the 1970s. We shall give the most space to social constructionism since this is by far the most utilized orientation of the three in social science.

Positivism and beyond

The concept of 'positivism' has been central in the philosophy-of-science debate since the beginning of the nineteenth century, when Comte (1844) introduced the term, and through the twentieth century when logical positivism (later called logical empiricism) was topical. The sense of the positivism concept has often varied depending on who was doing the describing. The term 'positivism' has often been used in a derogatory sense, serving as a general invective. There is, though, a conceptual core. More concise and inclusive is perhaps Nietzsche's (1901: 267) description of the approach as the doctrine that 'halts at phenomena: "there are only *facts*"'. To which Nietzsche promptly retorts: 'No, facts is precisely what there is not, only interpretations'. A little more elaborately, Feyerabend (1981: 16) describes positivism as 'any interpretation of science (and of theoretical knowledge in general), which applies an assumption equivalent to' the statement by the well-known positivist Hempel, 'Science is ultimately intended to systematize data of our experience.'

Etymologically, the word positivism comes from the Latin *positum*,[1] the supine form of *pono*, put, set, place or lay. Thus, something is put, set, placed or laid; this something is given facts or data, and the one they lie in front of is the researcher.

Data are consequently something that *exists*, *is (already) there*, and the task of the researcher thus becomes to gather and systematize them. The underlying harvest metaphor is palpable. The researcher, as it were, collects the crops of the earth which are already there, and then prepares them into a tasty dish. Various positivist approaches have put different emphasis on these two processes, the gathering and the systematizing, and have also described them in different ways. For positivist historians in the nineteenth century, data collection was more important than systematization, a systematization that was never allowed to lead as far as to theory, since this would mean the abandonment of facts in favour of speculation. In contrast, for Comte and also for the logical positivists in the twentieth century, theory, the systematization of data, was central.

Current social science positivists focusing on statistical analysis are found somewhere in between these positions: theory, seen as a summing of data, is accepted, but the theoretical propositions are both less encompassing and less systematized than the logical positivists' prescriptions of universally valid, formalized axiomatic systems (prescriptions that the positivists' later inheritors in the philosophy of science have sharply criticized, see Suppe (2000)).

Data or facts should, according to positivism, be observable, and here is the link to empiricism (Harré, 1981). For modern positivism, what is observable also includes what is measurable or possible to register through some kind of instrument (Braithwaite, 1953: 8n). One approach within positivism, operationalism, even went as far as to reduce facts to measurable phenomena. A critical point against identifying observability with measurability is of course that this is all right when we talk about telescopes or microscopes; but even for these, a lot of interpretation beyond normal seeing is required. For other instruments, for instance a survey, the element of observation appears more distant or problematic. The logical positivists made a sharp distinction between theoretical language and observation language (reflecting the dichotomy between theory and empirical facts). The former was supposed to be translatable to the latter through so-called correspondence rules. As we shall see, this distinction was put in doubt by critics of positivism, who pointed out that all facts are theory-laden. If we talk about results of measurements, this already presupposes both theories about the instruments that measure and theoretical preconceptions of what we measure (otherwise we would not know what to measure). For surveys, for instance, statistical theory lies at their basis, and the variables that are part of the measurements presuppose various social-scientific theories. The correspondence rules were also criticized for being a 'heterogeneous confusion of meaning relationships, experimental design, measurement, and causal relationships, some of which are not properly part of theories', while on the other hand more vague or diluted interpretations were criticized for being logically inconsistent (Suppe, 2000: 103).

Critics of positivism

In the post-war era, the positivist approach, and particularly logical empiricism, long dominated the scientific-philosophical discussion in the Anglo-American sphere. From the end of the 1960s, however, positivism was the target of strong and growing

criticism, in particular from the Marxist left. The criticism did not diminish but rather increased in strength after the Marxist wave had ebbed away in the political disillusion of the 1970s. Structuralism, hermeneutics and phenomenology became the new banners under which many social scientists and humanists gathered. For the proponents of change, not as least important was the access to something of a fifth column in the very camp of the enemy: Kuhn's ideas of paradigms and paradigmatic revolutions – ideas that had emerged from within a positivist environment. Kuhn's contributions, however, have somewhat unfairly come to obscure other important authors in the post-positivist tradition. Names like Feyerabend (1975), Hanson (1958) and Toulmin (1953, 1961) deserve to be mentioned in this context. Kuhn, Feyerabend, Hanson and Toulmin have been given the umbrella term 'historical relativists' (Suppe, 1977), since they held that scientific knowledge is historically and socially conditioned, and so is not absolutely true but relative in character.

The 'heart' of logical empiricism stopped beating on the 26 March 1969 at the opening day of a symposium in Illinois, when one of its foremost standard-bearers, Carl Hempel, openly admitted that he no longer accepted the basic theses of this approach (Suppe, 2000). (This does not of course prevent several of these theses from surviving *post mortem* in the social sciences.) Kuhn himself, even though his ideas are very topical in social science, has become more or less superseded by later developments in the philosophy of science: post-Kuhnian critics of positivism have turned away from historic relativism and, for good or bad, instead tried to form more general, timeless principles for theoretical knowledge (Preston, 2004).

If there is any common feature to be found in the just mentioned various alternatives to positivism, it is the following. The purpose of scientific activity no longer stands out as a statistical putting together of surface phenomena in an observed reality. The important thing rather becomes to conceive this reality as an expression for, or a sign of, deeper-lying processes. For Hanson and Toulmin in the post-positivist school, the latter took the shape of law-like 'patterns' lying behind and explaining the manifestations of observed reality. The structuralists sought to trace structures that made their imprints on the matter of reality. The hermeneuticians interpreted the meanings that form the backdrop to and bring understanding of our language and actions. The Marxists and other dialecticians focused their interest on the hidden driving forces and mechanisms that in the form of in-built laws of movement, generated by contradictions, govern and develop the systems.

Thus, a transcending tendency is characteristic of the approaches critical of positivism: the observed reality is not all there is – and the researcher can reach behind it and reveal more fundamental layers, of which what we 'see' is a kind of projection or reflection. Such a way of looking at things was almost by definition excluded by the positivists, since (knowledge of) empirical reality was all that existed for them, and everything else was subjective constructions. Even scientific theories were conceived as complex of statements (systems of axioms) about generally observed relationships between surface phenomena. The parentheses around the words 'knowledge of' indicates the existence of differences in nuances between positivists in their attitudes to so-called theoretical entitities, i.e. what we have called underlying patterns or deep structures. On the one hand, there is the far-reaching

opinion that these do not *exist*; on the other hand there is the opinion that we cannot *know* whether they exist or not (McMullin, 1982: 19). In both cases the attitude leads to a resistance to the use of theoretical entities as part of the scientific process. One might argue that at least the former kind of positivists would accept the use of such entities as a heuristic aid to find their theories. The theoretical entities could thus serve as a kind of useful fiction for the construction of theories. As an extension of this line of reasoning we might ask what the difference is in practice between positivists and their critics here. It is difficult to see why a positivist should take the roundabout way via deep structures and underlying non-observable tendencies, when it is much simpler and less time-consuming just to summarize data, which is the goal anyhow. The critics on their side aim precisely at theories that are *not* compilations of data.

The positivists' reduction to that which is observable (or even stronger, measurable) in reality is, in the view of the critics, not very justified. If there are hidden patterns, underlying rule formations, which govern the observed parts of reality, and whose exploration can contribute to explaining these observed parts, then this seems to be a legitimate area for research. Rather than beginning with survey-based measurements of large amounts of empirical data, or with guesses of what the connections are between such data – both positivist approaches – another way becomes more reasonable: to carry out intensive studies of a small number of cases in order to retrieve through analysis the underlying patterns that are arguably reflected in the surface structures. These lines of thought have been held by post-positivists (Hanson, 1958; Toulmin, 1953), structuralists (Chomsky, 1968; Lévi-Strauss, 1962) and dialecticians (Marx, 1967).

Later orientations such as different forms of postmodernism and poststructuralism (see Chapter 6) to a certain extent hark back to positivism by rejecting the idea of deep structures or underlying patterns. There are only surfaces (which can perhaps 'fold' – the fold metaphor is common in postmodernism/poststructuralism). In, for example, the case of Foucault (1972), this look in the rear-view mirror happens explicitly, since he refers to himself as a 'positivist'. Even Latour (1996) can be mentioned in this context. A counter-reaction against surface thinking and a contention that underlying structures exist is found in the topical 'critical realism', which we discuss later in this chapter.

To take up the thread of postmodernism again, things are not that simple in this line of thought. As we shall see in Chapter 6, many proponents of postmodernism denies or brackets the existence of anything real outside language, to which linguistic statements would refer: texts only refer to other texts, not anything 'out there' (at least not accessible for research), as the jargon goes. If there is no extra-linguistic reality, there are also no hidden patterns to which the statements might refer. At the same time the postmodernists focused on tracing hidden but decisive cracks in the seemingly solid texts they studied: the so-called 'deconstruction'. The focus on the hidden behind the immediately familiar, palpable, is thus obvious even here.

Post-Kuhnian theoreticians (e.g. Suppe, 2000) in the influential 'semantic conception of science' have gone rather far in rejecting the idea of theories as direct reflections of reality. They introduce a third or middle term: *models*, which for them are more central than either theory or empirical data, and even constitute a kind of 'autonomous agents' (Morgan and Morrison, 2000). According to these 'model theoreticians', as

they have also been called (Chakravarty 2001), researchers never directly compare theory and empirical data, as the logical empiricists argued; they compare *one the one hand* theory with models and *on the other hand* models with empirical data. For this line of thought, theories are almost kinds of Platonic 'non-linguistic' entities soaring over models and empirical data (Suppe, 2000). That theories contain non-observables – references to entities that cannot be observed – becomes nearly self-evident from such a stance. (The view of theories as non-linguistic has been criticized, for instance, by Hendry and Psillos (2004) and Chakravarty (2001).)

Beside the semantic conception of science there is also another post-Kuhnian orientation worth mentioning, the 'evolutionary' conception, in which the formation of theories are seen as a kind of Darwinian natural selection – the false or less true theories are weeded out by a kind of natural selection process. Through its evolutionary aspect, this orientation retains the time perspective on the scientific generation of knowledge that characterized Kuhn and other historical relativists, but without the relativism, for example Kuhn's idea that later paradigms are hardly more true than those they replace. Evolution and natural selection are thought to favour (ever) truer theories.

Theory vs empirical 'facts': verification, falsification and beyond

The clash between verificationists and falsificationists (or Carnapians and Popperians) in the philosophy of science is well known. Equally well known are Kuhn's paradigms. Lakatos's research programmes and Feyerabend's methodological anarchism are also important ingredients in post-positivist thinking, as are Toulmin's ideas. Less well known outside the philosophy of science as an academic field, post-Kuhnian approaches deserve more attention from social scientists.

With Carnap (1962) in the forefront, the logical empiricists had asserted the necessity of 'verifying' theories and hypotheses with positive examples. As against this, Popper's (1934) ideas had a delayed but very strong impact, inverting everything and stating that theories, on the contrary, can only be 'corroborated' by repeated attempts at falsification. With his criticism, Popper directed a fundamental blow against logical empiricism, and theoretically refuted its most central theses.

Later critics, however, have had at least as strong objections to Popper's ideas. Kuhn (1970) showed in his analyses of scientific history that even the falsifications Popper had advocated as alternatives to the verifications he rejected, never occur in real research processes. He distinguished between two types of research. On the one hand the so-called 'normal research', where everything is focused on solving 'puzzles' within the frame of a thought structure – a 'paradigm' dogmatically regarded as given and the truth of which is beyond questioning. On the other hand the so-called 'paradigm shifts', scientific revolutions when one thought structure substitutes another, not on the basis of falsification, but for quite other reasons, including the capacity to attract supporters, who in time will come to dominate the research community.[2] Imre Lakatos (1970) tried to reconciliate falsificationism with Kuhn's ideas, using his concept of 'research programmes'. These are reminiscent of complicated systems of fortifications, the aim being to protect the 'hard core' of the theory by

different kinds of 'immunization strategies'. Such strategies were just what Popper had criticized. According to Feyerabend (1975), even great scientists sometimes act like a kind of skilful con man or Machiavellian politician, who are good at hoodwinking the general public and their peer researchers by manipulating and forging data in various ways, and using micro-political tactics. This, too, is of course contrary to Popper's ideas of falsification.

An even earlier critic of positivism than Kuhn, Feyerabend and Lakatos was Stephen Toulmin. Avoiding the sometimes irrational overtones of the three other 'historical relativists', Toulmin (1953) held that theories are entities that are *used*, rather than entities that are *tested*. According to Toulmin, theoretical propositions can be compared with descriptions of rules. For rules, we first delimit the general area – here called the *domain*, indicating the type of cases for which the rule is valid. We do not say that the rule must necessarily hold for all these cases. In this way we stake out the rule's *area of application*. At American universities, there can, for example, be a general prohibition to walk on the lawns (the domain) – but not for 'Fellows', who fall outside the area of application. To deprive other people of liberty is generally forbidden – but not if they have committed crimes that lead to prison, something which falls outside the area of application. To kill is generally prohibited – except in war, which falls outside the area of application. And so forth.

In theoretical propositions, which thus express a kind of rules according to which reality functions, we first delimit the general domain, that is the types of cases in which the rules are applicable, and then successively try to map out the area of application within the domain, or, in other words, the cases within the general type for which the rule really is valid, and the cases that are exceptions. (Toulmin gives many examples of how this is done in scientific practice.) Positivism, on the contrary, regarded (in its deductive-nomological model) theoretical propositions as simply statements of universal relations like the prime example 'all swans are white'. If the statement is valid, it is true, otherwise false. As we have seen, matters are not as simple as that. Toulmin shows in many ways how positivism has ended up beside the road of scientific practice by failing to distinguish between domain and area of application for theoretical propositions. In particular this has been the case with physics, even though the latter has been the ideal model for positivism. Instead of verification/falsification, a procedure that, as we have pointed out (Kuhn, Feyerabend) seldom occurs in scientific practice, this research strategy therefore involves something else: *the successive establishing of a theory's area of application within a certain domain*. Concerning a *rule*, nobody asks 'it is true or false', but '*when does it apply*?'

Is it possible to generalize beyond the empirical base? Generalization of qualitative case studies is often called into question or regarded as unfeasible, something which has been seen as a weakness compared to quantitative setups. However, this depends on the epistemologic point of view. It also depends on what is meant by generalization. If we only accept surface regularities, there is of course no reason why a pattern that has previously been established should hold true for more occasions. Only a statistical study that can establish the probability that the findings have not emerged by chance is then justified to make a generalization – with stated probability. In a perspective that accepts non-observables in the form of patterns

and tendencies, common to and underlying several surface phenomena, *successive expansions of the empirical area of application within a certain domain* are both possible and desirable, even in qualitative studies. (For concise arguments in favour of generalizing qualitative case studies on the basis of knowledge realism, see also Tsoukas (1989) and Danermark et al. (2002, Chapter 5).)

Many difficulties of the social sciences appear to be caused by importing a positivist view of how science 'should' be practised, a view that in its turn has been based on an erroneous picture of how natural sciences really work. This is true of the issue we are discussing here, as well as of the issues of theory legitimation vs theory generation and the structure of explanations. The battle will then be for or against this supposed natural-science picture of scientific practice, whereby a lot of 'anti' approaches will ensue. This polarization risks losing what is really common to various branches of science, despite different subject matters. It is to the credit of the philosophy of science that emerged after logical empirism – for instance in the shape of post-positivism (Hanson and Toulmin) – that it has increasingly focused on how real research processes happen, contrary to previous, more prescriptive approaches.

In other words, positivism, by prescribing a formal logical form for theoretical propositions (universal judgements) has simply, as Toulmin (1972 and 1974, passim) points out, followed its tradition to *identify the rational with the logical*. (Something which also manifests itself in other ways, for instance in the requirement of theories to be formalizable as axiomatized systems.) But everything that is rational does not need to be formally logical even though the reverse is the case. If the research process is ensnared in a formal logical straitjacket, there is a risk that the qualitative counter-currents which necessarily will follow as a reaction, in the most extreme case, take on irrational overtones.

Among post-Kuhnians, adherents of the 'semantic conception of science' have particularly strongly, as we have seen, maintained that theory is never compared directly with empirical data, but with models; and models with empirical reality. There is also a two-way traffic between, on the one hand, theory and models and, on the other hand, models and empirical data: theories can be adjusted if they do not fit the results of the model, or the model can be adjusted; models can also be revised if they do not correspond to empirical results, and new empirical results can be sought out for further checking if the current ones do not agree with the model (errors in instruments, registrations, etc. can occur).

This conception of science, however, is 'methodologically naturalistic', i.e. it presupposes natural science (and especially experimental physics) as the paragon for all science; the models should, for example, be mathematically formulated and the theory be expressible in terms of state (or phase) spaces (Suppe, 2000), whence the semantic conception of science seems less applicable to qualitative method in social science. McKelvey (2003) gives examples of why organization theory could be reformed from the semantic conception, and this seems rather quantitative. However, it should be kept in mind that the problematic is not unambiguous or simple; there is indeed qualitative mathematics (set theory, abstract algebra and topology are examples). The *basic* ideas in the semantic conception of science should also be applicable in qualitative research. What first comes to mind is of

course qualitative models, but other ideational artefacts should also be considered. Metaphors, analogies and narratives often seem to play a similar role in qualitative research as quantitative models in quantitative research, pointing out a way to appropriate the semantic conception of science *malgré lui* for qualitative purposes. Such artefacts, re-presentations of, on the one hand empirical 'reality', and, on the other hand, theory, should then be considered as entities in their own right, irreducible to and potentially more important to the research process than either of these two, yet playing a mediating role between them.

The semantic conception of science claims to be epistemologically neutral, that is to be compatible with both realism and non-realism. Thus, the individual researcher can adhere to one or the other, in line with personal preferences, without either of them being at variance with the semantic conception of science. In particular when it comes to realism, one advantage is said to be the avoidance of the problems relating to the language–reality linkage (the problematic of representation), since theories are said *not to be linguistic* (and language includes also mathematical language). The argument for scientific theories' 'non-linguistic' character is that a theory can be expressed in different languages, and therefore it must be something that, so to speak, lies behind language. The argument seems doubtful – were it true, fiction in literature, for instance, would not be linguistic either, since it can be translated (albeit not always perfectly). Moreover, the problematic of representation with language is not avoided, since theories must always be expressed in some language (Japanese or Mathematese or …) and the same goes for models, whose representations of reality, even if these are visual, must always be expressed in words (Chakravarty, 2001).

The semantic conception of truth seems, despite its asserted neutrality between a realist and a non-realist view, to have clear preferences for the former. Its compatibility with a non-realist view is open to question, since the very root to the conception is a correspondence between theory and reality, based on Tarski's correspondence theory for truth.

Social constructionism

Social constructionism has been associated with all the orientations that we take up in Chapters 3–6 below. Its roots are in phenomenology, but it has more recently been related to postmodernism. There are also attempts to launch a social constructionist grounded theory, and sometimes both critical theory and hermeneutics have been associated with social constructionism, which also shares tangential points with ethnomethodology and with Foucault. As social constructionism is thus a very broad and multi-faceted perspective and furthermore has often been contrasted to, compared to, and seen as an alternative to, on the one hand, positivism, on the other hand, critical realism, we treat it together with these two in the present chapter. For social constructionism, reality – or at least selected parts thereof – is not something naturally given. The study of how reality is socially constructed therefore becomes crucial for social constructionists.

The basic thrust of social constructionism can be described in the following four steps Hacking taken from (1999: 6, 12; we have changed Hacking's numbering):

1 In the present state of affairs X is taken for granted; X appears to be inevitable.

Social constructionist texts regularly begin with something that is regarded as self-evident, a taken-for-granted truth. The very point of social constructionism is then to prick a hole in this self-evidence by going further and showing that:

2 X need not have existed, or need not be at all as it is. X, or X as it is at present, is not determined by the nature of things; it is not inevitable.

This gives the 'aha experience' which is the main point of social constructionist texts. Many – though far from all – social constructionist texts then take one or two steps further, first to:

3 X is quite bad as it is.

And then to:

4 We would be much better off if X were done away with, or at least radically transformed.

Since social constructionism is so multi-faceted, these overarching characteristics are important in order to see, as it were, the forest and not just the single trees of the approach. (As to steps 3 and 4 above, it should be added that social constructionists are generally less systematic and confrontational in their criticism of societal phenomena than, for example, critical theorists (Chapter 5). Some try to be neutral or are only mildly sceptical. We will now look in more detail into Berger and Luckmann's classical work, which has become something of a cult book within the movement. After this we take up two more recent important authors: Gergen and Latour. This is followed by a presentation of the variety of social construction. Finally we present a few critical reflections and points of discussion.

Berger and Luckmann: reality as a social construction

Central author duo and pioneers for social constructionism,[3] Peter Berger and Thomas Luckmann in 1966 published their classical book *The Social Construction of Reality*. The main inspiration for Berger and Luckmann was phenomenological (see Chapter 3 below). It was mediated by the Austrian Alfred Schutz, who in the 1930s became strongly influenced by the father of phenomenology, Edmund Husserl. Schutz, fleeing from nazism, emigrated to the USA, and in his philosophy applied phenomenology to the common-sense world of everyday life. Berger and Luckmann developed this thinking in the area of sociology.

They were also influenced by other authors who have anticipated or been active within the area of knowledge sociology, such as Marx, Nietzsche, Scheler and Mannheim. All these, who from the present perspective can be regarded as a kind

of 'forefathers' to social constructionism, called into question the existence of a purely rational, objective knowledge, arguing instead that knowledge arises from processes more related to ideology, interests, or power.

Yet other sources of inspiration for Berger and Luckmann included sociological 'mega classics' like Durkheim, Weber and Mead. Durkheim presented as his basic rule that it was necessary to view 'social facts as things'. Weber, on the other hand, held that the subjective meaning content in social actions was the central issue. In their social constructionism, Berger and Luckmann wanted to join together these two polarized standpoints between objective macro relationships and subjective micro relationships (a polarity that has been something of a leading theme in much social science). Their solution laid the emphasis on the individual level and the social facts; the institutions remain secondary.

Berger and Luckmann held that:

> common-sense 'knowledge' rather than 'ideas' must be the central focus for the sociology of knowledge. It is precisely this 'knowledge' that constitutes the fabric of meanings without which no society could exist. The sociology of knowledge, therefore, must concern itself with the social construction of reality. (1966: 27)

Through this, the at first sight unsolvable conflict between Durkheim's and Weber's sociological position might be solved

> The central question for sociological theory can then be put as follows: How is it possible that subjective meanings *become* objective facticities? … How is it possible that human activity … should produce a world of things …? In other words, an adequate understanding of the 'reality sui generis' of society requires an inquiry into the manner in which this reality is constructed. (1966: 30)

The authors begin to solve this question by first 'attempt[ing] to clarify the foundations of knowledge in everyday life, to wit, the objectivations of subjective processes (and meanings) through which the *inter*subjective common-sense world is constructed' (1966: 34). The 'phenomenological analysis' they hereby use is termed 'descriptive' and 'empirical', not 'scientific' (1996: 34). All consciousness is intentional – i.e., it is always directed against some object. But these objects can present themselves for us in different spheres of reality, for instance in dreams or the waking state. We live in several different realities – among others in that of the dream – but the basic one is the usual everyday world. This is experienced as in various ways pre-structured, objectified. It also presents itself for us as an 'intersubjective world, a world [we] share with others.' (1966: 37). We continually remove ourselves from the everyday world to other, more secondary realities, or 'finite provinces of meaning' in Berger and Luckmann's terminology (taken from Schutz). These exist in unlimited numbers but a few important examples can be mentioned – theoretical science, art, religion, and the previously mentioned state of dreaming.

We share the everyday world with others. These others are experienced most characteristically in face-to-face situations. We '*typify*', according to Berger and Luckmann, (Part 1, Chapter 2) these others in various ways, for instance as English, Brasilian, man, woman, child, grown-up, professor, nurse, police officer, etc. Everyday

life is also filled with 'objectivations' (Part 1, Chapter 3). In these, a certain meaning content is given a material expression, which then becames more permanent and transcends the immediate, concrete face-to-face relationship. Signs, symbols and language are examples of such objectivations. (We return to the objectivation concept below.) Language is of course especially important in the building-up of 'a social stock of knowledge' (1966: 56). A prominent role in this social stock of knowledge is played by the build-up of routines for acting in various situations, something which functions as a sort of 'recipe knowledge' for actions (1966: 56).

Human beings differ from other animals in their less developed instinctual behaviour and in their great flexibility. In order not to become chaotic, human acting must therefore be confined by some form of stability. This happens through a 'social order'. Human beings are social in their nature, and Berger and Luckmann argue (inspired by the father of social interactionism, Herbert Mead) that even the experience of a 'self' is developed in meaningful interaction with others. The social order is thus a human product, or more specifically 'an ongoing human product'; it is not something inherent in the 'nature of things', nor does it express any 'natural law'. People alienate, or *externalize*, themselves by necessity in their actions, and the social order is an expression of this (1966: 69–70).

Central to the social order is the process through which institutions emerge – *institutionalization*. What, then, is an institution? Berger and Luckmann mean that the forming of habits and routines, 'habitualization', happens continually in human acting. All the time, we develop habits, through which a certain way of acting can be repeated in similar situations. Berger and Luckmann describe institutionalization as 'a reciprocal typification of habitualized actions. Put differently, any such typification is an institution' (1966: 72). In every institution, actions of a certain type are supposed to be carried out by a certain type of actor. For example, our legal system as an institution stipulates certain penalties for individuals above a certain age who are aware of the consequences of their actions and commit certain crimes. Academic institutions stipulate certain rules of admittance for certain types of actors (students) and conditions of employment for others (researchers, teachers, administrators). And so forth. Through institutionalization, we are subject to social control, for example the incest prohibition forbids certain kinds of sexual action. But this social control is already preceded by the typification whereby we define certain kinds of sexual action as incest and not others. This typification will of course vary between different cultures and societies.

We create within our social relations all the time new habits and routines in our actions, as well as new categories in our observing of others and their actions. Or in Berger and Luckmann's terminology, we habitualize and typify; these habitualizations and typifications – these habits, routines, and categorizations – spread between actors, and as they do this, institutions, that is fixed patterns of thought and action, emerge: institutionalization occurs, for instance in the shape of family, religion, legal systems, sports, school systems, health care, hunting, etc. These institutions, originally created by people, by and by begin to be perceived as something external, objective, and given, that is, there occurs also an *externalization* and an *objectivation* (1966: 78).[4] Berger and Luckmann here draw on Hegel's and Marx's concept of alienation, in

which people are viewed as estranged from their own activity, their praxis, which has been separated from them and therefore falsely comes to be understood as something external to themselves.

Through the need for meaningful mutuality in the social interaction, and for a coherent life – a 'biography' – there arises a need for coherence and unity, for integration, not only within but also between the institutions. Still, this integration, as different forms of 'institutional logic', is created by people, and is not the expression of any functionality or effectivity in the institutions themselves. Such institutional logics are the legitimizations of institutions, and a particularly important means for legitimizations is language. Whole 'bodies of knowledge' develop in this way, for instance theoretical formations in science, but there are also pre-theoretical bodies of knowledge that integrate knowledge in various areas on a pre-theoretical basis. The knowledge that is in this way alienated – externalized – from individuals will then be carried back to them, be internalized:

> Knowledge, in this sense, is at the heart of the fundamental dialectic of society. It 'programmes' the channels in which externalization produces an objective world. It objectifies this world through language and the cognitive apparatus based on language, that is, it orders it into objects to be apprehended as reality. It is internalized again *as* objectively valid truth in the course of socialization. Knowledge about society is thus a *realization* in the double sense of the word, in the sense of apprehending the objectivated social reality, and in the sense of ongoingly producing this reality. (1966: 83–84, note omitted)

This knowledge is then transmitted not only betwen the individuals in society at a certain time, but also over time between generations, which is how traditions arise.

More specifically, experiences and knowledge are stored as memory layers in and between individuals, or as Berger and Luckmann (1966: passim) say with a geological metaphor, inspired by Husserl, they are '*sedimented*'. Language, through its intersubjective transferring of meaning, is an important means for collective sedimentation. The transferring of institutional meaning is an important aspect of this. Knowledge of the sense and meaning of institutions is transferred by special typifications – for example teacher and pupil – and by special control apparatuses. Rituals and symbols of various kinds are used as carriers of institutional, sedimented knowledge, for example 'fetishes and military emblems' (1966: 88). The sediments of knowledge in a society are legitimized, but these legitimations can differ from time to time. For example, at one time prisons can be legitimized from their punishment function, at another time from their reforming function. Under a certain epoch universities can be legitimized from their educational function, under another epoch from their economic role in society.

In their typifications, individuals create different *roles* for themselves and others. Institutions cannot exist without being realized by human enactments in roles. Conversely, roles represent institutions. Institutions are also represented by many other things, like linguistic symbols, physical artefacts, and so on. But only human enactment in roles make the institutions, so to speak, come to life. 'The institution, with its assemblage of "programmed" actions, is like the unwritten libretto of a drama'. (1966: 92)[5] Roles are very important for the development of the individual's self,

since they are internalized and together will form a whole self, a subject. The roles further illustrate and mediate the basic dialectics between the institutional and the individual level of society. 'By playing roles, the individual participates in a social world. By internalizing these roles, the same world becomes subjectively real for him [sic!]' (1966: 91).

The extent of institutionalization can vary between different societies and times. Some societies are more or less pervaded by institutions; in others the institutionalization takes place mostly around a core. The institutionalization can also be segmented, so that, for instance, a certain institution is reserved for certain people or groups and is closed to outsiders. Cults are of course the extreme case of this. Institutions can also vary in the degree that they are '*reified*', that is are perceived as physical things. (Berger and Luckmann here took their inspiration from Marx and his concept of 'Verdinglichung', which would translate as something like 'thingification'.) The reification is described as an extreme case of objectivation, but it is not always easy to see the difference, for example when the authors describe overarching theoretical formations as reifications – which originators would see these as physical 'things', one wonders?

Legitimization constitutes another layer in the objectivation of meaning. It integrates disparate meanings to a connected whole. This takes place both at the level of the single individual's biography and at the level of institutions. Legitimization becomes necessary when meaning is to be mediated to new generations for which it is no longer self-evident. Explanations and justifications therefore become possible, and this is the process of legitimization. Legitimization is therefore both cognitive and normative. Four levels of legitimization can be discerned: The first and the most rudimentary level is built into language: our very vocabulary, the words we use, legitimize ('counter-terrorism' instead of 'oil war', 'subprime loans' instead of 'reckless loans', for instance). The second level consists of proverbs and sayings, maximes, legends, etc. The third level contains explicit theories. The fourth and most important level creates entire *symbolic universes*. Such a symbolic universe orders and integrates within its framework '*all* socially objectivated and subjectively real meanings' (1966: 114).

Thus, individuals create their reality, the institutions and their legitimizations, but this created reality in turn creates the individuals. This happens through socialization, the social influence through which individuals internalize social norms and knowledge. In the primary socialization, the child learns the basics of what is important in society, and in the secondary socialization, the process is fine-tuned for the grown-up. In the primary socialization, the child learns via 'significant others'. The identity is built up through role-taking – another term from social interactionism – we see one another with the eyes of significant others, reflect over this, and successively generalize the experiences.

In secondary socialization we appropriate 'sub-worlds', rather than 'base-worlds' as in the primary socialization (Berger and Luckmann, 1966: 158). The secondary socialization, for instance the school system, involves less of significant others, and is more formalized and abstract; the people included in this are often interchangeable (for instance teachers, as compared to parents). Yet, even here life partners, for instance, can take on the role as significant others, maintaining the person's subjective

reality, a reality that is always fragile and threatened by experiences that do not seem to fit in. More peripheral others function as a sort of 'choir' around the central, significant others. Another important means to maintaining a person's subjective reality is conversation. Through conversation with others, and perhaps above all through what is *not* said in conversation but is implied, we continually confirm our picture of reality. The need to maintain a subjective reality of course also means that it can change. The change can be continuous in the secondary socialization, when the present is interpreted in terms of the past. The change can also be radical, transforming, when instead the past is interpreted in terms of the present. Examples of such radical changes are said to be religious conversion, political brainwash and therapy.

Socialization can fail. One extreme case is stigmatized individuals: lepers, pedophiles, mentally ill people, etc., under various periods. Less conspicuous examples can occur because of discrepancies in the socialization. The discrepancy can take place between significant others, so that for instance father and mother or parents and nursemaid convey different messages. It can also take place between primary and secondary socialization, for instance between parents and teachers. If several discrepant worlds for socialization are accessible in a society, then this paves the way for individualism and relativism.

Berger and Luckmann see the human organism as a 'biological substratum', which sets limits to the individual's sociality – the need to feed, sexuality, death, etc. But the social world also sets limits to the individual's biology, in how we eat, how we have sex, when and how we die, etc. Socialization itself constitutes an ongoing invasion of the sociological world in the biological one, by regulating time and space, against the spontaneous tendencies or active resistance of the organism.

The authors conclude by saying that they view their contribution primarily as a re-definition of knowledge sociology. Beyond this, they hope that their book will eventually become an important complement to structural analyses in sociology. They do not want to deny the importance of these analyses, or maintain that social constructionism must always be a part of them; however, despite a certain ambivalence, they are not enthusiastic over macro sociological approaches like structuralism, functionalism and systems theories, which they see as always running the risk of reification. Berger and Luckmann maintain that the dialectic between society and individual that Marx already pointed out 'in fact and generally, does exist' (1966: 209), but that it is necessary to move on and develop this dialectic on the basis of sociological tradtion. In this work, they have sought to integrate other classical theories such as Durkheim's macro sociology, Weber's focus on individuals and understanding, and Mead's interactionist social psychology. (It should be added that this integration is done with a pervading phenomenological colour.) Finally, Berger and Luckmann hold that sociology is a humanistic discipline that must be carried on in a dialogue with philosophy and historiography.

Berger and Luckmann's book is very well written and its theses are unfolded with verve and enthusiasm. Its often suggestive terminology – 'finite provinces of meaning', 'symbolic universes' and so on – contribute to its rhetorical power. It integrates elements from major social-scientific classics such as Marx, Durkheim, Weber and

Mead. This integration is not just an eclectic putting together, but gains its unity through the philosophic canvas it is painted on, more precisely the phenomenology that was introduced by Husserl and further developed by Schutz in the area of everyday life. The book includes a whole spectrum of topical approaches in social science at the end of the 1960s. It is also a micro sociological protest against various macro sociological trends dominating at the time the book was written. It is an attempt to solve a basic problematic of social science: the contradiction between micro and macro, individual and society.

That said, several basic aspects of Berger and Luckmann's book are problematic and can be called into question. We will leave these critical comments until the end of the present section on social constructionism. For now, we turn to two newer social constructionist thinkers, stemming from different national cultures and academic disciplines – the American social psychologist Kenneth Gergen and the French sociologist of knowledge Bruno Latour.

Gergen: a persistent critic of positivism

Gergen (1978, 1996, 2004) strongly emphasizes the importance of language, and in this respect his thinking borders on that of the postmodernists. Influenced, among others, by philosophers like Gadamer, Kuhn and Rorty, Gergen has struggled for many years against the dominating positivist orientation of his discipline. According to Gergen, knowledge is never abstract, objective and absolute, but always concrete, situated and tied to human practice. There is no Truth, only local truths.

The important theories that have formed everyday thought and defined the problems of social science have, as Gergen (1982) shows, contained very little data, the most obvious example being Freud's works. These important theories have instead offered persuasive conceptions and ideas about central issues of life, often calling into question both prevailing assumptions and predominating values. They have often led to intense and long-lasting debates. This should not come as a surprise, since Kuhn has pointed to roughly similar conditions within the natural sciences. There are reasons why certain theories are accepted rather than others, but this is not just a question of facts. The extension and the use of certain theories, and even the results of these, is better explained by popularity cycles, boredom, career needs, and social and economic relationships. It is important for a theory to challenge established conceptions and question assumptions in previous theories to appear interesting and (reasonably) surprising, which is central to becoming influential (Davis, 1971). Empirical support is less important.

Gergen (1978, 1982), Rorty (1979) and many others have pointed out the insufficiency of theoretical claims for representation and of the hypothetical-deductive model as a way to think about the choice of theory. The various assumptions that are made – about the primacy of objective facts, the requirement of verification, the goal to reach universal atemporal results, and the impartial spectator – hide the nature and values of theories (Gergen, 1978, 1982). 'Facts' mean either the end or a suppression of a conflictual negotiation process which includes different interests and participants in the research process. A 'scientific procedure' often suppresses

the element of negotiation and prevents us from investigating its character. The testing of hypotheses becomes, to a certain extent, self-fulfilling over time since the theory shapes that towards which the attention is directed and people react interactively in testing situations. All results are historical artefacts, both because of the theory and because people, partly as a consequence of social-scientific reports, change over time. The question is only if we accept the conditions and practices that are required to create the scientific artefact. Every theory carries with itself the values of a research community which often lets its concepts and interpretations substitute for those that are lived by the research subjects.

Gergen seems to us to have fought so long – four decades! – against positivism in psychology that he, in contrast to Latour for instance, risks becoming a bit negatively dependent on it – something of an inverted mirror image. With some justification, he has been criticized for relativism, rather angrily by Ratner (2005) – a proponent of the older type of psychology that Gergen (2004) compares with dinosaurs on their way to extinction. (For a counter-reply to Ratner, see Zielke (2005)).

Gergen strongly advocates qualitative methods, which he believes have often been marginalized in favour of quantitative methods. He emphasizes the importance of a reflexive dialogue to set in motion hardened taken-for-granted assumptions which have emerged through collective processes of knowledge. Contrary to modernism, but like postmodernism, Gergen also very much emphasizes the instability and fragility of the human self, which he thinks has become particularly accentuated in our time of ever faster technological development and its influence on the individual. For Gergen (1989), this is a question of the ways in which human beings present their own (and others') inner self and gain credibility and legitimacy: 'What we take to be the dimension of self ... are symbolic resources for making claims in a sea of competing world construction' (1989: 75). Within more recent social constructionism, Gergen provides a contrast and represents something of a counter pole to Latour, to whom we now turn.

Latour and ANT: the 'second wave' of social constructionism

Within the sociology of knowledge, social constructionism has been particularly influential. Here, Bruno Latour with his investigations of scientific knowledge is incomparably the most known name. He has been responsible for what can be characterized as a second wave of social constructionism, in which also non-human actors such as technical artefacts and the like can play an active role in the construction. By using social constructionist lines of thought on natural-science activities, Latour has aroused fury in some quarters, especially among American scientists, and strongly contributed to the so-called 'science war' between, on the one hand French social constructionists and postmodernists, and on the other hand American natural-science realists, for whom the laws of nature are absolute, objective truths. A good early example of this is Latour and Woolgar's (1979) famous and controversial study of how knowledge is constructed socially in a laboratory. There are predecessors – already at the beginning of the 1930s, Ludwig Fleck (1934), who was a physician himself, investigated how medical knowledge is constructed; but Fleck's book is

long forgotten and it is Latour who made the great breakthrough. Latour followed *Laboratory Life* up with a book on the *Pasteurization of France* (1988), in which he showed that the reception and success of Louis Pasteur's pathbreaking ideas about micro organisms were socially constructed.

Latour has launched a very successful methodological programme, 'the actor-network theory' (ANT), where the actors do not only need to be humans. This idea is inspired by Greimas's (1983) semiotic theory of 'actants', a kind of generalized actors who do not need to be people but can be artefacts, etc. The idea of non-human actors can at first sight seem a bit fantastic, almost science fiction, but on a closer look it appears less bizarre. In the context of organizations, for example, most of us know that organizations, organizational subunits, groups, etc. take on something of a life of their own. The same is true of technical systems such as IT systems in organizations or plans and projects: they have a life of their own and are not mere passive products. Machines and devices of various kinds can also be actants. To take a simplified illustration: when we stop at a red traffic light, we are influenced by the device on the street corner, which is thus not only a passive receptor of impulses from human subjects in its initial construction but also in its turn influences human subjects. Actor-network theory is also sometimes called the 'actant-network theory'. For example, in Latour's (1996) book *Aramis* the main actor is a futurist personal rapid transit system which was later axed, and the question becomes, 'Who killed Aramis?'

In his extensive account of ANT, Latour (2005) strongly argues for a micro sociological 'bottom-up' perspective where the single actors, events and processual aspects play a decisive role. He rejects both macro sociology of the Durkheimian type (in which 'society' becomes a kind of metaphysical substance according to Latour) *and* postmodern deconstructionism. The latter is described as ruins built on ruins. Instead ANT aims at following the traces of associations between actants; associations that are always in the process of dissolving and re-emerging. Latour describes himself as a social constructionist, on the condition that the word 'social' should not be misunderstood as some sort of macro phenomenon which is already there, instead of being created at the micro level. Other authors related to ANT are Callon (1980; 1986), who seems to be the one who first published the idea, and Law (1994). As to method, ANT uses various ways of following the actor through interviews and ethnographic observation; there is also work with 'inscriptions', that is texts and the like (for instance data bases and graphical material). For current applications of ANT, see Czarniawska and Hernes (2005).

Compared to the perspectives of his first studies, such as *Laboratory Life*, Bruno Latour (2004a and b) has later developed in a (more) realist direction. He describes himself as being in permanent change and transformation, and provocatively refers to himself as a realist and a positivist[6] – even though he puts partly different meanings into this than the usual ones. The actants can be both human and non-human, and it would then seem strange to claim that the latter do not exist or can be 'reduced' to constructions. Moreover, reality is not neutral to operations on it, but resists, so its existence cannot be disregarded. However, reality is fluid, since the construction work continues all the time. Latour is extremely critical of postmodernism, seeing it as a sterile and destructive nihilism. But he is also critical of more theoretical ambitions, such as those of Bourdieu. Instead, it is crucial to keep to pure

descriptions of how the actants create their networks; not even explanations should enter the picture, since the research subjects know better than the researcher what goes on. The researcher must not get up on any high horses but must be humble and let the actants speak. Latour's approach is in this way reminiscent of the empirically oriented lines of thought which we take up in Chapter 3 – in particular ethnomethodology, for which he has sympathies.

Latour problematizes conventional interpretations and conceptions of both terms in 'social constructionism'. The 'social' does not mean that the constructions are *made* by, *consist* of, something social – whether that be insubstantial phantoms of the brain or manifestations of overarching societal power structures à la Bourdieu – but that the *construction process itself* is social, with several actants participating, co-constructing. However, also, the word construction is problematic at least as it has been conceived, and this has to do precisely with the participation of several, even non-human, actants. The constructor is not a god who blows his spirit into a material. The construction is real, no chimera, and what we have is, thus, a *realist constructionism*.

But this does not mean that there is a 'both-and' of construction and reality. Then we are just building further on a false opposition. Discourse and reality, 'words and worlds' are not dichotomic pairs but end points on a continuum consisting of practices carried out by (human and non-human) actants. Instead of getting stuck in this opposition, Latour holds, we should pose the question whether the construction is *good or bad*, something which is usually at the centre for questions concerning constructions in general, for example architectonic ones. (We might of course wonder: 'Good or bad *in relation to what*?') In any case, Latour's basic slogan, for which all other considerations must give way is: 'Follow the actants' (Latour, 1999). Latour plays with the idea of substituting 'composition' for '(social) construction'; but he inclines towards after all keeping the term construction since it is well established. Latour's reflection over, and problematization of, the concept and the term social constructionism thus results, after due deliberation, in his decision to keep it.

Generally, we find Latour's ideas exciting, for instance the view of artefacts as more active than is usually assumed. Yet, like some other more recent French thinkers, he seems to have a tendency to overdramatize his own lines of thought, in his case so that the artefacts are almost transformed into living entities and tend to assume a science-fiction character. A certain coquettishness with one's own position, a kind of hide-and-seek towards the reader, is another part of the style. Latour is also among the most antitheoretical of the social constructionists, which is based in his reductionist catchphrase about following the actants – everything else is to be rejected. As is the case with grounded theory, we wonder how it is possible to do research – or any mental activity – without theoretical preconceptions, and what the point is with pure descriptions. Books like *Aramis* can be fun to read – for a while. Then the amount of describing voices becomes a bit wearying.

The variation of social constructionism

Social constructionism has successively spread to most areas of social science and in many cases – where not everybody follows Latour or Gergen – has become more or

less dominant. An alliance, or convergence, with the postmodern orientation and Zeitgeist has contributed to its success – even though the coupling between these two lines of thought is not unproblematic, as we have seen. We will not try to list encyclopedically all conceivable authors in the social constructionist domain, since they are legion. Ian Hacking has playfully exemplified the host of social constructionist texts with an alphabetic sample. Thus, there are texts about The Social Construction of:

- Authorship
- Brotherhood
- The child viewer of television
- Danger
- Emotions
- Facts
- Gender
- Homosexual culture
- Illness
- Knowledge
- Literacy
- The medicalized immigrant
- Nature
- Oral history
- Postmodernism
- Quarks
- Reality
- Social homicide
- Technological systems
- Urban Schooling
- Vital statistics
- Youth homelessness
- Zulu nationalism

(Hacking, 1999: 1, references omitted)

Further examples can be entered ad lib for most letters of the alphabet. Hacking comments that he has not been able to find a title with 'The Social Construction of X', but that was before the era of Googling. We found 'The Social Construction of X-rated films' (Kurti, 1983). There is great variety not only in the empirical examples authors use but also in their theoretical approach. Sometimes all possible orientations that can have any point in common with social constructionism are included in the latter, such as deconstructionism, Foucauldianism, grounded theory, poststructuralism, discourse analysis, etc. (see, for instance, Burr, 2003). Most social scientists probably adhere to the idea that society and its institutions are not given, but in some (wide) sense socially created. In this way, most of us are social constructionists. We have, however, wanted to conceive social constructionism as a fairly delimited approach. Nonetheless, it is necessary to point out the considerable variation even within the rather diffuse core area(s) of social constructionism. We can discern – with an increasing degree of radicality – social construction as a critical perspective, a sociological theory, a theory of knowledge and a theory of

reality (Barlebo Wenneberg, 2001). This makes four degrees of radicality within social constructionism: a critical, a social, an epistemological and an ontological. The critical variant is the mildest, and means impugning the 'natural' in what has previously and commonly been regarded as self-evident and natural, and instead showing that this is socially constructed. Youth or race can serve as examples. It is obvious that conceptions of these vary a lot, and that they tend to create the phenomena in question. The social variant means arguing that society is in some sense produced and reproduced by shared meanings and conventions and thus socially constructed. The epistemologic variant means as the name indicates that knowledge is socially constructed. In the ontological variant, finally, reality itself is a social construction.

Barlebo Wenneberg now contends that these four degrees of social constructionism make a kind of inclined plane, where it is easy to slip or glide inconspicuously from the first, relatively innocuous position, all the way to the most radical position, that reality, including natural reality is nothing but a social construction. From a relatively trivial remark that certain phenomena do not occur naturally but are social creations, we are driven to reflect over how these constructions in their turn have emerged. We are then into some kind of social theory such as, for example, Berger and Luckmann elaborate in the later part of their book (1996). But a theory of society must to be consequent, must also tackle the issue of knowledge in society and how it is created. Then we have taken the step to the third variant of social constructionism, the epistemological one, in which knowledge is maintained to be a social construction. But if that is the case, it is a close step to start considering whether or not the object of knowledge, reality itself, is a social construction. Then we have taken the step to the fourth, ontological position, in which reality is a social construction.

*

Social constructionism is, as we have said, complex, varied and the overall picture is fragmented. One could probably say in today's use of social constructionism there is a shift of emphasis from the former to the latter poles within the three mentioned areas, that is from critical perspective to ontology, from cognition to language and from constructions of social phenomena to construction as a central aspect of research projects and claims. This coincides with an increased interest in postmodernism during recent decades, even though this interest has dropped during the last decade.[7]

Critique of social constructionism

Social constructionist texts and studies have the great merit of often being both fun to read and interesting, as well as exciting in their contents. They challenge common sense and not infrequently surprise the reader. Their value of attraction is usually higher than the mostly boring, not to say tinder-dry statistical investigations on the quantitative side, to which they have often been presented as the major qualitative alternative; but they are also more alluring than, for instance, texts in grounded theory generally are. This has probably contributed to the success of this orientation. But the success has its price, and we now come to our concluding critical points.

The Social Construction of Reality (Beger and Luckmann,1966), to start with the basic book of the movement, has, as we noted above, considerable merits and is a pioneering work, but critical reflections are possible at several points. To begin with, the central second part of the term 'social construction' remains not only undefined but unexplained. What is 'construction'? The authors never give any answer to this; the term is just abruptly introduced in the text, like many other suggestive terms of the book. Etymologically, the term comes from the Latin *con-struo*: staple, pile or order together, build. 'Construction' is a metaphor associated with planned activities such as those related to the erection of a building. The result of the building or construction process is an artefact, and the result of the social construction analogously would be a kind of social artefact, a social 'building' – an institution. Here, though, the metaphor limps, because social constructions, especially in the social constructionist sense, are not planned activities. According to Hacking (1999), the very point of social constructionism and that which has brought the approach such success is that it shows how various, seemingly 'natural' phenomena are not at all natural but social. This recurs time and again in Berger and Luckmann's book. But through the metaphor of 'construction' they actually take one step further: the terminology suggests that various natural phenomena are in fact not just social but intentionally planned, thus almost manipulatively created: they are, as it were, human *fabrications*, and the *disclosure* of these manipulatory or arbitrary fabrications becomes an important part of social constructionism's enticement.

It is also possible to question the very starting point of the book, which becomes something of an axiom for the authors, namely the primacy of the individual, the individual as the one where everything begins and ends. The basic problem for Berger and Luckmann was how it is possible that subjective meanings become objective facticities. From this point of departure, and using the construction metaphor, they go on to think that the study of society must focus on how individuals construct society. As we have seen, the metaphor is skewed. But also, the primacy of the individual is anything but a matter of course. The authors seem to experience it as self-evident, but it is not. How do we know, for example, that the individual is not secondary and that overarching structures are the primary constructors – or rather creators – of individuals in the first place? But this cannot possibly be so? Well, Fuchs (2001), among others, has presented an elaborate theory of how individuals and individuality are created – or as he says, 'constructed' – by networks. Networks are, in this theory, the place where everyting begins and ends. And there are other examples of approaches where individuals do not have primacy. For postmodernists and poststructuralists, individuals are created by texts or discourses; the text/discourse is thus primary. Individuals and cognitions are seen as the 'result' of language (which is also a main point for non-cognitivist constructionists like Gergen). For more recent (alethic) hermeneutics (see Chapter 4), the very situation of understanding is primary, and the individual is a result of this. As to the macro–micro problematic which has permeated so much of social science, many attempts have been made to solve it, both before and after Berger and Luckman. Among the most successful are Bourdieu's (1979) theory of practice and Giddens's (1984) theory of structuration. Both introduce a third element – a processual aspect which mediates between individual and society.

Even were we to accept the basic axiom of Berger and Luckmann that society and its institutions are built out of individual meanings, it does not follow that social science must necessarily be limited to studying how society is constructed. To take the analogy with houses: we can research into many other things than how the house was built, for example, architecture, strength of materials, plumbing, ventilation, price of the house, rents, how people behave in the house, etc.

A problem in any form of critical discussion of social constructionism more generally is that the orientation includes so many different positions, in between which advocates often move rhetorically under the pressure of various critiques and counter arguments (Barlebo Wenneberg, 2001), This is a common immunity strategy against critics. In addition there seems to be quite a lot of 'genuine' confusion and indecisiveness (Fleetwood, 2005). Correspondingly, there are many different versions of what people mean by 'construction' as well as 'social'. This means that any critique of social constructionism risks not 'sticking' as a consequence of the ambiguity and slipperiness of the target, and the critique may only be relevant for minor parts of the intended goal. What follows should therefore be read with some caution, as we also have problems in 'fixing' our object of critique.

Social constructionists hold that since social reality is a social construction, the only thing worth investigating is how this construction is carried out. This has profound consequences in that it leads to anti-theoretical tendencies (descriptivism and to a reduction to the individual level of analysis). But since knowledge is always theory-laden and we never 'observe' anything (including social constructions) without theories, this neglects the decisive role of our theories in research. Theories that tell us anything about social phenomena 'beyond' the construction of these, hardly become possible. Reflection over our theories, and the ensuing development of them, in order better to understand what we study, is an integrating part of research.

For this reason it can even be argued that theory is the most important aspect of research. Social constructionists tend, unfortunately, as Bourdieu has pointed out about micro sociologists generally, to stop where the real fun begins, instead of posing questions such as: 'Why do people construct society in the way they do?' and 'How do these constructions function, as patterns of social reality, once they have been constructed'?[8]

While some constructionists 'neutrally' and, at times, amusedly seek to point at construction processes, others often tend to put on a more sceptical perspective, regarding the patterns studied as something basically bad or evil, which we should not study as given (other than how they are *constructed*), just change or abolish. '[M]ost people who use the social construction idea enthusiastically want to criticize, change, or destroy some X that they dislike in the established order of things' (Hacking 1999: 7). The abolishment is rather easy, since the patterns do not truly exist but are, so to speak, only make-believe; that is they are constructions – hence a penchant for voluntarism.[9]

Social constructionists pursue, to a greater or lesser degree, a nominalist line of thought, according to which reality is amorphous, without qualities, and is only provided with arbitrary patterns by the researcher (see, for instance, Hacking, 1999).[10] This anti-realism, more or less adhered to, is self-destructive, implying as it does that

social constructionism itself is just an arbitrary pattern, invented by researchers. If everything is a social construction, then social constructionism is too, and there is no reason to believe in it, rather than any other taken-for-granted assumption. (See, for instance, Willmott's (1994b) criticism of the waverings and paradoxes this gives rise to, in his comments to Shotter and Gergen (1994).) The interesting thing, following social constructionism itself, would instead be to study how social constructionism has been socially constructed or how specific social constructionist studies construct others' constructions. For some reason, though, such studies are conspicuous by their absence.

For radical constructionists view the many more 'conventional' social constructionists as a target for critique, when these only focus on the social constructions of the society being investigated. That also the researchers in their knowledge could be seen as an example of social constructions in operation is hardly considered. Some critics here talk of 'trivial construction' (von Glasersfeld, 1991). Potter (1996: 13) suggests, for example, that Berger and Luckmann have a view on their own knowledge contributions as free from the studied citizens' social constructions, remarking that even though the authors spend considerable time considering the assumptions of the experienced reality of, for example, a car mechanic, Berger and Luckmann themselves seem to be able to look around the corner without any difficulties. Radical proponents of this direction are then also eager to indicate their own social constructioning. This increases consistency and awareness in the research approach but at the cost of reduced opportunities to say something of social phenomena 'as such', that is beyond the construction work of the researcher.

Social constructionists – in all variants – strongly object to what is called the 'essentialism' of other approaches. By essentialism is meant the opinion that various phenomena have some kind of immutable core of properties, their 'essence' (Latin *essentia*). As against this, social constructionists argue that everything is instead constructed. The question is, however, whether social constructionism itself does not adhere to such an essence, one that is marred by real existence and is not just constructed but is 'out there'. We are thinking of the *construction*. This is something which is said to be perpetually ongoing, and to be central to what happens; the mystical force behind the curtain, to allude to Hegel. Without a continually ongoing construction, no social constructionism. The construction has, in fact, obvious characteristics of a social constructionist 'essence', an inherent, unchangeable, constant property of our reality. The social constructionist criticism of essentialism thereby has a boomerang effect; its fundamental views become self-destroying here too. This essentialism has its roots in the phenomenology which is the ideational background to social constructionism. An important ingredient of phenomenology was the so-called 'intuition of essences' (*Wesensschau*) behind phenomena, and in a corresponding way, social constructionists try to intuit hidden constructions-essences behind social phenomena.

A possible counter-picture could be that construction of a specific phenomenon happens sometimes, but that for a major part of our time we do not indulge in construction of the world in terms of men and women, competence, leisure time or whatever it may be. On the other hand, one might always argue that it is typical of

essences precisely that they do not always come to expression, so in that sense, the counter-picture is not a counter-picture but rather a confirmation.[11]

The question is also if social phenomena, which are always dependent on mutual, subjective attributions of meaning, cannot have a real, objective existence. Must they, as (inter-)subjective phenomena be mere collective creations of our mental processes, chimera of our imagination … social constructions? Even if we agree that all social phenomena inevitably depend on mental processes and are infused with meaning, it is still possible to regard them as objectively, really, existing. For example, Searle has shown how there can 'be an objective reality that is what it is only because we think it is what it is' (1998: 113), through the 'collective attribution of status functions'. By this is meant that we, through collective intentions, assign to physical entities various symbolic functions. Take for instance money. When we have agreed that a certain type of paper slip, or a certain pattern of signs on a computer screen, represent money, we can heap all kinds of complex monetary functions on top of this. And we can – and do – link money to many further status functions, such as corporations, markets, governments, etc. This is a never-ending game, with real pieces from social life.

Critical realism

Critical realism, originated in writings by the philosopher Roy Bhaskar and in part inspired by Marx's view of science, has the ambition to be a more theoretical but also more realistic substitute for positivism and social constructionism in offering principles and ideas for science. Critical realists consider positivism and social constructionism as too superficial and non-theoretical in their way of doing research; analysis of underlying mechanisms and structures behind phenomena is what it takes to create theories that are not just concentrates of data. This orientation also has a radical vein: what is important is not just to explain the world but also to change it.

Overview

Critical realism is more and more often suggested as a counterweight and alternative to social constructionist ideas, and its increasing popularity can, to a large extent, be seen as a reaction against the spread of social constructionist and (overlapping with these) postmodernist ideas. Critical realism provides a ground for critique of social constructionism, in a time when positivism has lost its appeal for most scholars. Some see critical realism as an attractive alternative, at least to the more radical versions of social constructionism. Critical realism is sometimes used as a stick to beat what is taken to be the ambiguous, confused and imprecise mixture of standpoints characterizing social constructionism and postmodernist thinking (e.g. Fleetwood, 2005).

Although critical realism has received a certain international attention, it is still, primarily, a British tradition, The English philosopher Roy Bhaskar, who is considered

its founding father, has been developing the approach since the 1970s. It is intended to provide a philosophical grounding for science as well as an alternative to positivist and interpretive/constructionist approaches. The original target of critical realism was positivism, but nowadays more radical versions of social constructionism have become a main target of criticism, at least within the social sciences.

Critical realists stress the generalizing task of scientific activity. However, their stand is not to be confused with that of positivism, with its interest in predictable patterns. Instead, critical realism seeks to identify those deeper lying mechanisms which are taken to generate empirical phenomena. Bhaskar describes this as a shift from epistemology to ontology, and within ontology, as a shift from events to mechanisms. He thus turns against what he understands as misleading and antropocentric views, which give priority to epistemology, that is, questions concerning what and how we are able to know. Bhaskar refers to this as the 'epistemic fallacy'; by which he means the tendency to couple ontology and epistemology and to confuse that which exists with the knowledge we have about it (what we believe). These things should be kept separate, according to Bhaskar. Now, of course, science is a product of the social – moulded by a range of social, ideological and political conditions – 'but the mechanisms that it identifies operate prior to and independently of their discovery' (Bhaskar, 1998: xii).

The notion of reality as consisting of three domains – the empirical, the *actual* and the real – is a central one within critical realism. The empirical domain includes that which we can observe – things that happen and exist according to our immediate experience. The *actual* domain is a broader one, and refers to that which transpires independently of the researcher or any other observer who might record it. Finally, the domain of the real includes those mechanisms that are productive of different events and other 'surface phenomena'. According to critical realism, the task of science is to explore the realm of the real and how it relates to the other two domains. The empirical domain is more narrow and can be seen as a site of expression of the other two domains. 'Scientific work is instead to investigate and identify relationships and non-relationships, respectively, between what we experience, what actually happens, and the underlying mechanisms that produce the events in the world' (Danermark et al., 2002: 21).

It is the interest in mechanisms of a 'deeper dimension', which distinguishes critical realism from other traditions. It shares the interest of positivism in the objective world, patterns, generalization, and in finding causalities, but it also diverges from this tradition in claiming that the study of the observable is too superficial, as it disregards the unobservable mechanisms that produce the phenomena that positivists seek to measure and explain. It is not possible to reduce the world to observable objects and facts, critical realists argue. Moreover, they do not accept a distinction between theory and observation, nor the interest in finding all-encompassing laws. Instead critical realism takes an interest in complex networks of theoretical and observable elements characterizing efforts going beyond the surface of social phenomena. It shares with a great number of qualitative approaches an interest in synthesis and context, but it also strongly emphasizes the objective nature of reality, and it argues that a focus on social constructions is insufficient and misleading.

Indeed, most aspects of interest transpire beyond individuals' conception and definition of situations. Social structure entails things that lie behind individual consciousness and intention. In other words, causal mechanisms operate largely independently from the mind and action of individuals.

Critical realists emphasize strongly the reality as such, as distinct from our conceptions of it. They talk about an intransitive dimension – the object of scientific inquiry – and a transitive – that is our conceptions of that object. Many versions of social constructionism assume that such distinctions are artificial, and that societal phenomena are integrated with our conceptions of these, including those expressed by the research community, which contributes to the production of social reality. For critical realists, however, reality exists independently from researchers' ideas and descriptions of it.

The relevance of knowledge is dependent on the nature, power and mechanism of the objective reality. However, this is not to say that research, if it only has good intentions and methodology, is flawless or stands in an unproblematic relation to that which is researched.

> While it is evident that reality exists and is what it is, independently of our knowledge of it, it is also evident that the kind of knowledge that is produced depends on what problems we have and what questions we ask in relation to the world around us. (Danermark et al., 2002: 26)

Social constructions, while they are acknowledged to exist by critical realists, are framed in an objectivist manner, and are granted a rather limited role. Constructions are taken to be constructions of something, for example a discourse, a social practice, or physical reality, a reality that exists independent of how the constructions look like. The fact that it is socially defined and produced does not make a societal phenomenon any less real, critical realists argue. The way they see it, there *are* sellers, buyers, men, women, entrepreneurs, paid workers, carers, social outcasts, the unemployed, etc. Put differently, constructions are objective phenomena. A contrasting view, embraced by constructionists and many interpretive researchers, would be to approach constructions as volatile processes, which are then understood in terms of their subjective grounding. Focus is then placed, not on discourses or physical phenomena as such, but rather on the interpretation of these. Research is no exception here. According to this (constructionist) view, objects of knowledge are constructed by researchers through different procedures and tactics, not least discursive ones. To Bhaskar, however, the question is rather: What characteristics of societies make them possible as knowledge objects?

The real is central to critical realism. There is a strong conviction regarding the real and the possibility of identifying it. Something is real if it has a causal effect, that is, if it affects behaviour and makes a difference. Reality does not just consist of material objects. Ideas and discourses are real and can have causal effects. Ideas about, for example, race, men and age can explain patterns in the labour market, and they are real in the sense that they exist and work as mechanisms with causal effects. It is possible to identify at least four different types of realities; material, ideational, artefactual and social. A given entity can consist of several of these realities.

Artefactual reality, for example, refers to a synthesis of the physically, ideationally and socially real. The socially real 'refers to practices, states of affairs or entities for short, such as caring for children, becoming unemployed, the market mechanism, or social structures' (Fleetwood, 2005: 201). Social structure is used to capture configurations of causal mechanisms, rules, resources, powers, relations and practices. Causality refers to the nature of an object, which tells us what a certain object can or cannot do in terms of its effect. And a causal account 'does not deal with regularities between distinct objects and events (cause and effect), but with what an object is and the things it can do by virtue of its nature' (Danermark et al., 2002: 55). Objects have power connected to their structure; the mechanisms which produce effects are outcomes of this structure. But sometimes the effect of mechanisms does not show on the level of the empirical, that is, as an observable event.

Critical realism emphasizes the ideal and possibility of causal explanation. However, as noted earlier, the approach is still different from that of positivism, which seeks to establish predictable patterns and the exact relation between cause and effect. To critical realists relations are complex and causality can exist on different levels. They generate tendencies rather than inevitable, specific and measurable conditions. Critical realism examines the different mechanisms which have implications in terms of different effects and events, the forces and characteristics that mechanisms produce, and the intricate connections between different structural levels, that contribute to the complexity of causal forces, and that make possible the treatment of these as single, isolated factors. Causality should thus not be understood in terms of universal, predictable patterns, but rather as contextual and emergent, in changeable societies. According to critical realists, social reality is often slow in changing, but still emergent and varied as a consequence of the different processes that are part of producing it. As part of the project of accounting for typical patterns, while avoiding the misconception of statistical regularity and predictability, critical realists sometimes use the expression semi-regularity, which indicates 'the occasional, but less than universal, actualization of a mechanism or tendency, over [a] definite region of time-space' (Bhaskar and Lawson, 1998: 13).

The term mechanism is central within critical realism. A generative mechanism can be loosely defined as that which is capable of making things happen in the world. Mechanisms are taken to exist, even when they are not triggered (at work), or when their effect is impossible to trace, due to the effect of other mechanisms. Normally, mechanisms exist as part of complex compositions, whose outcome might vary or even fail to appear. Danermark et al. (2002) take the example of a match. It has the causal effect of being able to catch fire, if that mechanism is triggered, but for that to happen action has to be taken and in addition objects with other capacities must not intervene (e.g. by wetting the match). The same goes for social phenomena although it gets somewhat more complicated here, due to the dependency on human conceptions and actions (which of course, to some extent, goes for the match and its flaming capacity as well). To illustrate the point, Danermark et al. (2002) take the example of paid work. The structure of paid work is claimed to have the causal effect of forming the situation of people in our type of society, by making us reason and act in certain ways. It makes us want and apply

for jobs, and to acquire a suitable education, and it makes us go to work every working day. 'And each time acts in this way, the mechanism which reproduces the wage labour structure is triggered, which in turns generates new actions of the same kind, and so on' (2002: 56). At the same time there are other conditions that counteract the above mentioned mechanisms, for example the need to care for small children in the home, self-sufficiency, unsatisfactory work, or a football game on TV.

Supporters of critical realism look at the research process as a constant digging in the ontological depth of reality. In other words, reality is taken to be layered and research approaches which linger at a surface level are therefore discarded, be it social constructionism, hermeneutics, or positivism, which all depart from what is empirically given.

> In terms of the explanatory programme, the stratified nature of reality introduces a necessary historicity (however short the time period involved) for instead of *horizontal* explanations relating one experience, observable or event to another, the fact that these themselves are conditional upon antecedents, requires *vertical* explanations in terms of the generative relationships *indispensable for their* realization … (Archer, 1998: 196)

Critical realism distances itself from both methodological individualism (focus on the actor level) and holism (focus on the collective level), in emphasizing the social as relational and emergent. It is especially critical towards the former, arguing that 'actors' accounts are both corrigible and limited by the existence of unacknowledged conditions, unintendended consequences, tacit skills and unconscious motivations' (Bhaskar, 1998: xvi). Or, as Archer puts it 'we do not uncover real structures by interviewing people in-depth about them' (1998: 199).

According to critical realists the experiment has much to command it, in terms of its capacity to generate elementary knowledge. (This is similar to the 'semantic science approach' that was touched upon in the previous chapter.) Through experiments it is possible to isolate and identify mechanisms. However, given the fact that objects of study in social science can be considered open systems and, in addition, very complex, experiments are still not considered that relevant and useful in the study of social conditions. Bhaskar goes as far as to say that in open systems: 'positivism's instrumentalist-predictive-manipulative approach to phenomena is completely out of place' (1991: 141). Examples of more suitable alternatives are:

- Counterfactual thinking, through which one tries to imagine what could be – 'What would it be like if X did not exist?'
- Social experiments, for example anticipating reactions in breaking norms, an approach that is also favoured by ethnomethodologists (to be addressed in Chapter 3 below).
- The study of pathological or extreme cases.
- Comparative analysis of different cases (Danermark et al., 2002).

In general, however, critical realism does not engage with methodological matters much. It is a philosophy that cannot directly contribute to the disclosure of structures and mechanisms that produce and impact a certain, chosen, object of study. Still it is a philosophy for and not about science. It is generally prescriptive and it can support

research by offering an overall frame of reference and by 'affecting the questions put to reality, and the manner in which this is done' (Bhaskar and Lawson 1998: 7).

Critical realism does not deny the value of definitions of the social reality that is produced. Of course, social phenomena are acknowledged to be different from those studied in the natural sciences, but the active construction of social reality by individuals, and collectives thereof, is still downplayed. It is argued that the structures that guide the reproduction and transformation of social activities should be studied in their own right. In addition, a division between structure and agency is emphasized. These two aspects should be studied separately rather than together, as suggested by structuration and action theories. Structural impact, it is argued, mediates an objective influence and thus forms actions and provides actors with guidance. Structures are consequently taken to precede and determine actions, which in turn are seen as capable of gradually changing the former. Proponents of critical realism look at the sharp distinction between structure and human action as important to the analysis and enlargement of a space for action, which in turn is connected to the critical agenda of critical realism.

Although the centrality of such agendas can be discussed, it is often argued that critical realism encourages the transgression of existing social patterns by placing emphasis on the emergent theme. The approach can contribute to an ideology critique by going beyond common conceptions and by showing the workings of mechanisms, as well as the predefinition of our space for action, by structures. The indebtedness to Marxist thinking of critical realism shows clearly in its central concepts (underlying mechanisms, the level of articulation of reality), even if critical realism does not necessarily imply Marxist theory (some critical realists firmly reject connection to Marxism). According to Archer (1998: 203) critical realism 'has a "cutting edge" through identifying contextual constraints upon our freedoms and specifying strategic uses of our freedoms for social transformation'.

Critique

Below we will raise two areas within critical realism that we find particularly open to critique. The first is the strong claim by critical realism to grasp reality. It is always problematic to say something about causes which are not visible to us. The second area concerns central concepts which tend to be somewhat broad and diffuse, and do not really support the strong claims that critical realists are making. We will, however, on the whole refrain from going into the more philosophical critique that radical constructionists direct towards critical realists (e.g. Willmott, 2005), although inevitably we will touch upon it in discussing the other two themes.

Objectivism and exaggerated claims

In reading critical realism one is struck by the confidence with which its proponents use the concept of objective reality, as a point of departure and reference for the knowledge that is produced. It is argued that 'it is the nature of the object under study that determines what research methods are applicable, and also what knowledge

claims one may have' (Danermark et al., 2002: 70). And a central task of researchers is to identify 'the necessary, constituent properties of the study object, since these characteristics define what actions the object can produce' (2002: 70). It is almost as if the object of study discloses itself and then tells the researcher how it is most appropriately studied. So by defining the objective reality one is informed about what method to use and what outcomes in terms of knowledge can be expected. And by the same token, the correct identification of the 'necessary constitutive properties' will assist the researcher in figuring out the events that can be anticipated.

An alternative – and by far more realistic – approach departs from the notion that we can never describe the object as such, because we are always framed by our paradigmatic and methodological assumptions, a certain vocabulary and political stances (as we will discuss in the next chapter, preunderstanding and interpretation are also always central in research activities). To assume that the object of study, if appropriately defined, will direct the analysis, is a naïve conception, and to see the researcher as having privileged access to the object, seems pretentious. Different researchers have different views regarding the 'necessary constitutive properties' and even if one had the good fortune to find researchers sharing the assumption about such properties, they would most likely come up with different ideas on the nature of such properties, and they would probably also disagree over the events that the objects can be seen as capable of producing. Uses of different perspectives would probably lead to different properties and different produced objects.

To clarify our critique let us consider some of the accounts offered by critical realists that they argue illustrate their approach. Take the following statement for example: 'Structures divide the population – although seldom completely so – in to those with a positional interest in retaining and those with an interest in changing their structural location' (Danermark et al., 2003: 146). 'Beyond the main resource distributions there are relations behind ownership and those without property, those in power and those without, powerholders and the powerless, those discriminating and those discriminated against (2003: 148). According to this view it is the 'structure' that divides the population into this and that category. One could, however, argue that it is the author who does so. Let us explain. Without necessarily denying or even problematizing the idea of positional interests – although the concept of interest is not without its problems – it should be noted that the division of people, into those seeking change and those trying pre-empt it, is questionable in relation to the notion of a domain of the real (or the structure that is assumed to exist in it). Of course groups can be more or less arbitrarily defined, but still a group is never internally homogenous; there are always different or even conflicting interests within a given group, and, in addition, a certain individual ambivalence is surely to be expected. In effect, the division according to 'structure' hampers empirical diversity. Moreover, it could be argued that the notion of 'structure', as a grand divider in society, is in fact just another guise of the researcher, who, albeit often unconsciously, has a certain interest in dividing the population in this or that way. Thus, a more or less successful attempt to describe an observed empirical pattern is mistaken for a causal explanation of that same pattern.

Possibly this mistake also informs the second argument paraphrased above, regarding the distribution of resources. It is argued that behind the distribution of resources, there are relations of power (between those in possession and those lacking resources). These relations of power referred to are almost by definition about the distribution of resources, and consequently cannot be used to explain that very same distribution. Perhaps it is more accurate to depict the distribution of resources as being 'in front of' (rather than behind) the distribution. In other words, first there is a certain distribution of resources, and from this follows certain relations of power. But even so the division remains problematic. No doubt there are large economic differences in society, however, it is not clear how many people in today's Western societies (which are the main concern of the critical realists discussed here) can be considered completely lacking property and power.

Our point is thus that, unlike the arguments of critical realists, it is not so easy to assert the existence of structures, mechanisms, constitutive properties of objects of study, and so on.

Modest claims are not the trademark of critical realism. Bhaskar (1991) himself, for example, speaks of the necessity to reclaim reality which has been kidnapped by dangerous forces, led by skewed ideas. This reclamation should happen in two ways. Firstly, from 'philosophical ideologies which have usurped or denied it – reclamation in the sense of lost property'. Secondly, 'from the effects of those ideologies that have – like stagnant and muddy water – covered it up – reclamation in the sense of land reclamation' (Bhaskar, 1991: 144). Bhaskar finishes by stating that once reality has been reclaimed it should 'be used, nurtured and valued in an ecologically sustainable and humane way for human emancipation, happiness and flourishing' (1991: 144). It is not totally clear whether he is being ironic or not, but most likely that is not the case, because unlike, for example, social constructionists and postmodernists, critical realists are not known for their light-hearted and humorous forms of expression; to them the mission is much too important. Having said that it is interesting to note that the most loyal and orthodox critical realists, having observed the later Bhaskar becoming drawn towards spirituality, have half-jokingly suggested that he himself ought to get his membership to the critical realist club suspended.

The unproductive concepts of structure and mechanism

Structure and mechanism are two of the most central concepts within critical realism. In fact, the merits of the approach can, to a certain extent, be judged on the basis of how well these two concepts function within the research process. The notions of structure and mechanism are related. 'The objects have the power they have by virtue of their structure, and mechanisms exist and are what they are because of this structure' (Danermark et al., 2002: 55). Structure is thus the key to it all. The structure can mobilize force given the right input. So what is structure then?[12] Structure is a collection of internally related objects, such as teacher–student or employer–employee. With social structure, positions, practices and roles become associated. Social structure is that context in which actions and social interaction transpire.

At a first glance this image is quite appealing. Through observations we can surely spot indications of relations, roles and positions, we can grasp some of their implications. Indeed, language invites us to. Have you established that there are such things as teachers and students and that they are related in language use? But how much does this tell us? With the concept of structure, critical realism aims to do more than merely describe certain regularities and relations (conventionally, the concept of structure refers to such regularities). In fact, in their critique of positivism critical realists reject the importance of evident regularities and relations, but the latter is actually what Danermark et al. refer to in establishing the factual existence. But let us rest this discussion and move on to consider those 'forces' that the structure produces. If we exchange the term structure with its definition, as phrased by Danermark et al. (internally related object), then we can read that objects have forces as an outcome of internally related objects. Internally related objects can trigger forces.

Let us now consider this in less abstract terms, taking the teacher–student example from above. The teacher–student relation, that is the structure (internally related objects), should, according to the logic suggested by critical realists, be capable of activating forces and mechanisms. But the question is then, does this 'structure' trigger anything automatically, and if so, what would that be – ambitious students, subordination to authority, or maybe daydreaming related to lack of interest in what the teacher has to say, or perhaps attempts to disrupt tediousness through pranks, absenteeism or unruly behaviour in class, which in turn might cause the teacher to accumulate sick-leave days? Maybe what goes on in school between teachers and students is far more precarious, varied and processual, than the images suggested by objects, forces and mechanisms. In other words, the idea of Danermark et al. (2003) that the mechanisms exist as they are as an effect of this structure does not sit well with the variety and complexity of student–teacher relations that are likely to exist in many contemporary schools.

On paper and in an objectivist analysis, there are teachers and students, but a closer look may reveal a very different image. At a closer inspection we might find guards and troublemakers, part-time marketers with the objective of producing satisfied 'customers', and students that have adopted a customer perspective on education, or collaborators in a job creation scheme (with school as a tool for reducing/hiding unemployment) – that is, things that have less to do with learning and more to do with keeping young people off the streets. Perhaps teaching and learning are less central aspects of what actually goes on in many schools. In cases of inadequate teachers and very competent students the labelling, as we know it, might even verge on being misleading.

Surely critical realists can account for such conditions, and they do, through emergent structures and mechanisms. But the general stance shows a strong tendency to arrange the world in objective and sturdy categories. And while these refer to internally related objects, on the level of language, a closer and more open-minded look of what seems to be going on typically reveals a much more ambiguous view of the world. In fact, this is often the point with qualitative research; to transcend

seemingly objective definitions of the world, and to show their limited value in trying to explain what is going on.

The concept of mechanism also poses certain problems. The flaming capacity of matches is one thing, because this capacity is quite easy to relate to a mechanism. But what are the mechanisms that follow from structure, which determine the force of objects? Danermark et al. (2003) take the example of the organization and positions of paid work. It is argued that the structure of paid work has the causal power of forming the life conditions of people. If this structure is about the relation of employer–employed there are by definition certain elements in these that are highly unspecific and do not say much about specific conditions. The life situation of people is determined by factual conditions, such as legislation, attitudes, labour market conditions, specific relations between different people involved, the work organization, the machines, the organizational culture, managers, regional conditions, and the social security system, etc.

Proponents of critical realism would probably not deny that this is the case, and most likely they would attribute causal forces and mechanisms to all of these things. But is it reasonable to assume that a mechanism which reproduces the structure of paid work is triggered every time someone goes to work or applies for a job? Conceptualized this way, the mechanism metaphor becomes somewhat futile, that is, of course, if we disregard its, no doubt unintentional, comical value. 'I woke up this morning, feeling wretched not wanting to go to work. After some breakfast, with the feeling of agony still haunting me, I decided to activate (or the structure did so) the mechanism and reproduce the structure of paid work (= to go to work)'. It is our impression that critical realism uses the mechanism term in a much too literal sense. It appears to see the match example as a fairly good description of what is going on in social life. This point is, in our view, debatable.

Although critical realism acknowledges that social science, unlike natural science, mostly deals with open systems, a quite substantial part of the critical realist framework appears to be inspired by the world of physics. And when it leaves the realm of natural science and enters that of the social, the vocabulary of forces and mechanism does not work quite as well. Described in these terms, social phenomena come across as mechanical and often they run the risk of being overly simplified. They do not work quite as well in trying to explain complex matters, such as social relations in school or the organization of paid work.

To make the point even clearer, compare a recent comment made in an interview by Lisa Randall, one of the major names in theoretical physics at present, that it is simpler and easier to understand the universe than to understand the gender relations in a university organization.

Then again it is not always clear what mechanisms are, for instance as distinguished from underlying patterns (the latter are said to be revealed by abduction, the former via 'retroduction', and it is rather unclear what the difference really is). Not infrequently the hidden mechanisms in practical applications (for instance Danermark et al., 2003) become rather trivial and something that many positivists or social constructionists could well work with. The difference with, for example, Marx's more through-going analysis becomes evident.

Finally, it is our view that critical realists make too grand claims. They are utterly convinced about their approach to (what they take to be) objective science. Now, of course, they are aware of the precarious nature of research (as inevitably problematic and arguable), however, little space is granted to such discussions, apart from occasional confessions that come across as highly peripheral to what they otherwise consider themselves to be doing. In many ways critical realism expresses a view of self as a deep-digging project, which exposes reality and frees it from the filth and mess caused by positivists, postmodernists and their like. But as a noble reality saving project the approach runs the risk of becoming rigid and lacking in terms of reflexivity, presenting subjective and arbitrary representations as self-evident and robust findings.

Some of these things can probably be explained by the orientation's underdog position against more established competitors and the need to use rather heavy polemics and simplifications. As an alternative to various antitheoretical currents, critical realism all the same constitutes a stimulating and provocative counter-picture.

So far then, quite a lot of critical comments, however, there are also merits to the approach. One is its position as a clear alternative to positivist and constructionist approaches. As a result, options within research become clearer and proponents of other frameworks are forced to think through their own perspectives. For too long anti-positivism has constituted a sufficient argument for choosing a qualitative approach, however, with the emergence of critical realism social science is faced with yet another option. This framework also offers a relevant critique against research that refrains from leaving the surface level; approaches that never go outside or beyond the empirical, to analyse other aspects that affect this level. Critical realism thus challenges social constructionists and wants to stimulate researchers to rise above and beyond the empirical, to move on to more daring and theoretical analysis.

Brief comparison

Postivism, social constructionism, and critical realism thus diverge substantially. Still one could say that they are all interested in reality, real facts in the first case, the social reality in the second, and the objective 'big' world in the third: the *factual* world is put against the social world, and against this, the real, deep reality stands.

Many current debates frame social constructionism and critical realism as two main alternatives in outlining the development of social science. (Fleetwood, 2005; Willmott, 2005). Two objections can be made against this argument. First of all, social constructionism is a very broad field, and the range of different approaches that fall under this label includes quite diverse viewpoints. Second, the range and influence of critical realism is still quite limited outside of Great Britain. Ideas of real, objective structures are perhaps not entirely in line with the contemporary sceptical spirit of our time (unless, of course, this is about to change).

Some commentators emphasize the similarities when comparing the orientations. For example, positivism and critical realism both maintain that natural science and social science can use the same philosophy of science. Critical realism accepts a

constructionist element in research. There are streaks of realism within positivism, even though these are not predominant (its mainstream ontology is a subjective idealism – to regard data as subjective sense data and theories as instrumentalist summaries of data). Somebody might argue that the opposite of realism is not constructionism but idealism, in which the ideational is emphasized. Constructionism focuses on some phenomenon or object that is being constructed (Crotty, 1998). However, the construction itself is not a material thing, and neither, then, is the phenomenon or object under construction; they are both (inter)subjectively ideational. Delanty (2005) thinks that the main line of division is not between social constructionism and critical realism but between more radical and moderate variants of the former. Only extreme constructionists, such as the early Latour, deny the existence of underlying structures. We can of course establish various basic lines of division and options as to structures, constructions, interpretations, and so on. Extreme vs moderate social constructionism is here definitively a ground for drawing a main line of division. Yet, an even more fruitful distinction is probably that between a main or classical variant of social constructionism (exemplified by standpoints like those of Berger and Luckmann) and critical realism. The emphasis on (inter)subjective construction processes and outcomes, where socially defined and negotiated realities are central, generates rather different studies and understandings than critical realism's toning down of these in favour of an exploration of objective deep structures and mechanisms.

Final words

In this chapter we have considered positivism and post-positivism; social constructionism, a perspective that has become a dominating one within several disciplines of social science; and critical realism, a framework that competes with and has named itself as the leading alternative to and successor of the other two. This last claim remains to be realized. It can be mentioned, for example, that the approach is not very well known in the US. Social constructionism, on the other hand, is doubtless a broad framework and there are different varieties and connections, many of these have much in common with, for example, hermeneutics and post-modernism, as well as some more recent versions of grounded theory. We have also raised some doubts regarding the term social constructionism and its, over time, more and more opaque (over)use (see also Fleetwood, 2005). This does not stop social constructionist thinking from being central to the social sciences of today, and it is important to consider how one relates to this approach. Social researchers also have reasons to consider their relation to critical realism, and its reflections regarding the deep dimensions of knowledge and reality, as an incipient alternative to positivism and social constructionism.

*

So far in this first part of the book, we have focused the discussion on the concepts of surface structure, underlying patterns, perspectives, knowledge-sociological

conditions, relativism and dialogue. This is no coincidence; these concepts also point to the main emphases in the central chapters (3–6) of the book. Note that it is a question of emphasis; elements of the concepts appear in all these chapters. Even though the borderlines are of course fluid, the following generally is the case. In Chapter 3 about empirical orientations, empirical data and surface structures are at the forefront. Chapter 4 on hermeneutics contains an exploration of underlying patterns as its main aspect. The critical theory in Chapter 5 investigates various perspectives from a knowledge-sociological point of view, focusing on power and ideology. Postmodernism in Chapter 6, finally, takes up the problematic of relativism. Critical theory and postmodernism also emphasize, albeit each in different ways, the importance of dialogue. Critical theory strives for a rational dialogue, and to resist disturbances connected with power, ideology and dogma. Postmodernism emphasizes instead seminal clashes of meaning and disharmonies as a goal; in a wider perspective, the drift of this whole orientation is dialogue between texts, single individuals being seen as mere arenas for such dialogues.

Notes

1. See Heidegger's (1961) interpretation.
2. MacKenzie and House (1979) propose Popper's falsificationism with 'crucial experiments' as a 'paradigm' for social science in a Kuhnian spirit. This brave idea appears to put it mildly, neck-breaking, given that Kuhn among other things showed with his paradigm theory precisely that Popper's falsification method and 'crucial experiments' do not hold water.
3. There is a certain linguistic confusion about the concept itself; the terminology varies – sometimes the words 'social constructionism' are used, sometimes 'social constructivism'. We have chosen the former expression since 'constructivism' other orientations in, for example, mathematics and developmental psychology. In this, we join Kenneth Gergen (see below), who strongly advocates the use of 'constructionism'.
4. Berger and Luckmann here use the term 'objectivation ' in a somewhat different sense than before (see above), when it referred to material objects as carriers of meaning.
5. Here the authors are carried away by their verbal drive. A libretto is, of course, not the basis of a drama in general but of a musical performance – an opera, operetta, musical, etc.
6. In this context, it is interesting to note with Hacking (1999) the subterraenean connections between positivism and social constructionism. The main work of the central figure in logical positivism, Rudolf Carnap, has, for example, the title '*Der Logische Aufbau der Welt*', which would translate to '*The Social Construction of the World*'. This book has traits in common with the somewhat later ideas of one of the inspir of social constructionism – T.S. Kuhn. As Hacking writes: 'The roots of social constructionism are in the very logical positivism that so many present-day constructionists profess to detest' (1999: 42–43).
7. This tendency does not hold true for Latour's influential ANT approach, where language and researcher are given a less dominating place and postmodernism is totally dismissed.
8. A social constructionist might reply that the constructions are never finished, but are a continually ongoing process. This may be true, as it is true of house constructions (houses are repaired, altered etc.), but, as in the latter case, there are zones of relative stability constituting patterns for research.

9. The constructions being social and rooted in collective contexts can, however, be difficult to change (Czarniawska, 2005).
10. If the researcher provides the world with patterns, then what is there to abolish, we might ask. The answer is that also 'common' people, not only researchers, provide the world with patterns, and that it is these patterns – illusory or damaging ideas and ideologies – that are to be abolished. At a more philosophical level, the problematic harks back to Hegel and Marx with their thinking about alienation. The social world, which has originally been created by human beings, progressively becomes alien to them, and towers over them as an external threatening phenomenon. For Marx, the solution was, as is well known, to revolutionize this inauthentic world.
11. According to Latour, however, construction goes on all the time, and it has, as we have seen, real constructors as well as results (even though the latter are changing all the time, just like a city landscape).
12. The structure concept can also refer to 'small' structures, as in a workgroup structure, personality structure or emotional structure, but here we limit ourselves to social structures.

Diagnosing your Research Paradigm
Taken from: Collis and Hussey (2003)

Place a tick if you agree, or a cross if you disagree in the boxes on the right of the 10 statements to indicate if you agree or disagree with each statement. No right or wrong answers, and don't take it too seriously.		
1.	Quantitative data is more scientific than qualitative data.	
2.	It is important to state the hypothesis before data collection.	
3.	Surveys are probably the best way to investigate business issues.	
4.	Unless a phenomenon can be measured reliably, it cannot be investigated.	
5.	A good knowledge of statistics is essential for all approaches to business research.	
6.	Case studies should only be used as a pilot project before the main research is conducted.	
7.	Using participant observation to collect data is of little value in business research.	
8.	Laboratory experiments should be used more widely in business research.	
9.	It is impossible to generate theories during the course of research into business issues.	
10.	Researchers must remain objective and independent from the phenomena they are studying.	
Add the total number of ticks and crosses. If you have more ticks than crosses then you are likely to adopt a positivistic paradigm. If you have more crosses than ticks then you are more likely to adopt a phenomenological stance. If there are questions you are uncertain about, then you need to do more reading.		

REFERENCES

© Academy of Management Review, 1989, Vol. 14, No. 4, 532–550

Building Theories from Case Study Research

KATHLEEN M. EISENHARDT
Stanford University

This paper describes the process of inducting theory using case studies—from specifying the research questions to reaching closure. Some features of the process, such as problem definition and construct validation, are similar to hypothesis-testing research. Others, such as within-case analysis and replication logic, are unique to the inductive, case-oriented process. Overall, the process described here is highly iterative and tightly linked to data. This research approach is especially appropriate in new topic areas. The resultant theory is often novel, testable, and empirically valid. Finally, framebreaking insights, the tests of good theory (e.g., parsimony, logical coherence), and convincing grounding in the evidence are the key criteria for evaluating this type of research.

Development of theory is a central activity in organizational research. Traditionally, authors have developed theory by combining observations from previous literature, common sense, and experience. However, the tie to actual data has often been tenuous (Perrow, 1986; Pfeffer, 1982). Yet, as Glaser and Strauss (1967) argue, it is the intimate connection with empirical reality that permits the development of a testable, relevant, and valid theory.

This paper describes building theories from case studies. Several aspects of this process are discussed in the literature. For example, Glaser and Strauss (1967) detailed a comparative method for developing grounded theory, Yin (1981, 1984) described the design of case study research, and Miles and Huberman (1984) codified a series of procedures for analyzing qualitative data. However, confusion surrounds the distinctions among qualitative data, inductive logic, and case study research. Also, there is a lack of clarity about the process of actually building theory from cases, especially regarding the central inductive process and the role of literature. Glaser and Strauss (1967) and more recently Strauss (1987) have outlined pieces of the process, but theirs is a prescribed formula, and new ideas have emerged from methodologists (e.g., Yin, 1984; Miles & Huberman, 1984) and researchers conducting this type of research (e.g., Gersick, 1988; Harris & Sutton, 1986; Eisenhardt & Bourgeois, 1988). Also, it appears that no one has explicitly examined when this theory-building approach is likely to be fruitful and what its strengths and weaknesses may be.

This paper attempts to make two contributions to the literature. The first is a roadmap for building theories from case study research. This roadmap synthesizes previous work on qualitative methods (e.g., Miles & Huberman, 1984), the design of case study research (e.g., Yin, 1981,

1984), and grounded theory building (e.g., Glaser & Strauss, 1967) and extends that work in areas such as a priori specification of constructs, triangulation of multiple investigators, within-case and cross-case analyses, and the role of existing literature. The result is a more nearly complete roadmap for executing this type of research than has existed in the past. This framework is summarized in Table 1.

The second contribution is positioning theory building from case studies into the larger context of social science research. For example, the paper explores strengths and weaknesses of theory building from case studies, situations in which it

Table 1
Process of Building Theory from Case Study Research

Step	Activity	Reason
Getting Started	Definition of research question	Focuses efforts
	Possibly a priori constructs	Provides better grounding of construct measures
	Neither theory nor hypotheses	Retains theoretical flexibility
Selecting Cases	Specified population	Constrains extraneous variation and sharpens external validity
	Theoretical, not random, sampling	Focuses efforts on theoretically useful cases—i.e., those that replicate or extend theory by filling conceptual categories
Crafting Instruments and Protocols	Multiple data collection methods	Strengthens grounding of theory by triangulation of evidence
	Qualitative and quantitative data combined	Synergistic view of evidence
	Multiple investigators	Fosters divergent perspectives and strengthens grounding
Entering the Field	Overlap data collection and analysis, including field notes	Speeds analyses and reveals helpful adjustments to data collection
	Flexible and opportunistic data collection methods	Allows investigators to take advantage of emergent themes and unique case features
Analyzing Data	Within-case analysis	Gains familiarity with data and preliminary theory generation
	Cross-case pattern search using divergent techniques	Forces investigators to look beyond initial impressions and see evidence thru multiple lenses
Shaping Hypotheses	Iterative tabulation of evidence for each construct	Sharpens construct definition, validity, and measurability
	Replication, not sampling, logic across cases	Confirms, extends, and sharpens theory
	Search evidence for "why" behind relationships	Builds internal validity
Enfolding Literature	Comparison with conflicting literature	Builds internal validity, raises theoretical level, and sharpens construct definitions
	Comparison with similar literature	Sharpens generalizability, improves construct definition, and raises theoretical level
Reaching Closure	Theoretical saturation when possible	Ends process when marginal improvement becomes small

is an attractive research approach, and some guidelines for evaluating this type of research.

Background

Several pieces of the process of building theory from case study research have appeared in the literature. One is the work on grounded theory building by Glaser and Strauss (1967) and, more recently, Strauss (1987). These authors have detailed their comparative method for developing grounded theory. The method relies on continuous comparison of data and theory beginning with data collection. It emphasizes both the emergence of theoretical categories solely from evidence and an incremental approach to case selection and data gathering.

More recently, Yin (1981, 1984) has described the design of case study research. He has defined the case study as a research strategy, developed a typology of case study designs, and described the replication logic which is essential to multiple case analysis. His approach also stresses bringing the concerns of validity and reliability in experimental research design to the design of case study research.

Miles and Huberman (1984) have outlined specific techniques for analyzing qualitative data. Their ideas include a variety of devices such as tabular displays and graphs to manage and present qualitative data, without destroying the meaning of the data through intensive coding.

A number of active researchers also have undertaken their own variations and additions to the earlier methodological work (e.g., Gersick, 1988; Leonard-Barton, 1988; Harris & Sutton, 1986). Many of these authors acknowledge a debt to previous work, but they have also developed their own "homegrown" techniques for building theory from cases. For example, Sutton and Callahan (1987) pioneered a clever use of a resident devil's advocate, the Warwick group (Pettigrew, 1988) added triangulation of investigators, and my colleague and I (Bourgeois & Eisenhardt, 1988) developed cross-case analysis techniques.

Finally, the work of others such as Van Maanen (1988) on ethnography, Jick (1979) on triangulation of data types, and Mintzberg (1979) on direct research has provided additional pieces for a framework of building theory from case study research.

As a result, many pieces of the theory-building process are evident in the literature. Nevertheless, at the same time, there is substantial confusion about how to combine them, when to conduct this type of study, and how to evaluate it.

The Case Study Approach

The case study is a research strategy which focuses on understanding the dynamics present within single settings. Examples of case study research include Selznick's (1949) description of TVA, Allison's (1971) study of the Cuban missile crisis, and Pettigrew's (1973) research on decision making at a British retailer. Case studies can involve either single or multiple cases, and numerous levels of analysis (Yin, 1984). For example, Harris and Sutton (1986) studied 8 dying organizations, Bettenhausen and Murnighan (1986) focused on the emergence of norms in 19 laboratory groups, and Leonard-Barton (1988) tracked the progress of 10 innovation projects. Moreover, case studies can employ an embedded design, that is, multiple levels of analysis within a single study (Yin, 1984). For example, the Warwick study of competitiveness and strategic change within major U.K. corporations is conducted at two levels of analysis: industry and firm (Pettigrew, 1988), and the Mintzberg and Waters (1982) study of Steinberg's grocery empire examines multiple strategic changes within a single firm.

Case studies typically combine data collection methods such as archives, interviews, questionnaires, and observations. The evidence may be qualitative (e.g., words), quantitative (e.g.,

numbers), or both. For example, Sutton and Callahan (1987) rely exclusively on qualitative data in their study of bankruptcy in Silicon Valley, Mintzberg and McHugh (1985) use qualitative data supplemented by frequency counts in their work on the National Film Board of Canada, and Eisenhardt and Bourgeois (1988) combine quantitative data from questionnaires with qualitative evidence from interviews and observations.

Finally, case studies can be used to accomplish various aims: to provide description (Kidder, 1982), test theory (Pinfield, 1986; Anderson, 1983), or generate theory (e.g., Gersick, 1988; Harris & Sutton, 1986). The interest here is in this last aim, theory generation from case study ev-

Table 2
*Recent Examples of Inductive Case Study Research**

Study	Description of Cases	Research Problem	Data Sources	Investigators	Output
Burgelman (1983)	6 internal corporate ventures in 1 major corporation	Management of new ventures	Archives Interviews Some observation	Single investigator	Process model linking multiple organizational levels
Mintzberg & McHugh (1985)	1 National Film Board of Canada, 1939–1975, with 6 periods	Formulation of strategy in an adhocracy	Archives Some interviews	Research team	Strategy-making themes, "grass roots" model of strategy formation
Harris & Sutton (1986)	8 diverse organizations	Parting ceremonies during organizational death	Interviews Archives	Research team	Conceptual framework about the functions of parting ceremonies for displaced members
Eisenhardt & Bourgeois (1988)	8 microcomputer firms	Strategic decision making in high velocity environments	Interviews Questionnaires Archives Some observation	Research team Tandem interviews	Mid-range theory linking power, politics, and firm performance
Gersick (1988)	8 project groups with deadlines	Group development in project-teams	Observation Some interviews	Single investigator	Punctuated equilibrium model of group development
Leonard-Barton (1988)	10 technical innovations	Internal technology transfer	Interviews Experiment Observation	Single investigator	Process model
Pettigrew (1988)	1 high performing & 1 low performing firm in each of 4 industries	Strategic change & competitiveness	Interviews Archives Some observation	Research teams	In progress

* Examples were chosen from recent organizational writing to provide illustrations of the possible range of theory building from case studies.

idence. Table 2 summarizes some recent research using theory building from case studies.

Building Theory from Case Study Research

Getting Started

An initial definition of the research question, in at least broad terms, is important in building theory from case studies. Mintzberg (1979, p. 585) noted: "No matter how small our sample or what our interest, we have always tried to go into organizations with a well-defined focus—to collect specific kinds of data systematically." The rationale for defining the research question is the same as it is in hypothesis-testing research. Without a research focus, it is easy to become overwhelmed by the volume of data. For example, Pettigrew and colleagues (1988) defined their research question in terms of strategic change and competitiveness within large British corporations, and Leonard-Barton (1988) focused on technical innovation of feasible technologies. Such definition of a research question within a broad topic permitted these investigators to specify the kind of organization to be approached, and, once there, the kind of data to be gathered.

A priori specification of constructs can also help to shape the initial design of theory-building research. Although this type of specification is not common in theory-building studies to date, it is valuable because it permits researchers to measure constructs more accurately. If these constructs prove important as the study progresses, then researchers have a firmer empirical grounding for the emergent theory. For example, in a study of strategic decision making in top management teams, Bourgeois and Eisenhardt (1988) identified several potentially important constructs (e.g., conflict, power) from the literature on decision making. These constructs were explicitly measured in the interview protocol and questionnaires. When several of these constructs did emerge as related to the decision process, there were strong, triangulated measures on which to ground the emergent theory.

Although early identification of the research question and possible constructs is helpful, it is equally important to recognize that both are tentative in this type of research. No construct is guaranteed a place in the resultant theory, no matter how well it is measured. Also, the research question may shift during the research. At the extreme, some researchers (e.g., Gersick, 1988; Bettenhausen & Murnighan, 1986) have converted theory-testing research into theory-building research by taking advantage of serendipitous findings. In these studies, the research focus emerged after the data collection had begun. As Bettenhausen and Murnighan (1986, p. 352) wrote: ". . . we observed the outcomes of an experiment on group decision making and coalition formation. Our observations of the groups indicated that the unique character of each of the groups seemed to overwhelm our other manipulations." These authors proceeded to switch their research focus to a theory-building study of group norms.

Finally and most importantly, theory-building research is begun as close as possible to the ideal of no theory under consideration and no hypotheses to test. Admittedly, it is impossible to achieve this ideal of a clean theoretical slate. Nonetheless, attempting to approach this ideal is important because preordained theoretical perspectives or propositions may bias and limit the findings. Thus, investigators should formulate a research problem and possibly specify some potentially important variables, with some reference to extant literature. However, they should avoid thinking about specific relationships between variables and theories as much as possible, especially at the outset of the process.

Selecting Cases

Selection of cases is an important aspect of building theory from case studies. As in hypoth-

esis-testing research, the concept of a population is crucial, because the population defines the set of entities from which the research sample is to be drawn. Also, selection of an appropriate population controls extraneous variation and helps to define the limits for generalizing the findings.

The Warwick study of strategic change and competitiveness illustrates these ideas (Pettigrew, 1988). In this study, the researchers selected cases from a population of large British corporations in four market sectors. The selection of four specific markets allowed the researchers to control environmental variation, while the focus on large corporations constrained variation due to size differences among the firms. Thus, specification of this population reduced extraneous variation and clarified the domain of the findings as large corporations operating in specific types of environments.

However, the sampling of cases from the chosen population is unusual when building theory from case studies. Such research relies on theoretical sampling (i.e., cases are chosen for theoretical, not statistical, reasons, Glaser & Strauss, 1967). The cases may be chosen to replicate previous cases or extend emergent theory, or they may be chosen to fill theoretical categories and provide examples of polar types. While the cases may be chosen randomly, random selection is neither necessary, nor even preferable. As Pettigrew (1988) noted, given the limited number of cases which can usually be studied, it makes sense to choose cases such as extreme situations and polar types in which the process of interest is "transparently observable." Thus, the goal of theoretical sampling is to choose cases which are likely to replicate or extend the emergent theory. In contrast, traditional, within-experiment hypothesis-testing studies rely on statistical sampling, in which researchers randomly select the sample from the population. In this type of study, the goal of the sampling process is to obtain accurate statistical evidence on the distributions of variables within the population.

Several studies illustrate theoretical sampling. Harris and Sutton (1986), for example, were interested in the parting ceremonies of dying organizations. In order to build a model applicable across organization types, these researchers purposefully selected diverse organizations from a population of dying organizations. They chose eight organizations, filling each of four categories: private, dependent; private, independent; public, dependent; and public, independent. The sample was not random, but reflected the selection of specific cases to extend the theory to a broad range of organizations. Multiple cases within each category allowed findings to be replicated within categories. Gersick (1988) followed a similar strategy of diverse sampling in order to enhance the generalizability of her model of group development. In the Warwick study (Pettigrew, 1988), the investigators also followed a deliberate, theoretical sampling plan. Within each of four markets, they chose polar types: one case of clearly successful firm performance and one unsuccessful case. This sampling plan was designed to build theories of success and failure. Finally, the Eisenhardt and Bourgeois (1988) study of the politics of strategic decision making illustrates theoretical sampling during the course of research. A theory linking the centralization of power to the use of politics in top management teams was built and then extended to consider the effects of changing team composition by adding two cases, in which the executive teams changed, to the first six, in which there was no change. This tactic allowed the initial framework to be extended to include dynamic effects of changing team composition.

Crafting Instruments and Protocols

Theory-building researchers typically combine multiple data collection methods. While interviews, observations, and archival sources are particularly common, inductive researchers are not confined to these choices. Some investigators employ only some of these data collection methods (e.g., Gersick, 1988, used only obser-

vations for the first half of her study), or they may add others (e.g., Bettenhausen & Murnighan, 1986, used quantitative laboratory data). The rationale is the same as in hypothesis-testing research. That is, the triangulation made possible by multiple data collection methods provides stronger substantiation of constructs and hypotheses.

Of special note is the combining of qualitative with quantitative evidence. Although the terms *qualitative* and *case study* are often used interchangeably (e.g., Yin, 1981), case study research can involve qualitative data only, quantitative only, or both (Yin, 1984). Moreover, the combination of data types can be highly synergistic. Quantitative evidence can indicate relationships which may not be salient to the researcher. It also can keep researchers from being carried away by vivid, but false, impressions in qualitative data, and it can bolster findings when it corroborates those findings from qualitative evidence. The qualitative data are useful for understanding the rationale or theory underlying relationships revealed in the quantitative data or may suggest directly theory which can then be strengthened by quantitative support (Jick, 1979). Mintzberg (1979) described this synergy as follows:

> For while systematic data create the foundation for our theories, it is the anecdotal data that enable us to do the building. Theory building seems to require rich description, the richness that comes from anecdote. We uncover all kinds of relationships in our hard data, but it is only through the use of this soft data that we are able to explain them. (p. 587)

Also, of special note is the use of multiple investigators. Multiple investigators have two key advantages. First, they enhance the creative potential of the study. Team members often have complementary insights which add to the richness of the data, and their different perspectives increase the likelihood of capitalizing on any novel insights which may be in the data. Second, the convergence of observations from multiple investigators enhances confidence in the findings. Convergent perceptions add to the empirical grounding of the hypotheses, while conflicting perceptions keep the group from premature closure. Thus, the use of more investigators builds confidence in the findings and increases the likelihood of surprising findings.

One strategy for employing multiple investigators is to make the visits to case study sites in teams (e.g., Pettigrew, 1988). This allows the case to be viewed from the different perspectives of multiple observers. A variation on this tactic is to give individuals on the team unique roles, which increases the chances that investigators will view case evidence in divergent ways. For example, interviews can be conducted by two person teams, with one researcher handling the interview questions, while the other records notes and observations (e.g., Eisenhardt & Bourgeois, 1988). The interviewer has the perspective of personal interaction with the informant, while the notetaker retains a different, more distant view. Another tactic is to create multiple research teams, with teams being assigned to cover some case sites, but not others (e.g., Pettigrew, 1988). The rationale behind this tactic is that investigators who have not met the informants and have not become immersed in case details may bring a very different and possibly more objective eye to the evidence. An extreme form of this tactic is to keep some member or members of the research team out of the field altogether by exclusively assigning to them the role of resident devil's advocate (e.g., Sutton & Callahan, 1987).

Entering the Field

A striking feature of research to build theory from case studies is the frequent overlap of data analysis with data collection. For example, Glaser and Strauss (1967) argue for joint collection, coding, and analysis of data. While many researchers do not achieve this degree of overlap, most maintain some overlap.

Field notes, a running commentary to oneself and/or research team, are an important means of accomplishing this overlap. As described by

Van Maanen (1988), field notes are an ongoing stream-of-consciousness commentary about what is happening in the research, involving both observation and analysis—preferably separated from one another.

One key to useful field notes is to write down whatever impressions occur, that is, to react rather than to sift out what may seem important, because it is often difficult to know what will and will not be useful in the future. A second key to successful field notes is to push thinking in these notes by asking questions such as "What am I learning?" and "How does this case differ from the last?" For example, Burgelman (1983) kept extensive idea booklets to record his ongoing thoughts in a study of internal corporate venturing. These ideas can be cross-case comparisons, hunches about relationships, anecdotes, and informal observations. Team meetings, in which investigators share their thoughts and emergent ideas, are also useful devices for overlapping data collection and analysis.

Overlapping data analysis with data collection not only gives the researcher a head start in analysis but, more importantly, allows researchers to take advantage of flexible data collection. Indeed, a key feature of theory-building case research is the freedom to make adjustments during the data collection process. These adjustments can be the addition of cases to probe particular themes which emerge. Gersick (1988), for example, added several cases to her original set of student teams in order to more closely observe transition point behaviors among project teams. These transition point behaviors had unexpectedly proved interesting, and Gersick added cases in order to focus more closely on the transition period.

Additional adjustments can be made to data collection instruments, such as the addition of questions to an interview protocol or questions to a questionnaire (e.g., Harris & Sutton, 1986). These adjustments allow the researcher to probe emergent themes or to take advantage of special opportunities which may be present in a given situation. In other situations adjustments can include the addition of data sources in selected cases. For example, Sutton and Callahan (1987) added observational evidence for one case when the opportunity to attend creditors' meetings arose, and Burgelman (1983) added interviews with individuals whose importance became clear during data collection. Leonard-Barton (1988) went even further by adding several experiments to probe her emergent theory in a study of the implementation of technical innovations.

These alterations create an important question: Is it legitimate to alter and even add data collection methods during a study? For theory-building research, the answer is "yes," because investigators are trying to understand each case individually and in as much depth as is feasible. The goal is not to produce summary statistics about a set of observations. Thus, if a new data collection opportunity arises or if a new line of thinking emerges during the research, it makes sense to take advantage by altering data collection, if such an alteration is likely to better ground the theory or to provide new theoretical insight. This flexibility is not a license to be unsystematic. Rather, this flexibility is controlled opportunism in which researchers take advantage of the uniqueness of a specific case and the emergence of new themes to improve resultant theory.

Analyzing Within-Case Data

Analyzing data is the heart of building theory from case studies, but it is both the most difficult and the least codified part of the process. Since published studies generally describe research sites and data collection methods, but give little space to discussion of analysis, a huge chasm often separates data from conclusions. As Miles and Huberman (1984, p. 16) wrote: "One cannot ordinarily follow how a researcher got from 3600 pages of field notes to the final conclusions, sprinkled wth vivid quotes though they may be."

However, several key features of analysis can be identified.

One key step is within-case analysis. The importance of within-case analysis is driven by one of the realities of case study research: a staggering volume of data. As Pettigrew (1988) described, there is an ever-present danger of "death by data asphyxiation." For example, Mintzberg and McHugh (1985) examined over 2500 movies in their study of strategy making at the National Film Board of Canada—and that was only part of their evidence. The volume of data is all the more daunting because the research problem is often open-ended. Within-case analysis can help investigators cope with this deluge of data.

Within-case analysis typically involves detailed case study write-ups for each site. These write-ups are often simply pure descriptions, but they are central to the generation of insight (Gersick, 1988; Pettigrew, 1988) because they help researchers to cope early in the analysis process with the often enormous volume of data. However, there is no standard format for such analysis. Quinn (1980) developed teaching cases for each of the firms in his study of strategic decision making in six major corporations as a prelude to his theoretical work. Mintzberg and McHugh (1985) compiled a 383-page case history of the National Film Board of Canada. These authors coupled narrative description with extensive use of longitudinal graphs tracking revenue, film sponsorship, staffing, film subjects, and so on. Gersick (1988) prepared transcripts of team meetings. Leonard-Barton (1988) used tabular displays and graphs of information about each case. Abbott (1988) suggested using sequence analysis to organize longitudinal data. In fact, there are probably as many approaches as researchers. However, the overall idea is to become intimately familiar with each case as a stand-alone entity. This process allows the unique patterns of each case to emerge before investigators push to generalize patterns across cases. In addition, it gives investigators a rich familiarity with each case which, in turn, accelerates cross-case comparison.

Searching for Cross-Case Patterns

Coupled with within-case analysis is cross-case search for patterns. The tactics here are driven by the reality that people are notoriously poor processors of information. They leap to conclusions based on limited data (Kahneman & Tversky, 1973), they are overly influenced by the vividness (Nisbett & Ross, 1980) or by more elite respondents (Miles & Huberman, 1984), they ignore basic statistical properties (Kahneman & Tversky, 1973), or they sometimes inadvertently drop disconfirming evidence (Nisbett & Ross, 1980). The danger is that investigators reach premature and even false conclusions as a result of these information-processing biases. Thus, the key to good cross-case comparison is counteracting these tendencies by looking at the data in many divergent ways.

One tactic is to select categories or dimensions, and then to look for within-group similarities coupled with intergroup differences. Dimensions can be suggested by the research problem or by existing literature, or the researcher can simply choose some dimensions. For example, in a study of strategic decision making, Bourgeois and Eisenhardt (1988) sifted cases into various categories including founder-run vs. professional management, high vs. low performance, first vs. second generation product, and large vs. small size. Some categories such as size and product generation revealed no clear patterns, but others such as performance led to important patterns of within-group similarity and across-group differences. An extension of this tactic is to use a 2×2 or other cell design to compare several categories at once, or to move to a continuous measurement scale which permits graphing.

A second tactic is to select pairs of cases and then to list the similarities and differences between each pair. This tactic forces researchers to look for the subtle similarities and differences

between cases. The juxtaposition of seemingly similar cases by a researcher looking for differences can break simplistic frames. In the same way, the search for similarity in a seemingly different pair also can lead to more sophisticated understanding. The result of these forced comparisons can be new categories and concepts which the investigators did not anticipate. For example, Eisenhardt and Bourgeois (1988) found that CEO power differences dominated initial impressions across firms. However, this paired comparison process led the researchers to see that the speed of the decision process was equally important (Eisenhardt, in press). Finally, an extension of this tactic is to group cases into threes or fours for comparison.

A third strategy is to divide the data by data source. For example, one researcher combs observational data, while another reviews interviews, and still another works with questionnaire evidence. This tactic was used in the separation of the analyses of qualitative and quantitative data in a study of strategic decision making (Bourgeois & Eisenhardt, 1988; Eisenhardt & Bourgeois, 1988). This tactic exploits the unique insights possible from different types of data collection. When a pattern from one data source is corroborated by the evidence from another, the finding is stronger and better grounded. When evidence conflicts, the researcher can sometimes reconcile the evidence through deeper probing of the meaning of the differences. At other times, this conflict exposes a spurious or random pattern, or biased thinking in the analysis. A variation of this tactic is to split the data into groups of cases, focusing on one group of cases initially, while later focusing on the remaining cases. Gersick (1988) used this tactic in separating the analyses of the student group cases from her other cases.

Overall, the idea behind these cross-case searching tactics is to force investigators to go beyond initial impressions, especially through the use of structured and diverse lenses on the data. These tactics improve the likelihood of accurate and reliable theory, that is, a theory with a close fit with the data. Also, cross-case searching tactics enhance the probability that the investigators will capture the novel findings which may exist in the data.

Shaping Hypotheses

From the within-site analysis plus various cross-site tactics and overall impressions, tentative themes, concepts, and possibly even relationships between variables begin to emerge. The next step of this highly iterative process is to compare systematically the emergent frame with the evidence from each case in order to assess how well or poorly it fits with case data. The central idea is that researchers constantly compare theory and data—iterating toward a theory which closely fits the data. A close fit is important to building good theory because it takes advantage of the new insights possible from the data and yields an empirically valid theory.

One step in shaping hypotheses is the sharpening of constructs. This is a two-part process involving (1) refining the definition of the construct and (2) building evidence which measures the construct in each case. This occurs through constant comparison between data and constructs so that accumulating evidence from diverse sources converges on a single, well-defined construct. For example, in their study of stigma management in bankruptcy, Sutton and Callahan (1987) developed constructs which described the reaction of customers and other parties to the declaration of bankruptcy by the focal firms. The iterative process involved data from multiple sources: initial semi-structured telephone conversations; interviews with key informants including the firm's president, other executives, a major creditor, and a lawyer; U.S. Bankruptcy Court records; observation of a creditors' meeting; and secondary source material including newspaper and magazine articles and firm correspondence. The authors iterated between constructs and these data. They eventually developed definitions and measures for several constructs: disengagement, bargaining

for a more favorable exchange relationship, denigration via rumor, and reduction in the quality of participation.

This process is similar to developing a single construct measure from multiple indicators in hypothesis-testing research. That is, researchers use multiple sources of evidence to build construct measures, which define the construct and distinguish it from other constructs. In effect, the researcher is attempting to establish construct validity. The difference is that the construct, its definition, and measurement often emerge from the analysis process itself, rather than being specified a priori. A second difference is that no technique like factor analysis is available to collapse multiple indicators into a single construct measure. The reasons are that the indicators may vary across cases (i.e., not all cases may have all measures), and qualitative evidence (which is common in theory-building research) is difficult to collapse. Thus, many researchers rely on tables which summarize and tabulate the evidence underlying the construct (Miles & Huberman, 1984; Sutton & Callahan, 1987). For example, Table 3 is a tabular display of the evidence grounding the CEO power construct used by Eisenhardt and Bourgeois (1988), which included qualitative personality descriptions, quantitative scores from questionnaires, and quotation examples. The reasons for defining and building evidence for a construct apply in theory-building research just as they do in traditional, hypothesis-testing work. That is, careful construction of construct definitions and evidence produces the sharply defined, measurable constructs which are necessary for strong theory.

A second step in shaping hypotheses is verifying that the emergent relationships between constructs fit with the evidence in each case. Sometimes a relationship is confirmed by the case evidence, while other times it is revised, disconfirmed, or thrown out for insufficient evidence. This verification process is similar to that in traditional hypothesis testing research. The key difference is that each hypothesis is examined for each case, not for the aggregate cases. Thus, the underlying logic is replication, that is, the logic of treating a series of cases as a series of experiments with each case serving to confirm or disconfirm the hypotheses (Yin, 1984). Each case is analogous to an experiment, and multiple cases are analogous to multiple experiments. This contrasts with the sampling logic of traditional, within-experiment, hypothesis-testing research in which the aggregate relationships across the data points are tested using summary statistics such as F values (Yin, 1984).

In replication logic, cases which confirm emergent relationships enhance confidence in the validity of the relationships. Cases which disconfirm the relationships often can provide an opportunity to refine and extend the theory. For example, in the study of the politics of strategic decision making, Eisenhardt and Bourgeois (1988) found a case which did not fit with the proposition that political coalitions have stable memberships. Further examination of this disconfirming case indicated that the executive team in this case had been newly formed at the time of the study. This observation plus replication in another case led to a refinement in the emergent theory to indicate that increasing stabilization of coalitions occurs over time.

At this point, the qualitative data are particularly useful for understanding why or why not emergent relationships hold. When a relationship is supported, the qualitative data often provide a good understanding of the dynamics underlying the relationship, that is, the "why" of what is happening. This is crucial to the establishment of internal validity. Just as in hypothesis-testing research an apparent relationship may simply be a spurious correlation or may reflect the impact of some third variable on each of the other two. Therefore, it is important to discover the underlying theoretical reasons for why the relationship exists. This helps to establish the internal validity of the findings. For example, in her study of project groups, Gersick (1988) identified a midpoint transition in the lives of most project groups. She then used extensive quali-

Table 3
*Example of Tabulated Evidence for a Power Centralization Construct**

Firm	CEO Decision Description	CEO Power Score	CEO Power Distance[a]	CEO Dominated Functions	Story Decision Style[b]	Examples[c]
First	Strong Volatile Dogmatic	9.6	3.5	Mkt, R&D, Ops, Fin	Authoritarian	Geoff (Chairman) is THE decision maker. He runs the whole show. (VP, Marketing)
Alpha	Impatient Parental Tunes You Out	9.6	3.8	Mkt, R&D, Ops, Fin	Authoritarian	Thou shalt not hire w/o Presidential approval. Thou shalt not promote w/o Presidential approval. Thou shalt not explore new markets w/o Presidential approval. (VP, Operations)
Cowboy	Strong Power Boss Master Strategist	9.1	3.1	Mkt, R&D, Fin	Authoritarian Consensus	The tone of meetings would change depending upon whether he was in the room. If he'd leave the room, discussion would spread out, go off the wall. It got back on focus when he came back. (Director of Marketing)
Neutron	Organized Analytic	9.1	2.3	Mkt, Ops, Fin	Authoritarian	If there is a decision to make, I will make it. (President)
Omicron	Easy Going Easy to Work With	8.4	1.2	Fin	Consensus	Bill (prior CEO) was a suppressor of ideas. Jim is more open. (VP, Mfg.)
Promise	People-Oriented Pragmatic	8.9	1.3	Ops, Fin	Consensus	(My philosophy is) to make quick decisions involving as many people as possible. (President)
Forefront	Aggressive Team Player	8.3	1.2	None	Consensus	Art depends on picking good people and letting them operate. (VP, Sales)
Zap	Consensus-Style People-Oriented	7.5	0.3	Fin	Consultative	It's very open. We're successful most of the time in building consensus. (VP, Engineering)

[a] Difference between CEO power score and score of next most powerful executive.
[b] Authoritarian—Decisions made either by CEO alone or in consultation with only one person.
Consultative—Decisions made by CEO in consultation with either most of or all of the team.
Consensus—Decisions made by entire team in a group format.
[c] Individual in parentheses is the source of the quotation.
* Taken from Eisenhardt & Bourgeois, 1988.

tative data to understand the cognitive and motivational reasons why such abrupt and precisely timed transitions occur.

Overall, shaping hypotheses in theory-building research involves measuring constructs and verifying relationships. These processes are similar to traditional hypothesis-testing research. However, these processes are

more judgmental in theory-building research because researchers cannot apply statistical tests such as an F statistic. The research team must judge the strength and consistency of relationships within and across cases and also fully display the evidence and procedures when the findings are published, so that readers may apply their own standards.

Enfolding Literature

An essential feature of theory building is comparison of the emergent concepts, theory, or hypotheses with the extant literature. This involves asking what is this similar to, what does it contradict, and why. A key to this process is to consider a broad range of literature.

Examining literature which conflicts with the emergent theory is important for two reasons. First, if researchers ignore conflicting findings, then confidence in the findings is reduced. For example, readers may assume that the results are incorrect (a challenge to internal validity), or if correct, are idiosyncratic to the specific cases of the study (a challenge to generalizability). Second and perhaps more importantly, conflicting literature represents an opportunity. The juxtaposition of conflicting results forces researchers into a more creative, framebreaking mode of thinking than they might otherwise be able to achieve. The result can be deeper insight into both the emergent theory *and* the conflicting literature, as well as sharpening of the limits to generalizability of the focal research. For example, in their study of strategy making at the National Film Board of Canada, Mintzberg and McHugh (1985) noted conflicts between their findings for this highly creative organization and prior results at Volkswagenwerk and other sites. In the earlier studies, they observed differences in the patterns of strategic change whereby periods of convergence were long and periods of divergence were short and very abrupt. In contrast, the National Film Board exhibited a pattern of regular cycles of convergence and divergence, coupled with a long-term trend toward greater diversity. This and other conflicts allowed these researchers to establish the unique features of strategy making in an "adhocracy" in relief against "machine bureaucracies" and "entrepreneurial firms." The result was a sharper theory of strategy formation in all three types of organizations.

Similarly, in a study of politics, Eisenhardt and Bourgeois (1988) contrasted the finding that centralized power leads to politics with the previous finding that *decentralized* power creates politics. These conflicting findings forced the probing of both the evidence and conflicting research to discover the underlying reasons for the conflict. An underlying similarity in the apparently dissimilar situations was found. That is, both power extremes create a climate of frustration, which leads to an emphasis on self-interest and ultimately politics. In these extreme situations, the "structure of the game" becomes an interpersonal competition among the executives. In contrast, the research showed that an intermediate power distribution fosters a sense of personal efficacy among executives and ultimately collaboration, not politics, for the good of the entire group. This reconciliation integrated the conflicting findings into a single theoretical perspective, and raised the theoretical level and generalizability of the results.

Literature discussing similar findings is important as well because it ties together underlying similarities in phenomena normally not associated with each other. The result is often a theory with stronger internal validity, wider generalizability, and higher conceptual level. For example, in her study of technological innovation in a major computer corporation, Leonard-Barton (1988) related her findings on the mutual adaptation of technology and the host organization to similar findings in the education literature. In so doing, Leonard-Barton strengthened the confidence that her findings were valid and generalizable because others had similar findings in a very different context. Also, the tie to mutual adaptation processes in the education setting sharpened and enriched the conceptual level of the study.

Similarly, Gersick (1988) linked the sharp midpoint transition in project group development to the more general punctuated equilibrium phenomenon, to the literature on the adult midlife transition, and to strategic transitions within organizations. This linkage with a variety of literature in other contexts raises the readers' confidence that Gersick had observed a valid phenomenon within her small number of project teams. It also allowed her to elevate the conceptual level of her findings to the more fundamental level of punctuated equilibrium, and strengthen their likely generalizability to other project teams. Finally, Burgelman (1983) strengthened the theoretical scope and validity of his work by tying his results on the process of new venture development in a large corporation to the selection arguments of population ecology. The result again was a higher conceptual level for his findings and enhanced confidence in their validity.

Overall, tying the emergent theory to existing literature enhances the internal validity, generalizability, and theoretical level of theory building from case study research. While linking results to the literature is important in most research, it is particularly crucial in theory-building research because the findings often rest on a very limited number of cases. In this situation, any further corroboration of internal validity or generalizability is an important improvement.

Reaching Closure

Two issues are important in reaching closure: when to stop adding cases, and when to stop iterating between theory and data. In the first, ideally, researchers should stop adding cases when theoretical saturation is reached. (Theoretical saturation is simply the point at which incremental learning is minimal because the researchers are observing phenomena seen before, Glaser and Strauss, 1967.) This idea is quite similar to ending the revision of a manuscript when the incremental improvement in its quality is minimal. In practice, theoretical saturation often combines with pragmatic considerations such as time and money to dictate when case collection ends. In fact, it is not uncommon for researchers to plan the number of cases in advance. For example, the Warwick group planned their study of strategic change and competitiveness in British firms to include eight firms (Pettigrew, 1988). This kind of planning may be necessary because of the availability of resources and because time constraints force researchers to develop cases in parallel. Finally, while there is no ideal number of cases, a number between 4 and 10 cases usually works well. With fewer than 4 cases, it is often difficult to generate theory with much complexity, and its empirical grounding is likely to be unconvincing, unless the case has several mini-cases within it, as did the Mintzberg and McHugh study of the National Film Board of Canada. With more than 10 cases, it quickly becomes difficult to cope with the complexity and volume of the data.

In the second closure issue, when to stop iterating between theory and data, again, saturation is the key idea. That is, the iteration process stops when the incremental improvement to theory is minimal. The final product of building theory from case studies may be concepts (e.g., the Mintzberg and Waters, 1982, deliberate and emergent strategies), a conceptual framework (e.g., Harris & Sutton's, 1986, framework of bankruptcy), or propositions or possibly mid-range theory (e.g., Eisenhardt and Bourgeois's, 1988, midrange theory of politics in high velocity environments). On the downside, the final product may be disappointing. The research may simply replicate prior theory, or there may be no clear patterns within the data. The steps for building theory from case studies are summarized in Table 1.

Comparison with Other Literature

The process described here has similarities with the work of others. For example, I have drawn upon the ideas of theoretical sampling, theoretical saturation, and overlapped coding,

data collection, and analysis from Glaser and Strauss (1967). Also, the notions of case study design, replication logic, and concern for internal validity have been incorporated from Yin (1984). The tools of tabular display of evidence from Miles and Huberman (1984) were particularly helpful in the discussion of building evidence for constructs.

However, the process described here has important differences from previous work. First, it is focused on theory building from cases. In contrast, with the exception of Glaser and Strauss (1967), previous work was centered on other topics such as qualitative data analysis (e.g., Miles, 1979; Miles & Huberman, 1984; Kirk & Miller, 1986), case study design (Yin, 1981, 1984; McClintock et al., 1979), and ethnography (Van Maanen, 1988). To a large extent, Glaser and Strauss (1967) focused on defending building theory from cases, rather than on how actually to do it. Thus, while these previous writings provide pieces of the process, they do not provide (nor do they intend to provide) a framework for theory building from cases as developed here.

Second, the process described here contributes new ideas. For example, the process includes a priori specification of constructs, population specification, flexible instrumentation, multiple investigators, cross-case analysis tactics, and several uses of literature. Their inclusion plus their illustration using examples from research studies and comparison with traditional, hypothesis-testing research synthesizes, extends, and adds depth to existing views of theory-building research.

Third, particularly in comparison with Strauss (1987) and Van Maanen (1988), the process described here adopts a positivist view of research. That is, the process is directed toward the development of testable hypotheses and theory which are generalizable across settings. In contrast, authors like Strauss and Van Maanen are more concerned that a rich, complex description of the specific cases under study evolve and they appear less concerned with development of generalizable theory.

Discussion

The process of building theory from case study research is a strikingly iterative one. While an investigator may focus on one part of the process at a time, the process itself involves constant iteration backward and forward beween steps. For example, an investigator may move from cross-case comparison, back to redefinition of the research question, and out to the field to gather evidence on an additional case. Also, the process is alive with tension between divergence into new ways of understanding the data and convergence onto a single theoretical framework. For example, the process involves the use of multiple investigators and multiple data collection methods as well as a variety of cross-case searching tactics. Each of these tactics involves viewing evidence from diverse perspectives. However, the process also involves converging on construct definitions, measures, and a framework for structuring the findings. Finally, the process described here is intimately tied with empirical evidence.

Strengths of Theory Building from Cases

One strength of theory building from cases is its likelihood of generating novel theory. Creative insight often arises from the juxtaposition of contradictory or paradoxical evidence (Cameron & Quinn, 1988). As Bartunek (1988) argued, the process of reconciling these contradictions forces individuals to reframe perceptions into a new gestalt. Building theory from case studies centers directly on this kind of juxtaposition. That is, attempts to reconcile evidence across cases, types of data, and different investigators, and between cases and literature increase the likelihood of creative reframing into a new theoretical vision. Although a myth surrounding theory building from case studies is that the process is limited by investigators' preconceptions, in fact, just the opposite is true. This constant juxtaposition of conflicting realities tends to "unfreeze" thinking, and so the process has the potential to generate theory with less researcher bias than

theory built from incremental studies or armchair, axiomatic deduction.

A second strength is that the emergent theory is likely to be testable with constructs that can be readily measured and hypotheses that can be proven false. Measurable constructs are likely because they have already been measured during the theory-building process. The resulting hypotheses are likely to be verifiable for the same reason. That is, they have already undergone repeated verification during the theory-building process. In contrast, theory which is generated apart from direct evidence may have testability problems. For example, population ecology researchers borrowed the niche concept from biology. This construct has proven difficult to operationalize for many organizational researchers, other than its originators. One reason may be its obscure definition, which hampers measurability: ". . . that area in constraint space (the space whose dimensions are levels of resources, etc.) in which the population outcompetes all other local populations" (Hannan & Freeman, 1977, p. 947). One might ask: How do you measure an area in constraint space?

A third strength is that the resultant theory is likely to be empirically valid. The likelihood of valid theory is high because the theory-building process is so intimately tied with evidence that it is very likely that the resultant theory will be consistent with empirical observation. In well-executed theory-building research, investigators answer to the data from the beginning of the research. This closeness can lead to an intimate sense of things—"how they feel, smell, seem" (Mintzberg, 1979). This intimate interaction with actual evidence often produces theory which closely mirrors reality.

Weaknesses of Theory Building from Cases

However, some characteristics that lead to strengths in theory building from case studies also lead to weaknesses. For example, the intensive use of empirical evidence can yield theory which is overly complex. A hallmark of good theory is parsimony, but given the typically staggering volume of rich data, there is a temptation to build theory which tries to capture everything. The result can be theory which is very rich in detail, but lacks the simplicity of overall perspective. Theorists working from case data can lose their sense of proportion as they confront vivid, voluminous data. Since they lack quantitative gauges such as regression results or observations across multiple studies, they may be unable to assess which are the most important relationships and which are simply idiosyncratic to a particular case.

Another weakness is that building theory from cases may result in narrow and idiosyncratic theory. Case study theory building is a bottom up approach such that the specifics of data produce the generalizations of theory. The risks are that the theory describes a very idiosyncratic phenomenon or that the theorist is unable to raise the level of generality of the theory. Indeed, many of the grounded case studies mentioned earlier resulted in modest theories. For example, Gersick (1988) developed a model of group development for teams with project deadlines, Eisenhardt and Bourgeois (1988) developed a mid-range theory of politics in high velocity environments, and Burgelman (1983) proposed a model of new product ventures in large corporations. Such theories are likely to be testable, novel, and empirically valid, but they do lack the sweep of theories like resource dependence, population ecology, and transaction cost. They are essentially theories about specific phenomena. To their credit, many of these theorists tie into broader theoretical issues such as adaptation, punctuated equilibrium, and bounded rationality, but ultimately they are not theories about organization in any grand sense. Perhaps "grand" theory requires multiple studies—an accumulation of both theory-building and theory-testing empirical studies.

Applicability

When is it appropriate to conduct theory-building case study research? In normal science, theory is developed through incremental

empirical testing and extension (Kuhn, 1970). Thus, the theory-building process relies on past literature and empirical observation or experience as well as on the insight of the theorist to build incrementally more powerful theories. However, there are times when little is known about a phenomenon, current perspectives seem inadequate because they have little empirical substantiation, or they conflict with each other or common sense. Or, sometimes, serendipitous findings in a theory-testing study suggest the need for a new perspective. In these situations, theory building from case study research is particularly appropriate because theory building from case studies does not rely on previous literature or prior empirical evidence. Also, the conflict inherent in the process is likely to generate the kind of novel theory which is desirable when extant theory seems inadequate. For example, Van de Ven and Poole (in press) have argued that such an approach is especially useful for studying the new area of longitudinal change processes. In sum, building theory from case study research is most appropriate in the early stages of research on a topic or to provide freshness in perspective to an already researched topic.

Evaluation

How should theory-building research using case studies be evaluated? To begin, there is no generally accepted set of guidelines for the assessment of this type of research. However, several criteria seem appropriate. Assessment turns on whether the concepts, framework, or propositions that emerge from the process are "good theory." After all, the point of the process is to develop or at least begin to develop theory. Pfeffer (1982) suggested that good theory is parsimonious, testable, and logically coherent, and these criteria seem appropriate here. Thus, a strong theory-building study yields good theory (that is, parsimonious, testable, and logically coherent theory) which emerges at the end, not beginning, of the study.

Second, the assessment of theory-building research also depends upon empirical issues: strength of method and the evidence grounding the theory. Have the investigators followed a careful analytical procedure? Does the evidence support the theory? Have the investigators ruled out rival explanations? Just as in other empirical research, investigators should provide information on the sample, data collection procedures, and analysis. Also, they should display enough evidence for each construct to allow readers to make their own assessment of the fit with theory. While there are no concise measures such as correlation coefficients or F values, nonetheless thorough reporting of information should give confidence that the theory is valid. Overall, as in hypothesis testing, a strong theory-building study has a good, although not necessarily perfect, fit with the data.

Finally, strong theory-building research should result in new insights. Theory building which simply replicates past theory is, at best, a modest contribution. Replication is appropriate in theory-testing research, but in theory-building research, the goal is new theory. Thus, a strong theory-building study presents new, perhaps framebreaking, insights.

Conclusions

The purpose of this article is to describe the process of theory building from case studies. The process, outlined in Table 1, has features which range from selection of the research question to issues in reaching closure. Several conclusions emerge.

Theory developed from case study research is likely to have important strengths like novelty, testability, and empirical validity, which arise from the intimate linkage with empirical evidence. Second, given the strengths of this theory-building approach and its independence from prior literature or past empirical observation, it is particularly well-suited to new research

areas or research areas for which existing theory seems inadequate. This type of work is highly complementary to incremental theory building from normal science research. The former is useful in early stages of research on a topic or when a fresh perspective is needed, while the latter is useful in later stages of knowledge. Finally, several guidelines for assessing the quality of theory building from case studies have been suggested. Strong studies are those which present interesting or framebreaking theories which meet the tests of good theory or concept development (e.g., parsimony, testability, logical coherence) and are grounded in convincing evidence.

Most empirical studies lead from theory to data. Yet, the accumulation of knowledge involves a continual cycling between theory and data. Perhaps this article will stimulate some researchers to complete the cycle by conducting research that goes in the less common direction from data to theory, and equally important, perhaps it will help others become informed consumers of the results.

References

Abbott, A. (1988, September) *Workshop on sequence methods*. National Science Foundation Conference on Longitudinal Research Methods in Organizations, Austin.

Allison, G. (1971) *Essence of decision*. Boston: Little, Brown.

Anderson, P. (1983) Decision making by objection and the Cuban missile crisis. *Administrative Science Quarterly*, 28, 201–222.

Bartunek, J. (1988) The dynamics of personal and organizational reframing. In R. Quinn & K. Cameron (Eds.), *Paradox and transformation: Towards a theory of change in organization and management* (pp. 137–162). Cambridge, MA: Ballinger.

Bettenhausen, K., & Murnighan, J. K. (1986) The emergence of norms in competitive decision-making groups. *Administrative Science Quarterly*, 30, 350–372.

Bourgeois, L., & Eisenhardt, K. (1988) Strategic decision processes in high velocity environments: Four cases in the microcomputer industry. *Management Science*, 34, 816–835.

Burgelman, R. (1983) A process model of internal corporate venturing in a major diversified firm. *Administrative Science Quarterly*, 28, 223–244.

Cameron, K., & Quinn, R. (1988) Organizational paradox and transformation. In R. Quinn & K. Cameron (Eds.), *Paradox and transformation* (pp. 1–18). Cambridge, MA: Ballinger.

Eisenhardt, K. (in press) Making fast strategic decisions in high velocity environments. *Academy of Management Journal*.

Eisenhardt, K., & Bourgeois, L. J. (1988) Politics of strategic decision making in high velocity environments: Toward a mid-range theory. *Academy of Management Journal*, 31, 737–770.

Gersick, C. (1988) Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, 31, 9–41.

Glaser, B., & Strauss, A. (1967) *The discovery of grounded theory: Strategies of qualitative research*. London: Wiedenfeld and Nicholson.

Hannan, M., & Freeman, J. (1977) The population ecology of organizations. *American Journal of Sociology*, 82, 929–964.

Harris, S., & Sutton, R. (1986) Functions of parting ceremonies in dying organizations. *Academy of Management Journal*, 29, 5–30.

Jick, T. (1979) Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24, 602–611.

Kahneman, D., & Tversky, A. (1973) On the psychology of prediction. *Psychological Review*, 80, 237–251.

Kidder, T. (1982) *Soul of a new machine*. New York: Avon.

Kimberly, J. (1988) A review of Walter R. Nord and Sharon Tucker: Implementing routine and radical innovations. *Administrative Science Quarterly*, 33, 314–316.

Kirk, J., & Miller, M. (1986) *Reliability and validity in qualitative research*. Beverly Hills, CA: Sage.

Kuhn, T. (1970) *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Leonard-Barton, D. (1988) Synergistic design for case studies: Longitudinal single-site and replicated multiple-site. Paper presented at the National Science Foundation Conference on Longitudinal Research Methods in Organizations, Austin.

McClintock, C., Brannon, D., & Maynard-Moody, S. (1979)

Applying the logic of sample surveys to qualitative case studies: The case cluster method. *Administrative Science Quarterly*, 612–629.

Miles, M. (1979) Qualitative data as an attractive nuisance: The problem of analysis. *Administrative Science Quarterly*, 24, 590–601.

Miles, M., & Huberman, A. M. (1984) *Qualitative data analysis*. Beverly Hills, CA: Sage Publications.

Mintzberg, H. (1979) An emerging strategy of "direct" research. *Administrative Science Quarterly*, 24, 580–589.

Mintzberg, H., & McHugh, A. (1985) Strategy formation in an adhocracy. *Administrative Science Quarterly*, 30, 160–197.

Mintzberg, H., & Waters, J. (1982) Tracking strategy in an entrepreneurial firm. *Academy of Management Journal*, 25, 465–499.

Nisbett, R., & Ross, L. (1980) *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Perrow, C. (1986) *Complex organizations* (3rd ed.). New York: Random House.

Pettigrew, A. (1973) *The politics of organizational decision making*. London: Tavistock.

Pettigrew, A. (1988) Longitudinal field research on change: Theory and practice. Paper presented at the National Science Foundation Conference on Longitudinal Research Methods in Organizations, Austin.

Pfeffer, J. (1982) *Organizations and organization theory*. Marshfield, MA: Pitman.

Pinfield, L. (1986) A field evaluation of perspectives on organizational decision making. *Administrative Science Quarterly*, 31, 365–388.

Quinn, J. B. (1980) *Strategies for change*. Homewood, IL: Dow-Jones Irwin.

Selznick, P. (1949) *TVA and the grass roots*. Berkeley, CA: University of California Press.

Strauss, A. (1987) *Qualitative analysis for social scientists*. Cambridge, England: Cambridge University Press.

Sutton, R., & Callahan, A. (1987) The stigma of bankruptcy: Spoiled organizational image and its management. *Academy of Management Journal*, 30, 405–436.

Van de Ven, A., & Poole, M. S. (in press) Methods to develop a grounded theory of innovation processes in the Minnesota Innovation Research Program. *Organization Science*, 1.

Van Maanen, J. (1988) *Tales of the field: On writing ethnography*. Chicago: University of Chicago Press.

Yin, R. (1981) The case study crisis: Some answers. *Administrative Science Quarterly*, 26, 58–65.

Yin, R. (1984) *Case study research*. Beverly Hills, CA: Sage Publications.

Kathleen M. Eisenhardt (Ph.D., Stanford University) is Assistant Professor at Stanford University. Correspondence can be sent to her at the Department of Industrial Engineering and Engineering Management, 346 Terman Building, Stanford University, Stanford, California, 94305.

I appreciate the helpful comments of Paul Adler, Kennith Bettenhausen, Constance Gersick, James Frederickson, James Jucker, Deborah Myerson, Dorothy Leonard-Barton, Robert Sutton, and the participants in the Stanford NIMH Colloquium. I also benefitted from informal conversations with many participants at the National Science Foundation Conference on Longitudinal Research Methods in Organizations, Austin, 1988.

Questionnaire design.

For each decision you make when designing a questionnaire there is likely to be a list of points for and against just as there is for deciding on a questionnaire as the data gathering vehicle in the first place. Before designing the questionnaire the initial driver for its design has to be the research question, what are you trying to find out. After that is established you can address the issues of how best to do it.

An early decision will be to choose the method that your survey will be administered by, i.e. how it will you inflict it on your subjects. There are typically two underlying methods for conducting your survey; self-administered and interviewer administered.

A self-administered survey is more adaptable in some respects, it can be written e.g. a paper questionnaire or sent by mail, email, or conducted electronically on the internet.

Surveys administered by an interviewer can be done in person or over the phone, with the interviewer recording results on paper or directly onto a PC.

Deciding on which is the best for you will depend upon your question and the target population. For example, if questions are personal then self-administered surveys can be a good choice. Self-administered surveys reduce the chance of bias sneaking in via the interviewer but at the expense of having the interviewer available to explain the questions.

The hints and tips below about questionnaire design draw heavily on two excellent resources. SPSS Survey Tips, SPSS Inc (2008) and Guide to the Design of Questionnaires, The University of Leeds (1996).

The format of your questions will affect the answers;

Keep your questions short, less than twenty five words if possible. Keep questions understandable make sure the subject understands the terms used and importantly how the format of the questionnaire works (an already filled in example is often useful for this). Don't use “double negatives,” they can be confusing.

Choose appropriate question formats so they are understandable to the person answering and that enable you to analyse the resultant data.

Some questions can be easily answered with a simple single answer (e.g. do you smoke (y/n); what gender are you? (m/f), but others may require multiple choices a scale or, perhaps even a grid. Do make sure you know how to analyse the data you get, if you can't analyse the resulting data there was little point in collecting it. A research proposal should address analysis, a simple sentence "data will be analysed using SPSS" may pass the buck to SPSS but won't help much when you refer back to your plan. You should have an eye on the analysis when designing the questionnaire. Checking this is feasible should be part of the piloting; this will check that the data are arrangeable in the formats needed for analysis and that you have the resources to do it.

You might include open ended questions in the questionnaire, do though be aware that they will be "tainted" by the context of being in with strictly quantitative questions. The pilot is a good time to use more open questions to check there are sufficient options on multi choice answers and that there is sufficient discrimination in the questions, so not all the answers are the same when there is likely to be a range of views/responses.

Ambiguous questions.

Check for ambiguity in your questions, make sure what you're asking is obvious. Ambiguous questions not only yield no useful data but can frustrate the respondent and encourage them to give up! Avoid asking two questions at once. For example, “Are you happy with the amount and timeliness of feedback you receive from your tutors?” Analyzing the responses to such a question would be made practically impossible because you won’t be able to tell which part of the question the respondent was answering.

Leading Questions.

Leading questions will bias the results, this will reduce objectivity and hence the value of the research.

What is your opinion of the price of cinema admission? Very expensive - Expensive - Fair - Cheap - Very cheap

Cinema tickets are too expensive: Strongly agree - Agree - Disagree - Strongly disagree

You'll never get it 100% right, the question above has a rather subjective, "Fair" is open to interpretation - we might have used "About right" - it is hard to not be ambiguous and leave no room for interpretation.

Notice on the second of the two versions above that I didn't put a middle "neutral" value in. There is room for debate on this subject, not providing a fence for folk to sit on might encourage people to vote one way or another - but if a respondent has truly a neutral view they might choose to not fill in that question and so there is a bias in the data. The second version could be complimented by the same question asked in the opposite way, e.g. "Cinema tickets are not too expensive". We would expect to get a good level of negative correlation between the two versions, if so, this would indicate internal validity, if not it might indicate people were just clicking the same response to all the questions.

Layout and question types.

Be absolutely unambiguous about how the subject should fill in the question, e.g.

Do you hold a full driving licence? (Please circle the correct choice) YES NO

or probably better;

Do you hold a full driving licence? YES…☐ NO…☐

Use tick boxes rather than just blank space to solicit the subjects' choice, line them up with centred tabs, use, for example, the "Insert symbol" feature in MS Word to insert a box character. MS Word can offer more tricks, the "Forms" feature offers you a way to make the document interactive, useful if you intend to deliver and receive forms by email.

Strongly agree	Agree	Ambivalent	Disagree	Strongly disagree
☐	☐	☐	☐	☐

If your word processor doesn't offer box characters use brackets [ ].

An attractive survey form will be more appealing to the respondent and encourage a better quality of data. You can make a paper survey more inviting by enhancing readability, including white space to avoid large uninviting blocks of text, this increases readability. A very busy or cluttered questionnaire can confuse respondents. Colour might help in some cases, for example to delineate between sections.

Avoid using lots of different fonts, typically stick with Arial and use bold for headings, using lots of different text styles can make the document look scrappy and confuse the respondent.

Surveys conducted online have a greater variety of objects available to spice up the presentation but do make sure they don't detract from the basic data gathering agenda. The issue here is about your confidence in setting an online survey up and the issue of bias - it wouldn't be very good, for example, at assessing the level of computer confidence among a target group!

Try it out! Run a Pilot.

When you have created the ultimate questionnaire try it out. It is very unlikely to be right first time! Don't just pilot the survey but carry that data through to analysis to check that your analysis plan is capable of offering the results you are aiming for. Solicit comments from your pilot group, friends might be shy of being critical, make sure they feel it is OK to note the shortcomings.

How long should a questionnaire be?

How long is a piece of string? - there is no definite rule but as guidance the amount of time people will happily take in filling it in will depend on their interest or "stake" in it. I f you want to press me for a guide then twenty Likert type questions is probably

OK but forty is probably too many! It does depend partly on the target group. The real issue is how long does it take to fill it in? Another good reason to properly pilot it!

What kind of questions should I use?

They should fit two criteria; they should furnish the data required and they should give you data that can be arranged into a format you can analyse.

There are a couple of examples above, the Likert scale question and the yes/no question. It is vital that you consider how you will analyse the resultant data when adopting a question style. Yes/No and Likert questions are great, the Yes/No question yields categorical (Nominal) data. More specifically Yes/No or Male/Female are a specific type of category called a dichotomous category, one that can take just one of two values. You might meet others, e.g.

How did you get to work today (tick one only);

Walk	□
Car	□
Bus	□
Train	□
Other	□

The "Other" category is useful - if on the pilot you get a large contingent of "Other" then you might analyse these and introduce an extra named category.

Compare the question above to this one…

What transport do you use to travel to work (tick all that apply);

Walk	□
Car	□
Bus	□
Train	□
Other	□

This second version lets the respondent tick all the boxes they use or have used. The resulting data is more complex to analyse. It does have an advantage in that it lets us gauge the range of transport used, it doesn't though give us any discrimination between the popularity of the various modes of transport, if someone only used a car once this year they might sensibly still tick "Car" and "Bus" even if all their other journeys to work were buy bus.

Sorting and ordering questions.

Sorting and ordering questions tend to increase the complexity of analysis.

Rank the types of transport you use for travel to work,
1 = use most often, 5 = use least often;

Walk	□
Car	□
Bus	□
Train	□
Other	□

The data from this question will be richer than that from the earlier examples but as a consequence much more complex to analyse!

The question you must address is "am I making a rod for my own back?" i.e. don't make a questionnaire that you can't analyse, you have to get the results out of the data when it is all gathered!

Can I include open ended questions?

Many questionnaires place open-ended questions at the end, this makes analysis easier but do remember that these "qualitative" questions will be seen in the light of the quantitative ones that precede them - this is generally an issue when mixing qualitative and quantitative methodologies in the same questionnaire. The questions in the questionnaire might colour the thoughts of the respondent and influence their answers to the open questions.

So how do I analyse it then?

We can use a mixture of descriptive statistics and graphs and some nonparametric inferential statistics. Unlike examples when we have real measurements when we might be unsure about the wisdom of applying parametric methods, it is reasonable to apply nonparametric methods to the data collected from most questionnaires if the responses can be described as scores rather than true measurements. There is inevitable debate on this in the statistical community but I would suggest that you start from the basis of applying nonparametric methods rather than the other way round.

The data in the file *Students data 2001.sav* was gathered as part of a large project looking at the IT skills of new students. The data in the file are only a part of the data gathered, we have just kept a few sample question, but for these questions all the gathered data are in the file.

The part of the questionnaire that gathered the data is re-synthesised below, it is worth noting that when the data were gathered the university was split into schools, it has since been reorganised into a smaller number of faculties. Have a look at the questionnaire and check that you can see how it is related to the data file. When you analyse your own data you will have to translate the data from your questionnaires to a file on the computer. There are some general hints that might help;

- Each of your subjects/respondents will usually have one row in the data sheet.
- Each question will typically have one column (i.e. it will take up one variable).
- Responses will be stored as numbers (e.g. 1 to 5 for lickert scales) and the "Value Labels" will ascribe text labels to the numbers.
- If you have used Ranking or ordering questions then each option will take up a variable, this will also be the case when the respondent is asked to "tick all that apply"

We can use the *Students data 2001.sav* data file to have a go at some methods that might be useful…

First let's look at the file, there are 2614 entries in the file from first year students in the year 2001. Each entry takes up one row in the data sheet, this is usual for SPSS data, so in this file there are 2614 rows.

Depending upon the view of the data you have you will either see lists of words or numbers. You can toggle between the two views by choosing "Value Labels" from the "View" menu.

What school are you studying in?

EDS	Education	☐
HSC	Health and Social Care	☐
SCI	Science and Mathematics	☐
SED	Environment and Development	☐
SLM	Sport and Leisure Management	☐
CMS	Computing and Management Sciences	☐
SSL	Social Science and Law	☐
ENG	Engineering	☐
SCS	Cultural Studies	☐

What is your Gender?

Male	Female
☐	☐

How old are you?

18-24	25-30	31-40	41-50	51-60	60+
☐	☐	☐	☐	☐	☐

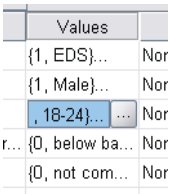
How do you rate your own basic computer use?

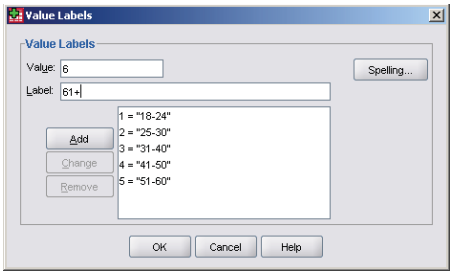
Below basic level	☐
Basic	☐
Competent	☐

How do you rate your ability to use statistics software? (e.g. Minitab, SPSS)

not competent	competent
☐	☐

To set these meanings behind the numbers you use the “Variable View” tab at the bottom of the screen. Click the “Values” column for the variable you want to create or alter labels for and then hit the small button that appears in the column, the “Value Labels” dialog box should appear. This is where you can type in each unique value and the corresponding text label. After typing in each pair click “Add” to add it to the list. You can also change and remove labels.



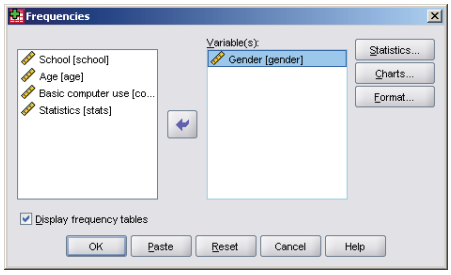


Spend some time on your data to get the labels correct, these labels will appear on your graphs and other output it is best to keep them reasonably short. SPSS will not automatically check the spelling of your labels.

Starting to look at the data.

It only takes SPSS a few seconds to do what might take all evening to do with questionnaires spread all over the dining room floor! So we can afford to play with the data to tease out meaning from it.

In our large sample of 2614 subjects we might want to do some basic demographic analysis, this is a useful preface to recording our results in any research project, it is where we tell the reader about the subjects who our results are based on. To analyse for simple percentages we can use the “Frequencies” command (choose Analyse then Descriptive Statistics, Frequencies). In this example I've put the Gender variable across to the variables box, have a go and hit the OK button. You might notice that the OK button is in a different place in this later version of SPSS, this change happened between versions 15 and 16, the functionality however is not altered. The output below is the resulting frequency table, it tells us that out of a total of 2614 respondents 1244 are female, 1342 are male and the data on gender is missing for 28. This accounts for all our 2614 subjects.



The percentage columns are of interest, the “Valid Percent” is calculated after the missing values are ignored. The “Cumulative Percent” isn't relevant for this analysis, but if we had data that were for example, an ordinal satisfaction scale, then this might be useful (*we might be making statements like “76% of responders were not dissatisfied”*). It can sometimes be helpful to think of the kind of statements that you might make about the results, this can help guide your analysis.

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	1342	51.3	51.9	51.9
	Female	1244	47.6	48.1	100.0
	Total	2586	98.9	100.0	
Missing	System	28	1.1		
Total		2614	100.0		

Which column would you use?

The valid percent leads us to statements like “51.9% of those responding to the question were male”, it would be sensible to offer the level of reply (in this case 98.9%) or (and I like this approach) put the results in a table, the actual figures can be put in brackets next to the percentages. A column can be made for the response rate for each question if you like.

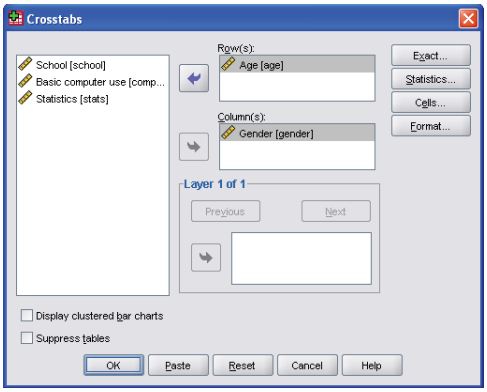
We could similarly look at the age profile of our respondents. Try this now. From the cumulative percentage column we can see that over 90% of respondents (91%) are 30 or younger. More importantly it gives us a good breakdown of the responses.

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-24	2161	82.7	85.1	85.1
	25-30	150	5.7	5.9	91.0
	31-40	160	6.1	6.3	97.3
	41-50	62	2.4	2.4	99.8
	51-60	6	.2	.2	100.0
	Total	2539	97.1	100.0	
Missing	System	75	2.9		
Total		2614	100.0		

Looking at two variables at once, for example; are the age profiles similar within the genders?

This is where crosstabulations come in useful.



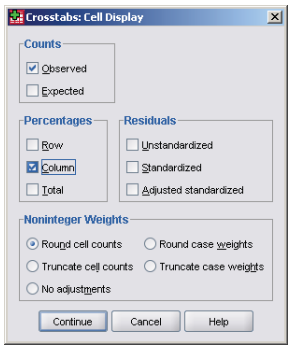
To create a crosstabulation pick "Crosstabs" from the Analyse, Descriptive Statistics menu, I've put the "Age" variable in the rows box and "Gender" in the columns. The output shows us the number of people in each age group but this time there is a column for each gender as well as a total column that should have the same figures in as the earlier frequency table we created unless age or gender data are missing.

Age * Gender Crosstabulation

Count

		Gender		
		Male	Female	Total
Age	18-24	1159	989	2148
	25-30	73	75	148
	31-40	60	98	158
	41-50	15	46	61
	51-60	2	3	5
Total		1309	1211	2520

This simple cross tabulation allows us to see that although there are slightly less females overall there are considerably more in the 31-40 and 41-50 age groups than there are males in those age ranges. We can get a better view of these results that will help us compare the gender/age relationship if we calculate percentages. We can ask SPSS to calculate the percentage of each gender in each age group. To do this go back to the Crosstabulation dialog box (Analyse, Descriptive stats, Crosstabs) and click the “Cells” button. Then click to add column percentages. The resulting table looks more complex because it gives both the raw number of respondents in each combination of gender and age group. You can if you want show percentages only by switching off the “Observed Counts” in the cells dialog.



In a results section you wouldn't simply copy and paste the output tables into the document, you might create a table including the output but in a more readable format, for example;

Crosstabulation of Age and Gender showing percentages within each gender.

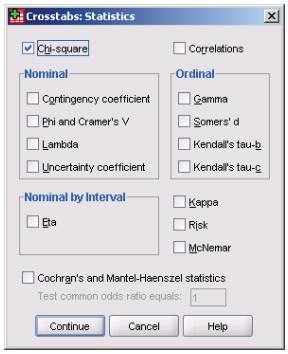
Age	Gender	
	Male	Female
18-24	1159 (88.50%)	989 (81.70%)
25-30	73 (5.60%)	75 (6.20%)
31-40	60 (4.60%)	98 (8.10%)
41-50	15 (1.10%)	46 (3.80%)
51-60	2 (0.20%)	3 (0.20%)

Make sure the title of your table clearly states what it intends to illustrate.

In this case we can see that larger percentages of females than males over 30 are becoming students.

We would have been surprised though if all the percentages were the same, some variability due to chance is inevitable. We can look to inferential statistics to tell us how likely we are to see such a difference in the percentages by chance. The statistic we will use for this is the Chi-square statistic.

The "Statistics" button on the Crosstabs dialog lets you request the Chi-square statistics. They come in various types, in our example here we don't need to worry about which to use, the p-value (Asymp. Sig) in each case is reported as ".000", we would report this as p<0.0005. (Note in this case the Pearson method has a note suggesting we use an alternative, we can though use the next one down.)

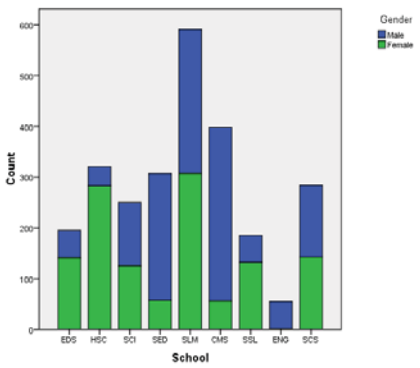


Chi-Square Tests

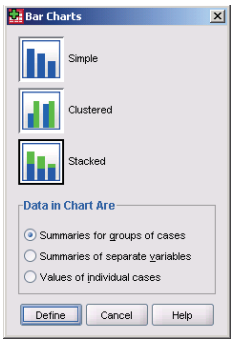
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.816[a]	4	.000
Likelihood Ratio	35.628	4	.000
Linear-by-Linear Association	32.798	1	.000
N of Valid Cases	2520		

a. 2 cells (20.0%) have expected count less than 5. The minimum expected count is 2.40.

A way to show this graphically…

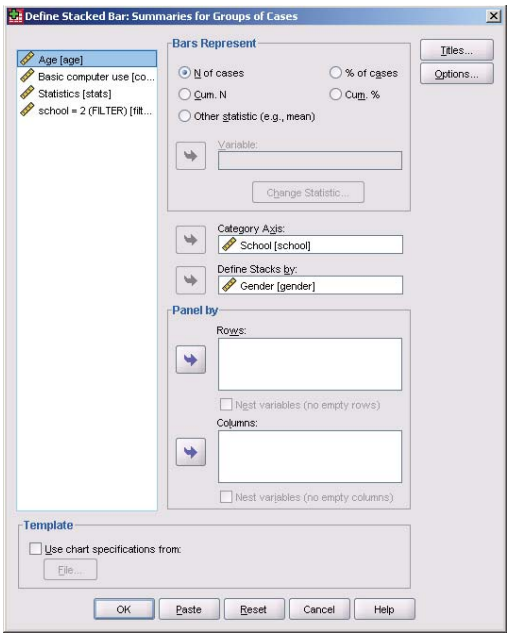


A bar chart would be useful to give an idea of the number of respondents from each school; we can go a step further and illustrate the male/female ratio at the same time. In the example here the height of the bar gives the number of respondent and the bar is stacked to show the male to female ratio for each bar.



To get this graph I used the old fashioned graph method, now tucked away under “Graphs, Legacy dialogs, Bar”, notice that in the initial dialog box for this method we have the option to go for a stacked bar chart, if you don’t want a stacked bar chart then leave it set at “simple” , the resulting dialog will not be as complex since you wouldn’t have to say which variable to use for the stacking.

This method could easily be applied to other questions where the answers were categorical, for example the question about travel.

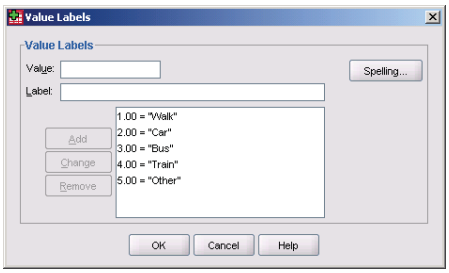


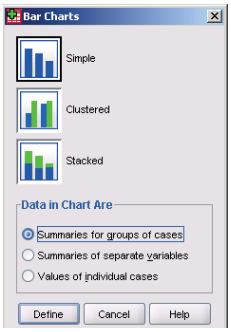
A brief recap about analysing “tick one only” type questions.

How did you get to work today (tick one only);

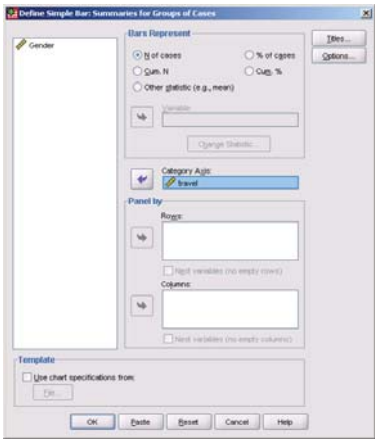
Walk	☐
Car	☐
Bus	☐
Train	☐
Other	☐

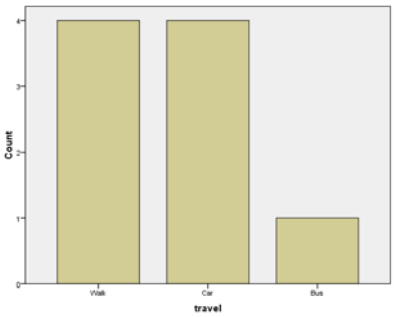
	Gender	travel
1	1	1.00
2	1	2.00
3	2	1.00
4	2	2.00
5	1	1.00
6	2	2.00
7	1	3.00
8	1	2.00
9	2	1.00





The data are coded into a single variable; this can take on one of five values in this example, depending upon the respondents’ choice. The numbers 1-5 used to code the data are given labels as previously described. This time a “Simple” bar chart can be requested from the legacy graph menu. The result isn’t too spectacular in this case, because the travel modes are similar in this small sample, a larger sample would have given more chance of people using the less popular methods. The next two ways of addressing the transport question give richer data but at a severe price in data handling complexity. The third type (ranking) can be simplified to “skim off” data similar to this example if it all gets too confusing.



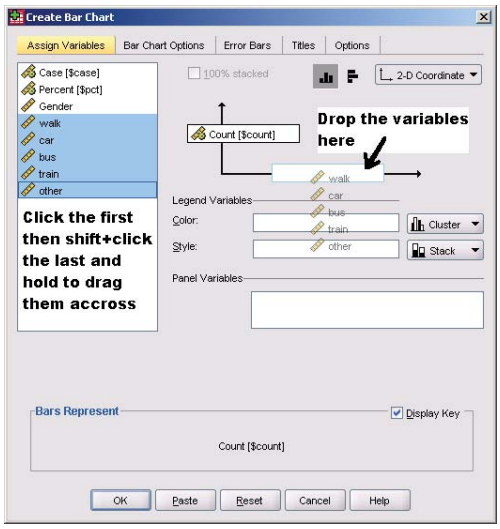


Analysing “tick all that apply” type questions.

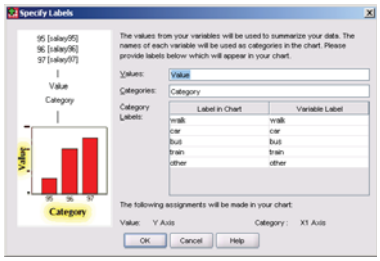
What transport do you use to travel to work (tick all that apply);	
Walk	☐
Car	☐
Bus	☐
Train	☐
Other	☐

	Gender	walk	car	bus	train	other
1	1	1	0	1	1	0
2	1	0	1	0	0	0
3	2	1	1	1	1	0
4	2	0	1	0	0	0
5	1	1	0	1	1	0
6	2	1	1	0	0	0
7	1	1	1	1	0	1
8	1	0	1	1	0	0
9	2	0	1	0	1	0
10						

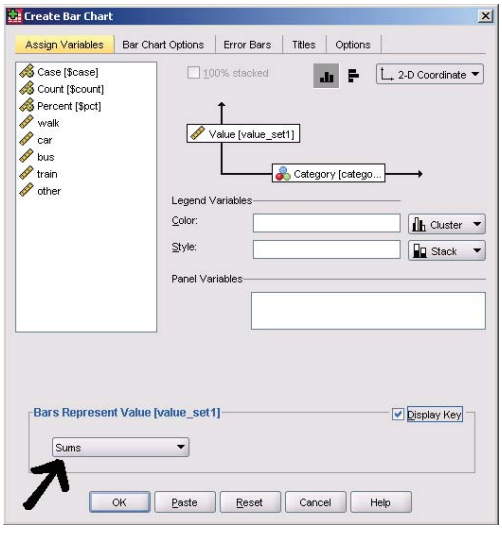
The above question could be coded and stored in SPSS by allocating one variable to each option (Walk, Car etc.) The file “Travel 2.sav” has some fictitious data in for you to play with. The responses are coded as 1=yes and 0 = no, the gender variable is included to illustrate that this is just part of a larger set of responses that might all have been stored in the same file. Analysing this structure is not as simple as when the respondent can only give one response.

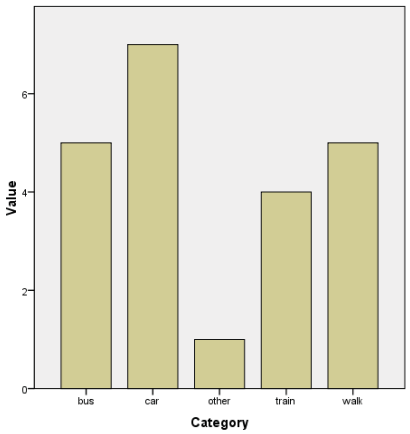


The Frequencies method (Analyse, Descriptive stats, frequencies) can be used to calculate the total number of votes for each type of transport by putting the five variables into the variables box and then clicking the “Statistics” button on the frequencies dialog box and asking for the “Sum” of each variable.



To get a graph you can use the Interactive legacy bar chart. The trick is to select all the necessary variables at once, do this by clicking on the first one then holding the shift key while clicking the last one. When they are selected drag them all to the horizontal axis (see the diagram). The “Specify labels” dialog should then appear, just OK this and finally, back in the “Create bar chart” dialog, select “Sums” instead of “Means” in the “Bars represent value” box at the bottom of the dialog box. You can now hit “OK”. The bars are in alphabetical order, this I expect can be altered, but frankly I’d rather not try! You might have a go at dragging the gender variable to the “Panel variables” box. Do though “Reset” this dialog before trying more tricks, it doesn’t like being used in this way.





Analysing "Ranking" type questions.

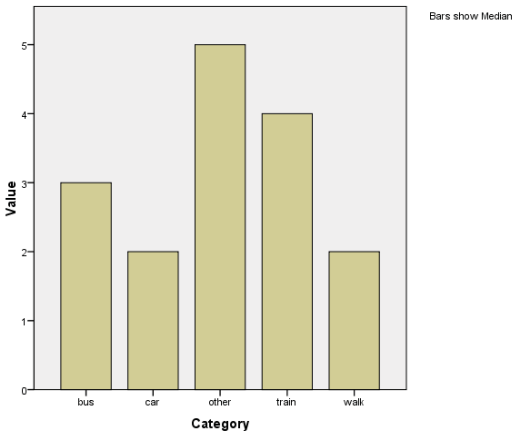
Rank the types of transport you use for travel to work,
1 = use most often, 5 = use least often;

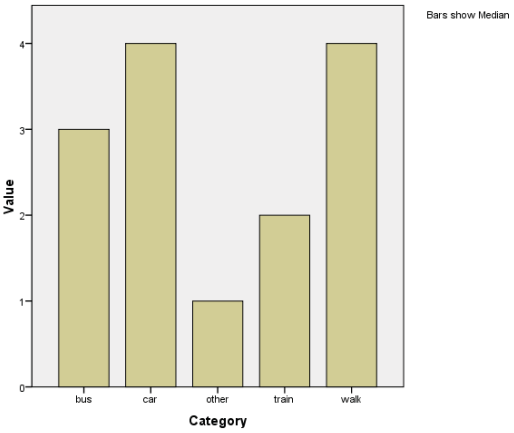
Walk ☐
Car ☐
Bus ☐
Train ☐
Other ☐

	Gender	walk	car	bus	train	other
1	1	1	2	3	4	5
2	1	2	1	5	3	4
3	2	1	3	2	5	4
4	2	3	1	2	4	5
5	1	1	2	4	3	5
6	2	2	1	4	3	5
7	1	4	2	1	3	5
8	1	5	1	3	4	2
9	2	1	5	3	4	2

The data for this question, again fictional, are stored in the file called "*Travel 3.sav*".

A similar graph can be constructed as above, in the "tick all that apply" example but you might consider using the Median rather than the sum. Another issue is that although the respondent is probably happier grading their most popular travel mode as a "one", rather like a league table, the reader of the resultant graph would typically expect to see the taller bars representing the more popular choices. This isn't the case unless you recode the data. Recoding data is a little involved, the command lives under the transform menu. The safe way to play with it is to use the "Save as" command to save a copy with a new name and play on that copy. The second graph here was done on recoded data and shows more clearly the popularity of each method.





In summary; Questionnaire data analysis.

What type of data do you have? Remember that different statistical procedures are appropriate for types of data and of course what you want to show! The choices are limited by the level of measurement of the variable(s) to be analysed.

Questionnaire derived data are likely to be nonparametric. The exception would be if you had people fill in their height or weight.

The categorical or nominal variables resulting from this method of data gathering provide a list of choices with no meaningful order to the list, e.g. our first travel question, or hair colour. The mean of a categorical variable is meaningless. Use the mode, frequency tables and crosstabulations with categorical variables. To illustrate this type of data, use bar charts (or pie charts if you wish to show proportion).

Ordinal variables have an implied order to the response choices. (e.g. 1= strongly agree, 2= agree, etc.) Typically use the median and mode for these variables, frequency tables (possibly even cumulative frequencies – but don't get carried away) and crosstabulations. Bar charts can display results usefully.

If your questionnaire yields some continuous variables (e.g. age in years where we know each year is the same distance apart from the next) we can apply many more statistics and if we really want we can condense them down into ordinal groups, (e.g. if we know the actual age we could reclassify the data into age groups.)

SPSS Inc. (2004). *SPSS Survey Tips* [online], SPSS Inc. Last accessed on 3/11/2008 at: http://www.spss.com/PDFs/STIPlr.pdf

University Computing Services, The University of Leeds (1996). *Guide to the Design of Questionnaires*. The University of Leeds Last accessed on (latest version) 3/11/2008 at: http://iss.leeds.ac.uk/info/312/surveys/217/guide_to_the_design_of_questionnaires/5

BUCKINGHAM, Alan and SAUNDERS, Peter (2004). *The Survey Methods Workbook.* Polity

---

The data and latest copy of the exercises will be available at;
http://teaching.shu.ac.uk/hwb/ag/resources/resourceindex.html

SHIKEN: JALT Testing & Evaluation SIG Newsletter. March 2011. *15*(1) 10-14.

Statistics Corner
Questions and answers about language testing statistics:

Likert items and scales of measurement?

James Dean Brown (University of Hawai'i at Mānoa)

Question: Many people have asked me this seemingly simple question: Are "Likert-scale" questions on questionnaires nominal, ordinal, interval, or ratio scales?

Answer: In preparing to answer this seemingly easy question, I discovered that the answer is far from simple. To explain what I found, I will have to address the following sub-questions:

1. What are scales of measurement?
2. What does the literature say about Likert items and scales of measurement?
3. What does common sense tell us about Likert items and scales of measurement?

What are Scales of Measurement?

Language researchers commonly describe the different ways they measure things numerically in terms of *scales of measurement*, which come in four flavors: nominal, ordinal, interval, or ratio scales. Each is useful in its own way for quantifying different aspects of language teaching and learning.

Nominal scales categorize. A nominal scale can be based on natural categories like gender (male or female) or artificial categories like proficiency (elementary, intermediate, or advanced proficiency groups). Nominal scales are also sometimes called *categorical scales*, or *dichotomous scales* (when there are only two categories).

Ordinal scales order or rank things. For instance, an item might ask students to rank ten types of classroom activities from most to least interesting (from 1 through 10). The most interesting activity would be first, followed by second, third, etc. (sensibly, ordinal scales are most often expressed as ordinal numbers). While the order is clear on such a scale, it is not clear what the distances are along the ordering. Thus the 1st activity might be much more interesting than the 2nd, but the 2nd activity might be only a little more interesting than the 3rd, and so forth. In short, ordinal scales show us the order, but not the distances between the rankings. Such ordinal scales are also sometimes called *ranked scales*.

Interval scales show the order of things, but with equal intervals between the points on the scale. Thus, the distance between scores of 50, 51, 52, 53 and so forth are all assumed to be the same all along the scale. Test scores are usually treated as interval scales in language research. Scales based on Likert items are also commonly treated as interval scales in our field.

Ratio scales differ from interval scales in that they have a zero value and points along the scale make sense as ratios. For example, a scale like age can be zero, and it makes sense to think of four years as twice as old as two years.

Researchers are often concerned with the differences among these scales of measurement because of their implications for making decisions about which statistical analyses to use appropriately for each. At times, they are discussed in only three categories: nominal, ordinal, and *continuous* (i.e., interval and ratio are collapsed into one category). [For more on scales of measurement, see Brown, 1988, pp. 20-24; 2001, pp. 17-18.]

What Does the Literature Say About Likert Items and Scales of Measurement?

Likert items were first introduced by Rensis Likert (1932). The following is an example of three Likert (pronounced /lɪkərt/) items:

Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. I understand the difference between Likert items and Likert scales.	1	2	3	4	5
2. I understand how to analyze Likert items.	1	2	3	4	5
3. I like using Likert items.	1	2	3	4	5

The example items have five options. They could equally well have 3, 4, 6, 7, or more options. [For more information on choosing the number of options and on how to write sound Likert items, see Brown, 2001, pp. 40-42, 44-54.]

When I first delved into the general literature on Likert items and scales of measurement, I found most articles were counter-intuitive and confusing. A number of articles argued or assumed that Likert items do not form an interval scale, but instead should be considered ordinal scales and should be analyzed accordingly (e.g., Coombs, 1960; Vigderhous, 1977; Jakobsson, 2004; Jamieson, 2004; Knapp, 1990; Kuzon, Urbanchek, & McCabe, 1996). Other articles proposed ways to get around this perceived ordinal/interval scale "problem" by proposing alternative Likert-like item formats such as the two-stage alternative offered by Albaum (1997) or the phrase completion alternative offered by Hodge and Gillespie (2003).

Despite all this discussion of the ordinal nature of Likert items and scales, most of the research based on Likert items and scales that I have seen in our field treats them as interval scales and analyzes them as such with descriptive statistics like means, standard deviations, etc. and inferential statistics like correlation coefficients, factor analysis, analysis of variance, etc. So you can see why I found the general literature counter-intuitive and confusing. For the most part, it says that we should treat Likert scales as ordinal scales, yet the research in my field consistently treats them as interval scales. How can these two positions be reconciled?

I believe that much of this ordinal/interval confusion arises from the fact that many authors use *Likert scale* to refer to both the Likert item type (items of the form shown above) and Likert scales (sums or averages of the results on sets of Likert items). For example, a questionnaire might have a total of 120 Likert items, divided into 12 Likert scales of 10 items each. If we carefully differentiate between Likert *items* and Likert *scales*, as I have done throughout this article, I think that much of the confusion will dissipate.

In addition, several papers have shown that *Likert scales* can indeed be analyzed effectively as interval scales (see for instance, Baggaley & Hull, 1983; Maurer & Pierce, 1998; and Vickers, 1999). Also, Allen and Seaman (1997, p. 2) support treating Likert scales as interval data with certain rather sensible provisos: "The "intervalness" here is an attribute of the data, not of the labels. Also, the scale item should be at least five and preferably seven categories. Another example of analyzing Likert scales as interval values is when the sets of Likert items can be combined to form indexes. However, there is a strong caveat to this approach: Most researchers insist such combinations of scales pass the Cronbach's alpha or the Kappa test of intercorrelation and validity. Also, the combination of scales to form an interval level index assumes this combination forms an underlying characteristic or variable."

In another vein, a number of authors have shown how Rasch analysis can be used to analyze and improve Likert scales as well as transform them into *true* interval scales. For more on this topic in the general literature, see Andrich (1978), Hagquist and Andrich (2004), Linacre (2002), Van Alphen, Halfens, Hasman, and Imbos (1994), and Waugh (2002); in the area of language research, see Sick (2006, 2009) or Weaver (2005, 2010).

What Does Common Sense Tell Us About Likert Items and Scales of Measurement?

Because they confuse Likert items with Likert scales, many authors look at a single Likert item and conclude that the 1 2 3 4 and 5 options form an ordinal scale at best, and therefore data based on these scales must be analyzed as though they are ordinal. I have two responses to that form of "logic".

When your read that MacArthur graduated first in the West Point class of 1903, that means he was at the top of his class ahead of whoever was second, third, fourth, fifth, etc. What is it about any Likert item 1 2 3 4 5 (much less an Likert scale) that can be expressed in ordinal numbers? Is *strongly agree* fifth, ahead of *agree* at fourth, and *neutral* at third, *disagree* at 2nd, and *strongly disagree* at 1st? This doesn't make sense, even at the Likert item level, much less at the Likert scale level.

From a Likert scale perspective, even if we were to accept the erroneous idea that Likert items are ordinal, saying that the resulting data must be analyzed as though they too are ordinal is like saying that test items that are scored right or wrong are nominal so data based on them must be analyzed as though they are nominal. Test scores are usually based on nominal right/wrong items, yet the total scores are always treated as interval data in our field. If the single argument (that Likert item options are ordinal) is wrong, then the compound argument (that Likert scales are ordinal [sic] because Likert items are ordinal [sic]) is doubly wrong.

The one 100% sensible treatment I have found for this set of issues is found in Carifio and Perla (2007). On page 114, they list "the top ten myths and urban legends about 'Likert scales' and the counter argument and 'antidote' for each myth and urban legend." According to the authors, the following myths are ***WRONG***:

> *Myth 1*—There is no need to distinguish between a scale and response format; they are basically the same "thing" and what is true about one is true about the other.
> *Myth 2*—Scale items are independent and autonomous with no underlying conceptual, logical or empirical structure that brings them together and synthesizes them.
> *Myth 3*—Likert scales imply Likert response formats and vice versa as they are isomorphic.
> *Myth 4*—Likert scales cannot be differentiated into macro and micro conceptual structures.
> *Myth 5*—Likert scale items should be analyzed separately.
> *Myth 6*—Because Likert scales are ordinal-level scales, only non-parametric statistical tests should be used with them.
> *Myth 7*—Likert scales are empirical and mathematical tools with no underlying and deep meaning and structure.
> *Myth 8*—Likert response formats can without impunity be detached from the Likert Scale and its underlying conceptual and logical structure.
> *Myth 9*—The Likert response format is not a system or process for capturing and coding information the stimulus questions elicit about the underlying construct being measured.
> *Myth 10*—Little care, knowledge, insight and understanding is needed to construct or use a Likert scale.

Notice in particular Myths 1, 5, and 6, which are directly related to the topic of this column. For more details about these 10 myths, you should of course refer to the original article.

Conclusion

The original question was: Are Likert-scale questions on questionnaires nominal, ordinal, interval, or ratio scales? My experience and my take on the literature lead me to believe that the following are true:

With regard to Likert *items* -

1. We must think about individual Likert items and Likert scales (made up of multiple items) in different ways.
2. Likert items represent an item format not a scale.
3. Whether Likert *items* are interval or ordinal is irrelevant in using Likert *scale* data, which can be taken to be interval.
4. If a researcher presents the means and standard deviations (interval scale statistics) for individual Likert items, he/she should also present the percent or frequency of people who selected each option (a nominal scale statistic) and let the reader decide how to interpret the results at the Likert-item level.
5. In any case, we should not rely too heavily on interpreting single items because single items are relatively unreliable.

With regard to Likert *scales* -

1. Likert scales are totals or averages of answers to multiple Likert items.
2. Likert scales contain multiple items and are therefore likely to be more reliable than single items.
3. Naturally, the reliability of Likert scales should be checked using Cronbach alpha or another appropriate reliability estimate.
4. Likert scales contain multiple items and can be taken to be interval scales so descriptive statistics can be applied, as well as correlational analyses, factor analyses, analysis of variance procedures, etc. (if all other design conditions and assumptions are met).

References

Albaum, G. (1997). The Likert scale revisited: An alternate version. *Journal of the Market Research Society, 39*, 331-349.

Allen, E., & Seaman, C. A. (2007). Likert Scales and Data Analyses. *Quality Progress, 40*, 64-65.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.

Baggaley, A., & Hull, A. (1983). The effect of nonlinear transformations on a Likert scale. *Evaluation & the Health Professions, 6*, 483-491.

Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about likert scales and likert response formats and their antidotes. *Journal of Social Sciences, 3*(3), 106-116.

Coombs, C. H. (1960). A theory of data. *Psychological Review, 67*, 143-159.

Hagquist, C., & Andrich, D. (2004). Is the Sense of Coherence-instrument applicable on adolescents? A latent trait analysis using Rasch-modelling. *Personality and Individual Differences, 36*, 955-968.

Hodge, D. R., & Gillespie, D. (2003). Phrase completions: An alternative to likert scales. *Social Work Research, 27*, 45-55.

Jakobsson, U. (2004). Statistical presentation and analysis of ordinal data in nursing research. *Scandinavian Journal of Caring Sciences, 18*, 437-440.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38*, 1212-1218.

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research, 39*, 121-123.

Kuzon, W. M. Jr., Urbanchek, M. G., & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery, 37*, 265-272.

Likert, R. (1932), A Technique for the measurement of attitudes. *Archives of Psychology, 140*, 1-55.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106.

Maurer, J., & Pierce, H. R. (1998). A comparison of Likert scale and traditional measures of self-efficacy. *Journal of Applied Psychology, 83*, 324-329.

Sick, J. (2006). *The learner's contribution: Individual differences in language learning in a Japanese high school*. Doctoral dissertation, Temple University, Philadelphia, PA.

Sick, J. (2009). Rasch measurement in language education part 3: The family of Rasch models. *Shiken: JALT Testing & Evaluation SIG Newsletter, 13*(1), 4-10.

Van Alphen, A., Halfens, R., Hasman, A., & Imbos, T. (1994). Likert or Rasch? Nothing is more applicable than a good theory. *Journal of Advanced Nursing, 20*, 196-201.

Vickers, A., 1999. Comparison of an ordinal and a continuous outcome measure of muscle soreness. *International Journal of Technology Assessment in Health Care, 15*, 709-716.

Vigderhous, G. (1977). The level of measurement and 'permissible' statistical analysis in social research. *Pacific Sociological Review, 20*(1), 61-72.

Waugh, R. F. (2002). Creating a scale to measure motivation to achieve academically: Linking attitudes and behaviours using Rasch measurement. *British Journal of Educational Psychology, 72*, 65–86

Weaver, C. (2005). Using the Rasch model to develop a measure of second language learners' willingness to communicate within a language classroom. *Journal of Applied Measurement, 6* (4), 396-415.

Weaver, C. (2010). *Japanese university students' willingness to use English with different interlocutors*. Doctoral dissertation, Temple University, Philadelphia, PA.

Where to Submit Questions:

Please submit questions for this column to the following e-mail or snail-mail addresses: *brownj@hawaii.edu*. Your question can remain anonymous if you so desire.

JD Brown, Department of Second Language Studies
University of Hawai'i at Manoa, 1890 East-West Road
Honolulu, HI 96822 USA

HTML: http://jalt.org/test/bro_34.htm / **PDF**: http://jalt.org/test/PDF/Brown34.pdf

Examples of Likert Scaled Responses Used in Data-Gathering

A variety of methods are available to assist evaluators in gathering data. One of those methods involves the use of a scale. One of the most common scale types is a Likert scale. A Likert scale is commonly used to measure attitudes, knowledge, perceptions, values, and behavioral changes. A Likert-type scale involves a series of statements that respondents may choose from in order to rate their responses to evaluative questions (Vogt, 1999).

Too little	OK as is	Too much
Ineffective	Uncertain	Effective
Not useful	Some impact	Useful
Will not do it	Undecided	Will do it
Definitely not	Undecided	Definitely will
Not essential	Makes no difference	Imperative
No	Maybe	Yes
Not at all	Very little	Some
Very hard	Hard	Neither hard nor easy
Yes	Somewhat	No

None	Slight	Considerable	Great
Poor	Fair	Good	Very good
Not important	Somewhat important	Important	Very important
None	A little	Quite a bit	Completely
Not aware	Somewhat aware	Usually aware	Very much aware
Not knowledgeable about	Somewhat knowledgeable about	Knowledgeable about	Very knowledgeable about

Strongly disagree	Disagree	No opinion or uncertain	Agree	Strongly agree
Very poor	Poor	Average	Good	Excellent
No value	Limited value	Average value	Much value	Extreme value
Very poorly	Poorly	Adequately	Well	Very well
Not valuable	Limited value	Average value	Valuable	Very valuable
Very much below average	Below average	Average	Above average	Very much above average
Inferior	Not good	Acceptable	Good	Superior
Very inferior	Inferior	Average	Superior	Very superior
Would not try	Poorly	Acceptably	Well	Very well
Very unhappy	Unhappy	Can take it or leave it	Satisfied	Highly satisfied
Very poor	Poor	Fair	Good	Very good
Not competent	Somewhat competent	Uncertain	Competent	Highly competent
False	More false than true	In between	More true than false	True
Hardly ever	Occasionally	Sometimes	Frequently	Almost always
Much less than most	Less than most	Above average	More than most	Much more than most
Poor	Fair	No opinion	Good	Excellent
Very bad	Bad	Average	Good	Very good
Very ineffective	Ineffective	Average	Effective	Very effective
Very slow	Slow	Average	Fast	Very fast
Poor	Unremarkable	Meets expectations	Better than expected	Outstanding
Excellent	Very good	Satisfactory	Very poor	Unacceptable

Decrease greatly	Decrease slightly	Stay the same	Increase slightly	Increase greatly

Very low	Low	Moderate	High	Very high

Little importance 1	 2	 3	 4	Great importance 5

Extremely dull	Very dull	Fairly dull	So-so	Fairly interesting	Very interesting

Not at all	Very little	Fairly well	Quite well	Very well	Perfectly

Exceptionally unfavorable	Unfavorable	Somewhat unfavorable	Somewhat favorable	Favorable	Exceptionally favorable

Excellent	Very good	Good	Satisfactory	Poor	Very poor	Unacceptable

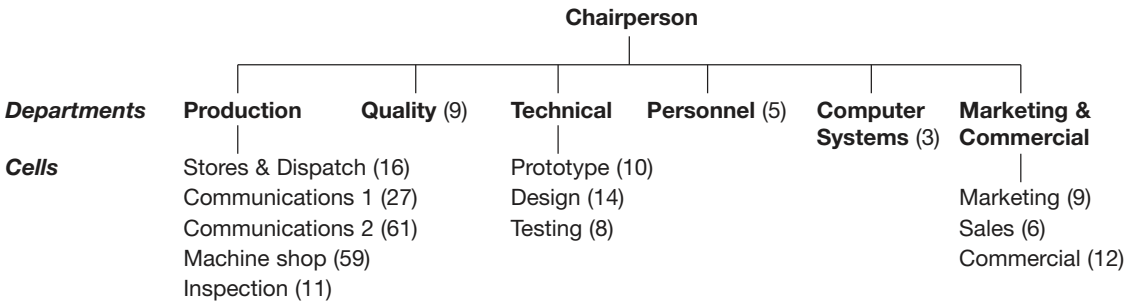
Vogt, W. Paul (1999). *Dictionary of statistics and methodology*. Sage: Thousand Oaks, California.

Case 7a
Change Management at Hattersley Electrics

Hattersley Electrics is a division of the Hattersley Group PLC, a United Kingdom based manufacturing conglomerate whose main markets were, until recently, the high tech aerospace and the defence industries. In recent years both the recession in the airline industry and the contraction in defence spending by European governments have hit the division. In order to overcome these Hattersley Electrics has embarked upon the process of repositioning itself within the electronics marketplace niche of advanced civilian communications equipment. As part of this it has changed its manufacturing foci and chosen to downsize and restructure its workforce.

Initially the workforce was not resistant to change due to the extensive efforts made by management to keep them involved and informed at all stages. This involved restructuring the division (Fig. 1) into cells and reducing the workforce from 380 to 250. After this restructuring the works council informed the division's management team that the workforce were concerned about the longer term effects of these changes. In particular they highlighted possible loss of employment, future job security, a lack of division identity, a lack of direction from senior management and a lack of employee involvement in recent decisions.

Fig. 1 Organisational Structure of Hattersley Group PLC Electrics Division



Figures in brackets refer to number of employees including managers.

A consultancy group was hired by the management team and set three objectives to be achieved (at a 95% level of certainty) within a fixed budget and a time scale of 10 weeks from start to finish. These were to establish:

1 the views of the division's managers and employees about the effectiveness of change at Hattersley Electrics;
2 general perceptions of the division's managers and staff about general aspects of the experience of the changes at Hattersley Electrics;
3 whether there were any differences in general perceptions between departments and between managers and employees.

The agreed methodology involved three stages of data collection. In the first stage interviews were undertaken with the division's chairperson and six managers and a sample of employees. One employee was selected at random from each cell. These interviews were to enable the consultants to understand the background to the changes, and to establish the full variety of opinions regarding its effectiveness. In addition they covered a range of other issues associated with general aspects of the experience of the changes.

In the second stage a questionnaire was distributed, which reflecting the issues generated by interviews in the first stage, to 50% of company managers and employees. This sample was selected from a sampling frame generated from the personnel department's staff database and was stratified by department and cell. Within each cell employees were listed in order of seniority. All employees other than those in the production department received a questionnaire; 28% of those employees in the production department received a questionnaire. The numbers sampled in each cell is given in Fig. 2. Overall there was a 97.6% response rate to the questionnaire, two non-responses coming from the Computer Systems department and one from the production department.

Fig. 2 Number of managers and employees in each cell who received a questionnaire

Production	**Quality** (9)	**Technical**	**Personnel** (5)	**Computer Systems** (3)	**Marketing & Commercial**
Stores & Dispatch (4)		Prototype (10)			Marketing (9)
Communications 1 (7)		Design (14)			Sales (6)
Communications 2 (18)		Testing (8)			Commercial (12)
Machine shop (17)					
Inspection (3)					

After the analysis of the questionnaire, three follow-up group interviews were undertaken, each group consisting of five or six people (stage 3). One group consisted of production employees, another of managers from all departments and a third of employees from all departments other than production. In the interviews issues that had arisen during the analysis of the questionnaire were probed and clarified.

Questions

1 **a**. Name the sampling techniques used at each of the three stages.
b. List possible reasons for the choice of each of these techniques.

2 Given the quoted response rate of 97.6% what is the accuracy (margin of error) of the questionnaire results for the division likely to have been?

3 **a**. What issues will need to be taken into account when generalising from the questionnaire results to:
i. all employees in each department of the Electrics division?
ii. all employees in the Electrics division?
b. Give reasons for your answers.

4 **a**. Would it be possible to use the results from the three sample surveys outlined in this case to generalise about the management of change in the Hattersley PLC group?
b. Give reasons for your answer.

Summarising Interview Transcripts

Questions	Interviewee 1	Interviewee 2	Interviewee 3	Key Themes
1. How would your staff and colleagues describe your leadership style?				
2. How have you influenced employees to follow your strategic vision for the organization?				

Summarising Interview Transcripts

Questions	Interviewee 1	Interviewee 2	Interviewee 3	Key Themes
3. What methods have you used to gain commitment from your team?				
4. How have you encouraged learning and development of employees?				

REFERENCES ON QUALITATIVE RESEARCH AND DATA ANALYSIS

- Bryman, A. & Burgess, R.G. (eds). (1994). Analyzing qualitative data. London: Routledge.
- Coffey, Amanda (1996). Making sense of qualitative data: Complementary research strategies. Thousand Oaks : Sage Publications.
- Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks: Sage.
- Davidson, C., & Tolich, M. (1999a). Making sense of data. In C. Davidson, & Tolich, M. (Ed.), Social science research in New Zealand: Many paths to understanding (pp. 143-164). Auckland: Longman.
- Denzin. N. & Lincoln, Y. (1994). Handbook of qualitative research. Thousand Oaks, CA: Sage.
- Feldman, M.S. (1995). Strategies for Interpreting Qualitative Data (Qualitative Research Methods Series. No 33). Thousand Oaks, CA: Sage.
- Grbich, C. (1999). Qualitative research in health: An introduction. St :Leonards, NSW, Australia: Allen & Unwin.
- Miles, M. B. & Huberman, A. M. (1994). Qualitative data analysis (2nd ed.). London: Sage.
- Morse, J. Ed., (1994). Critical Issues in Qualitative Research Methods. Thousand Oaks, CA: Sage.
- Opie, A. (1999). Unstructured interviewing. In C. Davidson, & Tolich, M. (Ed.), Social science research in New Zealand: Many paths to understanding (pp. 220-230). Auckland: Longman.
- Patton, M. Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage. Chapter 7.
- Pope, C., Ziebland, S., & Mays, N. (2000). Qualitative research in health care: Analysing qualitative data. British Medical Journal, 320, 114-116.
- Silverman, D. (1999). Doing qualitative research: A practical handbook. London: Sage.
- Strauss, A. & Corbin, J. (1990). Basics of Qualitative Research. Newbury Park: Sage.
- Waldegrave, C. (1999). Focus groups. In C. Davidson, & Tolich, M. (Ed.), Social science research in New Zealand: Many paths to understanding (pp. 231-242). Auckland: Longman.

VARIETIES OF QUALITATIVE RESEARCH

Source: Miles, M. B. & Huberman, A. M. (1994). Qualitative data analysis (2nd ed.). London: Sage. (excerpts from pp. 2-11)

Qualitative research may be conducted in dozens of ways, many with long traditions behind them. To do them all justice is impossible here. For our purposes the question is, What do some of the different varieties of qualitative research have to say about analysis? Can we see some common practices, some themes? (p. 5) .

Recurring features of qualitative research

... we suggest some recurring features of "naturalistic" research: Qualitative research is conducted through an intense and/or prolonged contact with a "field" or life

situation. These situations are typically "banal" or normal ones, reflective of the everyday life of individuals, groups, societies, and organizations.

- The researcher's role is to gain a "holistic" (systemic, encompassing, integrated) overview of the context under study: its logic, its arrangements, its explicit and implicit rules.
- The researcher attempts to capture data on the perceptions of local actors "from the inside", through a process of deep attentiveness, of empathetic understanding (Verstehen), and of suspending or "bracketing" preconceptions about the topics under discussion.
- Reading through these materials, the researcher may isolate certain themes and expressions that can be reviewed with informants, but that should be maintained in their original forms throughout the study.
- A main task is to explicate the ways people in particular settings come to understand, account for, take action, and otherwise manage their day-to-day situations.
- Many interpretations of this material are possible, but some are more compelling for theoretical reasons or on grounds of internal consistency.
- Relatively little standardized instrumentation is used at the outset. The researcher is essentially the main "measurement device" in the study.
- Most analysis is done with words. The words can be assembled, subclustered, broken into semiotic segments. They can be organized to permit the researcher to contrast, compare, analyze, and bestow patterns upon them. (pp. 5-7)

THE NATURE OF QUALITATIVE DATA

Source: Miles, M. B. & Huberman, A. M. (1994). Qualitative data analysis (2nd ed.). London: Sage. (excerpts from pp. 9-11)

In some senses, all data are qualitative; they refer to essences of people, objects, and situations. ... In this book we focus on data form of words - that is, language in the form of extended text. (Qualitative data also can appear as still or moving images, but we do not deal with these forms.)

The words are based on observation, interviews, or documents (or as Wolcott [1992] puts it, "watching, asking, or examining"). These data collection activities typically are carried out in close proximity to a local setting for a sustained period of time. Finally, such data are not usually immediately accessible for analysis, but require some processing. Raw field notes need to be corrected, edited, typed up; tape recordings need to be transcribed and corrected. (p. 9)

Strengths of Qualitative Data

What is important about well-collected qualitative data? One major feature is that they focus on naturally occurring, ordinary events in natural settings, so that we have a strong handle on what "real life" is like.

That confidence is buttressed by local groundedness, the fact that the data were collected in close proximity to a specific situation, rather than through the mail or over the phone. The emphasis is on a specific case, a focused and bounded phenomenon embedded in its context. The influences of the local context are not stripped away, but are taken into account. The possibility for understanding latent, underlying, or nonobvious issues is strong.

Another feature of qualitative data is their richness and holism, with strong potential for revealing complexity; such data provide "thick descriptions" that are vivid, nested in a real context, and have a ring of truth that has strong impact on the reader. Furthermore, the fact that such data are typically collected over a sustained period makes them powerful for studying any process (including history). We can go far beyond "snapshots" of "what?" or "how many?" to just how and why things happen as they do - and even assess causality as it actually plays out in a particular setting. And the inherent flexibility of qualitative studies (data collection times and methods can be varied as a study proceeds) gives further confidence that we've really understood what has been going on.
Qualitative data, with their emphasis on people's "lived experience", as fundamentally well suited for locating the meanings people place on the events, processes, and structures of their lives: their "perceptions, assumptions, prejudgements, presuppositions" and for connecting these meanings to the social world around them. (p. 10)

MAIN TYPES OF QUALITATIVE DATA GATHERING METHODS

Interview formats

- Individual
- Small groups
- Focus groups
- Whanau interviews
- Household interviews
- Interview structure
- Informal conversational interview
- General interview guide
- Standardised, open-ended interview

Observation

Styles of observation

- Ethnography
- Observing multiple settings and events
- Observing critical incidents and service delivery
- Noticing "subtle" behaviours
- Writing effective field notes
- Self-completion questionnaires
- client satisfaction surveys
- use of both open-ended & structured questions
- knowledge questions following training
- self-reports of behaviours and attitudes

Other qualitative methods

- case studies
- discourse analysis
- documents, archives and records
- media items (e.g., newspaper, magazine, audio)
- Focus Groups

Composition

- Usually about 5-10 people who have something in common
- A facilitator initiates and guides the discussion
- Special arrangements for recording the discussion
- Need a suitable room
- Usually 1½ - 2 hours long

Advantages

- generates ideas and elicits topics which are unlikely to arise with individual interviews
- can quickly identify a range of issues which are relevant to a research topic

Disadvantages

- generates large amounts of data
- needs a skilled facilitator to obtain good quality data and keep participants on topic

Selected References: Focus groups

- Kruegar, R. R. (1994). Focus groups: A practical guide for applied research (2nd. ed.). Thousand Oaks: Sage Publications.
- (Especially useful for evaluation research, good for overview of focus groups)
- Morgan, D. L (1997). Focus groups as qualitative research (2nd ed., Qualitative Research Methods Series, Vol. 16). Thousand Oaks: Sage (79 pp.)
- (Good overview of use of focus groups for research)
- Murphy, B., Cockburn, J., & Murphy, M. (1992). Focus groups in health research. Health Promotion Journal of Australia, 2(2), 37-40.
- Patton, M. Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage. (pp. 335-337)
- Quine, S. (1998). Focus groups: The role of the scribe and procedures for transcription. Health Promotion Journal of Australia, 8(3), 214-216.
- Stewart, D. & Shamdesani, P. (1990). Focus Groups: Theory and Practice. Thousand Oaks: Sage. (Has good links to theoretical literature in group dynamics and leadership)
- Vaughn, S., Schumm, J.S. & Singub, J. (1996). Focus Group Interviews in Education and Psychology. Thousand Oaks: Sage Publications.
- Waldegrave, C. (1999). Focus groups. In C. Davidson, & Tolich, M. (Ed.), Social science research in New Zealand: Many paths to understanding (pp. 231- 242). Auckland: Longman.

QUALITATIVE DATA GATHERING METHODS: Ethnography

Source: Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage. (pp. 58-61)

An ethnography is a description and interpretation of a cultural or social group or system. the researcher examines the group's observable and learned patterns of behavior, customs, and ways of life (Harris, 1968). As both a process and an outcome of research (Agar, 1980), an ethnography is a product of research, typically found in book-length form. As a process, ethnography involves prolonged observation of the group, typically through participant observation in which the researcher is immersed in the day-to-day lives of the people of through one-on-one interviews with members

of the group. The researcher studies the meanings of behavior, language, and interactions of the culture-sharing group.
…the ethnographer begins the study by looking at people in interaction in ordinary settings and by attempting to discern pervasive patterns such as life cycles, events, and cultural themes.... Culture is an amorphous term, not something "lying about" ... but rather something the researcher attributes to a group as he or she looks for patterns of daily living. It is inferred from the words and actions of members of the group and is assigned to this group by the researcher. It consists of looking for what people do (behaviors), what they say (language), and some tension between what they really do and what they ought to do as well as what they make and use (artifacts) (Spradley, 1980). Thus, the ethnographer gathers artifacts and physical trace evidence; finds stories, rituals, and myths; and/or uncovers cultural themes.
…the themes of structure and function guide research of social organizations. Structure refers to the social structure or configuration of the group, such as the kinship or political structure of the social-cultural group. Function refers to patterns of the social relations among members of the group that help regulate behavior.
To establish these patterns, the ethnographer engages in extensive work in the field, called fieldwork, gathering information through observation, interviews, and materials helpful in developing a portrait and establishing "cultural rules" of the culture-sharing group. As Wolcott (1996) comments, "They (researchers) establish what a stranger would have to know in order to understand what is going on here, or, more challenging still, what a stranger would have to know in order to be able to participate in a meaningful way" (p. 6). The ethnographer is sensitive to fieldwork issues ... such as gaining access to the group through gatekeepers, individuals who can provide entrance to a research site. the ethnographer locates key informants, individuals who provide useful insights into the group and can steer the researcher to information and contacts.
... the procedures in ethnography call for a detailed description of the culture-sharing group or individual, an analysis of the culture-sharing group by themes or perspectives, and some interpretation of the culture-sharing group for meanings of social interaction and generalizations about human social life (Wolcott, 1994b).
The final product of this effort is a holistic cultural portrait of the social group that incorporates both the views of the actors in the group (emic) and the researcher's interpretation of views about human social life in a social science perspective (etic).

The ethnography is challenging to use for the following reasons: Grounded Theory

Source: Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage. (pp. 55

- The researcher needs to have a grounding in cultural anthropology and the meaning of a social-cultural system as well as the concepts typically explored by ethnographers.
- The time to collect data is extensive, involving prolonged time in the field.
- In many ethnographies, the narratives are written in a literary, almost storytelling approach, an approach that may limit the audience for the work and may be challenging for authors accustomed to traditional approaches to writing social and human science research.

- There is a possibility that the researcher will "go native" and be unable to complete the study or be compromised in the study. This is but one issue in the complex array of fieldwork issues facing ethnographers who venture into an unfamiliar cultural group of system.-58)

... the intent of a grounded theory study is to generate or discover a theory, an abstract analytical schema of a phenomenon, that relates to a particular situation. This situation is one in which individuals interact, take actions, or engage in a process in response to a phenomenon. To study how people act and react to this phenomenon, the researcher collects primarily interview data, makes multiple visits to the field, develops and interrelates categories of information, and writes theoretical propositions or hypotheses or presents a visual picture of the theory.

The centerpiece of grounded theory research is the development or generation of a theory closely related to the context of the phenomenon being studied. Strauss and Corbin (1994), for example, mention that a theory is a plausible relationship among concepts and sets of concepts. This theory, developed by the researcher, is articulated toward the end of a study and can assume the form of a narrative statement (Strauss & Corbin, 1990), a visual picture (Morrow & Smith, 1995), or a series of hypotheses or propositions (Creswell & Brown, 1992).

The researcher typically conducts 20-30 interviews based on several visits "to the field" to collect interview data to saturate (or find information that continues to add until no more can be found) the categories. A category represents a unit of information composed of events, happenings, and instances (Strauss & Corbin, 1990). The researcher also collects and analyzes observations and documents, but these data forms are atypical. While the researcher collects data, she or he begins analysis. In fact, my image for data collection in a grounded theory study is a "zigzag" process - out to the fields to gather information, analyze the data, back to the field to gather more information, analyze the data, and so forth. The participants interviewed are theoretically chosen - in theoretical sampling - to help the researcher best form the theory. How many passes one makes to the field depends on whether the categories of information become saturated and whether the theory is elaborated in all of its complexity. This process of taking information from data collection and comparing it to emerging categories is called the constant comparative method of data analysis.

The process of data analysis in grounded theory research is systematic and follows a standard format:

In open coding, the researcher forms initial categories of information about the phenomenon being studied by segmenting information. Within each category, the investigator finds several properties, or subcategories, and looks for data to dimensionalize, or show the extreme possibilities on a continuum of, the property.

In axial coding, the investigator assembles the data in new ways after open coding. This is presented using a coding paradigm or logic diagram in which the researcher identifies a central phenomenon (i.e., a central category about the phenomenon), explores causal conditions (i.e., categories of conditions that influence the phenomenon), specifies strategies (i.e., the actions or interactions that result from the central phenomenon), identifies the context and intervening conditions (i.e., the narrow and broad conditions that influence the strategies), and delineates the consequences (i.e., the outcomes of the strategies) for this phenomenon.

A grounded theory study challenges researchers for the following reasons:

- The investigator needs to set aside, as much as possible, theoretical ideas or notions so that the analytic, substantive theory can emerge.
- Despite the evolving, inductive nature of this form of qualitative inquiry, the researcher must recognize that this is a systematic approach to research with specific steps in data analysis.
- The researcher faces the difficulty of determining when categories are saturated or when the theory is sufficiently detailed.
- The researcher needs to recognize that the primary outcome of this study is a theory with specific components: a central phenomenon, causal conditions, strategies, conditions and context, and consequences. These are prescribed categories of information in the theory.

References: Grounded Theory

- Lonkila, M. (1995). Grounded theory as an emerging paradigm for computer-assisted qualitative data analysis. In U. Kelle (Ed.), Computer-aided qualitative data analysis: Theory, methods and practice (pp. 41-51). London: Sage.
- Olshansky, E. F. (1996). Theoretical issues in building a grounded theory: Application of an example of a program of research on infertility. Qualitative Health Research, 6(3), 394-405.
- Strauss, A. (1987). Qualitative Analysis for Social Scientists. Cambridge: Cambridge University Press.
- Strauss, A. & Corbin, J. (1990). Basics of Qualitative Research. Newbury Park: Sage.

QUALITATIVE DATA ANALYSIS: Phenomenological Analysis

Source: Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage. pp. 51-55.

A phenomenological study describes the meaning of the lived experiences for several individuals about a concept or the phenomenon.

Phenomenological data analysis proceeds through the methodology of reduction, the analysis of specific statements and themes, and a search for all possible meanings. The researcher also sets aside all prejudgements, bracketing his or her experiences and relying on intuition, imagination, and universal structures to obtain a picture of the experience.

The researcher needs to understand the philosophical perspectives behind the approach, especially the concept of studying how people experience a phenomenon. The investigator writes research questions that explore the meaning of that experience for individuals and asks individuals to describe their everyday lived experiences. The investigator then collects data from individuals who have experienced the phenomenon under investigation.

The phenomenological data analysis steps are generally similar for all psychological phenomenologists who discuss the methods. The original protocols are divided into statements or horizonalization. Then, the units are transformed into clusters of meanings expressed in psychological and phenomenological concepts. Finally, these transformations are tied together to make a general description of the experience, the textural description of what was experienced and the structural description of how it was experienced.

The phenomenological report ends with the reader understanding better the essential, invariant structure (or essence) of the experience, recognizing that a single unifying meaning of the experience exists.

Example using phenomenology

Parsons, K. (1997). The male experience of caregiving for a family member with Alzheimer's disease. Qualitative Health Research, 7(3), 391-407.

Methodology (p. 393)
In keeping with one of the identified needs in caregiving research, "to hear from even more of the participating voices than we currently do" and "to turn directly to lived experience and the related and diverse situations and working local discourses of caregiving" … , the researcher used the phenomenological method as described by Van Manen (1990).
The aim of phenomenology is to explicate the meaning of human phenomena and to understand the lived structures of meanings of everyday experience. Going beyond the actual state of affairs, such as the how, where, what, when, or why something happened, phenomenology is concerned with the essence or nature of the lived experience for a particular individual. It is concerned with interpreting the meaning of the lived experience, our lifeworld (Van Manen, 1990). (p. 393)

Data analysis (p. 395)
Following the completion of each interview, the researcher transcribed the tapes. This transcription process helped immerse the researcher in the data and helped the researcher to think about what the interviewees were saying and how they were saying it. Each written transcript was read several times while listening to the corresponding audio tape to ensure accuracy of the transcribed tape and to come to a better overall understanding of each participant's experience. This process of transcribing and listening also prompted additional questions for a subsequent interview.
The specific approach used to uncover the thematic aspects of the caregiving experience was the selective or highlighting approach outlined by Van Manen (1990). In the selective reading approach, the text was read several times and statements that appear to be revealing about the phenomenon were underlined or highlighted. Themes were identified by highlighting material in the interview text that spoke to each man's experience. Next, the researcher selected each of these highlighted phrases or sentences and tried to capture as fully as possible what meaning the highlighted material conveyed.
Following the initial readings and preliminary identification of themes in each of the interviews, the researcher met with three other researchers to discuss the themes and any areas that required more investigation. These meetings helped to ensure that the researcher's "decision trail" … could be followed and the themes refined. The researcher next compared the themes in each interview, looked for commonalities and differences, and identified the overall themes that best described the experiences of these men as caregivers.
With the themes identified, the researcher then began the process of writing the themes and describing how they were interrelated. Rewriting continued until the researcher felt the themes (parts) and the relationship between the themes (whole) captured as accurately as possible the way these men experienced caregiving. (p. 395)

References: Phenomenology

- Dale, P., & Allen, J. (1998). On memories of childhood abuse: a phenomenological study. Child Abuse & Neglect, 22(8), 799-812.
- Holstein, J. A., & Gubrium, J. F. (1994). Phenomenology, ethnomethodology, and interpretive practice. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 262-272). Thousand Oaks: Sage.
- Jarrett, M. E., & Lethbridge, D.J. (1994). Looking forward, looking back: Women's experience with waning fertility during midlife. Qualitative Health Research, 4(4), 370-384.
- Kugleman, R. (1999). Complaining about chronic pain. Social Science & Medicine, 49, 1663-1676.
- Madjar, I., & Walton, J. A. (Eds.). (1999). Nursing and the experience of illness: phenomenology in practice. New York: Routledge.
- Smythe, L. (2000, June). Making sense of qualitative research. Kaitiaki Nursing New Zealand, 16-18
- Van Manen, M. (1990). Researching lived experience: human science for an action sensitive pedagogy. London, Ontario: Althouse Press.

Original article

doi: 10.1111/jcal.12036

Journal of Computer Assisted Learning

Learning in a mobile age: an investigation of student motivation

K. Ciampa

Teacher Education Department, Brock University, St. Catharines, Ontario, Canada

Abstract The purpose of this single-case study was to explore the lived experiences of a grade 6 teacher and students who used tablets as part of their classroom instruction. Malone and Lepper's taxonomy of intrinsic motivations for learning is used as a framework for examining whether and how this particular theory of motivation applies equally well for mobile learning. This study reports on the grade 6 teacher's and students' perceptions regarding the motivational affordances of using these mobile devices for learning. The findings are consistent with those of Malone and Lepper that motivation can be enhanced through challenge, curiosity, control, recognition, competition and cooperation. This model is helpful in informing our understanding of the motivating features of using mobile devices for learning and how mobile technologies can be used to enhance learners' motivation.

Keywords digital games, engagement, iPads, motivation, mobile learning.

> Often students will be yelling out loud in class, but when we use the tablets, they're much quieter because they're paying attention to it. (Mark, Student, Grade 6)
>
> My students are excited and enjoy using the tablets and the apps, and are motivated to work on the tablet, so they do not get off task. (Natasha, Teacher, Grade 6).

Introduction

Over the past decade, great strides have been made in investigating the cognitive processes involved in mobile learning (Shuler, 2009). During the same period, however, attention to motivational factors relating to mobile learning has been minimal (Sharples, 2007; Sharples, Arnedillo-Sánchez, Milrad, & Vavoula, 2009). In the opening quote(s), Mark, a grade 6 student, and Natasha, his grade 6 teacher, reflect on the observed differences in behaviour when students are engaged with tablets during language arts class. As evidenced by these self-reports, when children use tablets, they are generally found to be very engaged in the process: they are on-task and totally immersed in it with little or no awareness of the more general world around them (Beck & Wade, 2006; Csikszentmihalyi, 1990; Shaffer, 2006). Many learners are motivated and excited to use mobile devices; as yet, however, there is little understanding of what it is that makes learning with mobile devices so engaging and motivating to use. According to Malone and Lepper (1987), motivation is a necessary precondition for student involvement in any type of learning activity; what and how effectively students learn may be influenced by their level of motivation. Vogel, Kennedy, and Kwok (2009) claimed that students' motivation plays a significant role in engaging and sustaining students to use mobile devices for learning purposes. This study seeks to examine this taxonomy of intrinsic motivations through the lens of mobile learning. Malone and Lepper's taxonomy of intrinsic motivations for

Accepted: 2 June 2013
Correspondence: Katia Ciampa, Teacher Education Department, Brock University, 500 Glenridge Avenue, St. Catharines, ON, Canada L2S 3A1. Email: kciampa@brocku.ca

learning is used as a framework for examining whether and how this particular theory of motivation, which has been applied to non-mobile learning (Sharples, 2007), applies equally well for this new context. The purpose of this qualitative single-case study was to explore the lived experiences of a grade 6 teacher and her students who used tablets as part of their classroom instruction. More specifically, this study reports on the grade 6 teacher's and students' perceptions regarding the motivational affordances of using these mobile devices for learning. That said, this paper aims to describe how the use of mobile devices for classroom instruction relates to the theoretical accounts of what motivates students to learn. Accordingly, the following research question guides this study: What do elementary teachers and students perceive as the motivational affordances of using mobile devices for learning?

Theoretical framework: taxonomy of intrinsic and extrinsic motivations for learning

Learning that is fun appears to be more effective (Lepper & Cordova, 1992). Also, Quinn (1994) argues that for games to benefit educational practice and learning, they need to combine fun elements with aspects of instructional design that include motivational, learning and interactive components. Deci and Ryan (1985) have noted that self-determined learner behaviour can stem from both intrinsic motivation (i.e., the learner engages in an activity because it is interesting or enjoyable) and from extrinsic motivation (i.e., the learner engages in an activity because he or she desires the outcome and wants to achieve some instrumental end such as earning a reward). In the past, research on motivation has mainly focused on assessing student motivation in a traditional classroom environment (Dornyei, 2000). With respect to technology-supported learning environments, however, research focusing on students' motivation is limited. As we create information systems to support programs and curricula, it becomes imperative that we understand the scope of technology-supported learning activities on aspects of motivation.

In line with the aim of this research study, Malone and Lepper's (1987) work on games focused particularly on what makes games both fun and educational. This early work used the existing literature on motivation backed up by a number of empirical studies to develop a theory of intrinsically and extrinsically motivating instruction for games. Malone and Lepper's theory is based on six categories of individual motivations that make an activity both intrinsically and extrinsically motivating for a learner and ultimately contribute to the fun in games. As discussed below, Malone and Lepper proposed that the following elements make an activity both intrinsically and extrinsically motivating for a learner: challenge, curiosity, control, cooperation, competition and recognition. It is believed that Malone and Lepper's motivation theory may provide important clues as to how and why mobile technologies are perceived 'fun', which can become powerful catalysts for change as well as tools for redesigning our learning and instructional systems. In addition to its theoretical contribution, this research presents important practical contributions through the identification of important factors deemed to support students' motivation in (mobile) technology-supported learning environments.

Intrinsic motivations for learning

Challenge

While in a state of flow or while playing a game, learning is made possible through the use of concrete goals. To prevent the learner from wandering around aimlessly, a game creates goals that the user must meet before being able to progress. Malone and Lepper (1987) claimed that learners are more motivated when goals are clearly defined and when challenge is balanced in such a way that the learning process is neither too easy as to bore the learner, or too difficult such that success seems impossible. There are several ways in which an optimal level of challenge can be obtained. Malone and Lepper (1987) suggest that activities should employ varying difficulty levels of instruction, establish multiple levels of goals, vary time constraints, provide incomplete information and make the learner seek out the missing elements.

Most mobile games and 'apps' (applications used on mobile devices) also allow for self-selected differentiation of difficulty level at the start of the game (e.g., easy, medium, hard) where students can move through the levels at their own pace or automatically adjusted difficulty levels according to how the student performs (Chou, Block, & Jesness, 2012). The ability to adjust

content to student level and allow self-paced learning may thus lend mobile technology as an ideal tool for implementing differentiated instruction in the classroom.

Performance feedback and score keeping allows the individual to track progress towards desired goals. Finally, goals must be meaningful, personalized and specific to the individual; activities that are within the individual's zone of proximal development (Vygotsky, 1978) will stimulate the greatest intrinsic motivation (Malone & Lepper, 1987). Most mobile games and apps can also provide immediate feedback and thus provide continued motivation for those who are not motivated by traditional educational settings (Valk, Rashid, & Elder, 2010).

Curiosity

Curiosity is the most direct intrinsic motivation for learning. The concept of curiosity can be divided into two broad categories: sensory curiosity and cognitive curiosity (Malone & Lepper, 1987). Sensory curiosity involves the attention-attracting value of variations and changes in the light, sound or other sensory stimuli of an environment. When considering motivation within multimedia learning environments, both an individual's sensory or cognitive curiosity can be stimulated. Multimedia effects such as videos, audio, music, animation and interactive capabilities afforded by mobile devices evoke sensory curiosity (Liu, Toprac, & Yuen, 2009). Mobile devices such as the tablet also afford greater opportunities for haptic modality, a new channel for communication through mobile technology by utilizing the sense of touch (Wong, Chu, Khong, & Lim, 2010). The tablet, in particular, features flicking, tapping, pinching and stretching. These haptic touch features have enhanced the visual feedback which also enhanced the player's experience during interaction and gameplay (Wong *et al*., 2010).

Cognitive curiosity is also aroused when learners discover that their knowledge is incomplete or inconsistent, and they have the desire to explore and attain new information and competence with the technology (Malone & Lepper, 1987). Technology-enhanced environments afford individuals with almost limitless opportunities for exploration and ready access to information to support both sensory and cognitive curiosity (Liu *et al*., 2009). This desire for new information can lead to deepening levels of interest and vice versa (Malone & Lepper, 1987).

Control

The concept of control is another cornerstone of intrinsic motivation (Malone & Lepper, 1987). Deci, Betley, Kahle, Abrams, and Porac (1981) define intrinsic motivation as a striving for competence and self-determination (where self-determination means control). Researchers have indicated that locus of control is associated with motivation when students are given control over their learning (Klein & Keller, 1990). According to Malone and Lepper (1987), the 'mere illusion of control' significantly improves motivation and academic performance (p. 238). Control is determined by the range of choices offered by an activity, the extent to which outcomes are contingent on the responses of the player, and the inherent power of these responses (Joiner, Nethercott, Hull, & Reid, 2006). This motivation is best promoted when the activity provides 'a sense of personal control over meaningful outcomes' (Malone & Lepper, 1987, p. 258).

The role of choice in motivation is also well recognized (Gambrell, 1996). Opportunities for choice promote students' independence and versatility (Turner, 1995). Environments that provide choices and self-direction support the feeling of autonomy, which enhances intrinsic motivation (Deci & Ryan, 1985). Task engagement also increases when students are provided with opportunities to make choices about their learning (Deci & Ryan, 1985). The mobile user's ability to make his/her own choice is one of the pillars on which the success of ubiquitous mobile environments for learning rests. Mobile technologies have the potential to support and encourage the view of the student as a self-regulated learner and constructivist approaches to pedagogy both within and beyond the classroom by assisting the learner to interact with his/her environment, make independent choices and regulate their own learning (Beishuizen, 2008). In addition, the personal, multimodal, independent capabilities of devices such as the tablet offer the potential for 'anywhere, anytime' learning (Evans & Johri, 2008; Norris & Soloway, 2008).

Extrinsic motivations for learning

Although extrinsic rewards can be less effective than intrinsic motives, both intrinsic and extrinsic motives

play a role in determining learner behaviour. The goal is to develop learners who are self-directed and self-motivated, both because the activity is interesting in itself and because achieving the outcome is important. Where intrinsic motivation to learn is the educator's ultimate goal, extrinsic motivators such as cooperation, competition and recognition can and should also be considered when designing learning environments or selecting instructional materials (Malone & Lepper, 1987).

Cooperation

Pure cooperation is generally defined as involving a group of individuals working together to attain a common goal (Malone & Lepper, 1987). Many theorists have argued that cooperation should facilitate performance, especially when individuals hold interdependent goals (e.g., Malone & Lepper, 1987). According to Johnson and Johnson (2003), cooperation (compared with competitive and individualistic efforts) promotes greater effort exerted to achieve and greater productivity; more on-task behaviour, higher quality of relationships among participants (e.g., greater interpersonal cohesion, task-oriented and personal support) and greater psychological adjustment (e.g., greater social competencies, higher self-esteem).

Mobile technology can be a tool to deliver one-on-one instruction or serve as a medium for collaboration. Students can learn at their own pace, collaborate with others and offer advice to each other through various apps. Utilizing student-centred activities and apps that match with the curriculum to encourage student collaboration and creativity would create a student-centred, socially interactive classroom; all important skills of the 21st century (Chou *et al*., 2012).

Competition

Competition is one of the basic components; competition is a component of many intrinsically motivated 'play' activities. Csikszentmihalyi (1990) stated that achievement motivation (itself a complex of intrinsic and extrinsic motivation) involves competition against a standard of excellence. Competition is usually spoken in terms of two or more people or groups having directly opposing goals. However, Csikszentmihalyi made a similar distinction by differentiating the following two items: 'measuring self against others' (direct competition) and 'measuring self against own ideal' (indirect competition). In indirect competition, the individual or group struggles to perform well against an impersonal standard such as one's best previous performance or the performance norms for one's ability level. Direct competition, however, involves people struggling against one another. Insofar as one plays in order to win, rather than to play well, an extrinsic orientation dominates over an intrinsic one. As one would expect, success led to greater willingness for future participation than failure.

Similarly, when children set out to do a task, they can either proceed with a mastery orientation or a performance orientation (Dweck & Leggett, 1988). Children with a mastery orientation have learning goals – they are concerned with increasing their competence and abilities while mastering new tasks over time. Conversely, children with a performance orientation have performance goals – they are concerned with eliciting positive judgments about their work. There is strong evidence that a mastery orientation can boost children's academic performance, in the short- and long-term. In an experimental study, Dweck and Leggett (1988) manipulated fifth graders' orientation by highlighting either performance goals or learning goals, and by providing feedback indicating either high or low ability on a task. They found that in response to obstacles, mastery-oriented children tended to view challenging situations as an opportunity to acquire new skills or extend their mastery. This response caused them to seek challenges with a positive attitude and high persistence. Performance-oriented children, on the other hand, sought to avoid others' unfavourable judgments. They avoided failure by avoiding risk and difficult/challenging tasks. In response to failure, performance-oriented children were more likely to give up, because they saw failure as evidence of low competence (Dweck & Leggett, 1988). This study aims to describe whether and which of the two competitive forces (direct or indirect competition) and goal orientations (performance or mastery goals) plays a greater role in influencing students' motivation to learn with mobile devices.

Recognition

The final kind of intrinsic motivation that can be used in designing instructional environments is recognition (Malone & Lepper, 1987). There is some general agreement among traditional motivational theorists

(e.g., Deci & Ryan, 1985; Malone & Lepper, 1987) that learners enjoy having their efforts and accomplishments recognized and appreciated by others. In order for an environment to engage the motivation for recognition, the results of the individual's activities must be visible to other people (Malone & Lepper, 1987). This can be done in several ways: (1) the process of performing the activity may be visible, (2) the product of the activity may be visible, or (3) some other result of the activity may be visible (Malone & Lepper, 1987).

There is evidence that certain types of technology-enhanced environments provide affordances that support and engender both intrinsically and extrinsically motivated learning (e.g., Malone & Lepper, 1987; Reynolds & Harel Caperton, 2011). That said, this paper aims to describe how the use of mobile devices for classroom instruction relates to the theoretical accounts of what motivates students to learn. Accordingly, the following research question guides this study: What do elementary teachers and students perceive as the motivational affordances of using mobile devices for learning?

Methodology

Research design

Qualitative case study methodology (Creswell, 2012) was utilized in order to examine the perceived role of motivation in students' learning with mobile technology. Case studies are undertaken when educational researchers want to derive in-depth understandings of a particular phenomenon that is unique or unusual (Creswell, 2012; Merriam, 2001; Yin, 2003). This particular school is unique because of its geographical setting and the participants' technological expertise and experience. This single-case design was also collective (Stake, 1995) in that it tapped data from different sources, and it was descriptive (Yin, 2003) in that it sought to describe the natural phenomena. This case study aims to use a thick, holistic analysis to describe the perceptions of a grade 6 teacher and her students regarding the motivational affordances of using tablets for learning in their classroom.

Context of the study

This single-case study is a 3-year SSHRC-funded[1] research project on 21st century reading. This article describes the preliminary Year 1 findings of this longitudinal research study, which was carried out over a 5-month period in a sixth-grade class in a suburban Catholic elementary school in Southern Ontario, Canada. St. Martin Catholic Elementary School (pseudonym used) has a Kindergarten through eighth-grade population of approximately 400 students, with an average family income of $164 000. The school has a predominantly white, upper-to-middle class population with some diversity. The neighbourhood data related to St Martin indicates 13.9% lone parent families, and an unemployment rate of 5.6%. About 4.8% of the residents are recent immigrants, while 22% report a first language other than English or French (Rowsell, McQuirter-Scott, & Bishop, 2013).

Participants

Natasha, the grade 5 and 6 teacher participant featured in this article, is white, middle class and has been teaching for 7 years [6 of which have been within the Junior Division (Grades 4–6)]. Natasha's class contained 24 students, who ranged in age from 10 to 12 years old. There were 10 boys and 14 girls. Natasha had always been technologically adept and had some understanding of tablets before the study began, but she had not thought about using them for cross-curricular instruction. Technology use in the classroom had been previously limited to desktop computers; however, since participating in this study, Natasha's grade 6 class had regular access to ten tablets which were stored inside a locked cabinet behind Natasha's desk. A team of school district consultants, university faculty, the special education resource teacher and school administration met with Natasha regularly to provide technical, logistical and pedagogical support. Table 1 contains information (gathered through student and teacher interviews) on the demographic characteristics of the ten students who were randomly selected from Natasha's grade 5 and 6 classroom to participate in this study.

As shown in Table 1, participants included four female students and six male students. With the exception of Jeremy, the remaining ten participants owned at least one mobile device at home. The majority of participants stated that they used their mobile devices every day for playing games (apps) such as Angry Birds, accessing social networks (e.g., Facebook),

Table 1. Demographic Characteristics of Grade 5 and 6 Student Participants

Student	Grade	Gender	Student characteristics	Mobile technology ownership (home)	Mobile technology frequency and purpose of use
Samantha	5	Female	Mid-level reader	iPod, iPhone and iPad	Every day; plays games (e.g., Tap Galaxy, Cut the Rope)
Jeremy	5	Male	High-level reader	NA	NA
John	5	Male	Low-level reader	iPad	Every day; plays games (e.g., Minecraft, Angry Birds)
Mike	5	Male	Low-level reader	iPod	Every day; plays games (e.g., Angry Birds) and text messages to friends
Stephanie	6	Female	Mid-level reader	iPad and iPod	Some days; plays games (e.g., Angry Birds) and visits social networking sites daily
Kathy	6	Female	High-level reader	iPod	Every day; reads iBooks; visits social networking sites; plays word games (e.g., Bluster, Whirly Word)
Sarah	6	Female	High-level reader	iPod Touch and iPad	Some days; email, homework and research, plays games (e.g., Grammar/Word games); visits social networking sites
James	6	Male	Mid-level reader	iPad	Every day; plays games (e.g., Angry Birds)
Mark	6	Male	Low-level reader	iPad	Every other day; plays games (e.g., Zombie Farm)
Geronimo	6	Male	Low-level reader	iPod Touch	Every day; plays games (e.g., Angry Birds)

video chatting (e.g., FaceTime) and text messaging with their friends (e.g., iMessage). In fact, it was also discovered that several student participants who had access to mobile devices at home downloaded the same educational apps (e.g., Whirly Word, Bluster) that were used in Natasha's classroom.

Data collection and analysis

This project involved teacher and student interviews, a teacher blog, observational fieldwork and ecological surveys of the community. This article, however, reports only on the information contained in the teacher blog, as well as the teacher and student interviews conducted at the end of the 5-month (Year 1) study.

Student and teacher interviews

At the end of the 5 months, 15 min, semi-structured individual interviews were conducted in order to provide an in-depth understanding of the lived experiences of ten grade 6 students and their grade 6 teacher who used tablets in their classroom. Interviews were selected as a major data collection method, since it enabled the researcher to collect the participants' perceptions regarding the motivational affordances of using these mobile devices for learning. Interview questions focused on the impact of mobile technology use on student learning, motivation and engagement. Although students were more implicitly asked about this aspect of the research through questions of likes and dislikes, teacher interview questions were more specific and included questions such as 'How do you think your students' motivation to learn was impacted by their use of mobile devices in your classroom? How so?'

Teacher's blog

Natasha adopted an action research approach to this study in that she kept a research reflection blog to expand and refine her professional knowledge related to the use of mobile technology in her classroom (by semi-private, we mean that the researchers had access to the website). The inclusion of a semi-private blog as a regular reflection forum helped Natasha to consoli-

date and lift out how tablets affected her lessons. Natasha contributed to the blog regularly (e.g., once or twice a week). The blog was included in this project with the purpose of deepening the understanding or refreshing the teacher's perspective on the phenomenon. Since Natasha shared her experience and thoughts voluntarily in the discussion board without a feeling of 'being investigated', this type of data source might have included her emic issues on this project (Van Manen, 1990).

Data analysis

All interviews were audio-taped and transcribed by the researcher. Responses to the research question were triangulated from individual semi-structured teacher and student interviews and the teacher's blog. Data analyses consisted of coding and categorizing as described by Creswell (2012). The researchers coded all data independently, meeting subsequently to share individual interpretations and negotiate a shared understanding with any disagreements resolved through discussion until consensus was reached. Student and teacher data were analysed separately at first, and the results were compared to identify commonalities and differences in response patterns. After several readings of the teacher's blog posts and interview transcriptions, the researchers highlighted and coded recurring words, phrases and patterns. The codes represented categories that were in response to the research question. When the coding was complete, the codes were moderated and regrouped them into thematic clusters. The findings presented below were selected data excerpts from the interviews and blog entries that most closely represented Malone and Lepper's (1987) taxonomy of motivations for learning.

Findings

The overall research question involved: 'How do elementary students and teachers perceive their tablet use and what motivated(s) them to be a tablet user?' As such, the main focus of this study centred on students' and teacher's self-perceptions centring upon their experience with tablet use in the classroom. The findings are presented in clusters that describe the grade 6 teacher's and student participants' self-reported perceptions of the motivational affordances of using mobile technologies in the elementary classroom. Specifically, six categories emerged for the participants as the elements of mobile environments that stimulate intrinsic and extrinsic motivation, which also coincided with those of Malone and Lepper (1987). These categories were challenge; control; sensory and cognitive curiosity; competition, cooperation and recognition.

Challenge

One of the main findings of this research is that challenge and immediate feedback played a major role in making the mobile apps and games engaging, enjoyable and motivating for the students (Csikszentmihalyi, 1990; Malone & Lepper, 1987). As the students' skill level increased during the game, so did the challenges the student was faced with. Thus, flow was gradually increased over the course of the game in until either the challenge became too great (frustration) or the student's skill outpaced the challenges the game offered (boredom). This occurred to Mark when he first used the tablet; the novelty of the technology and games wore off and boredom began to set in.

> When I first used the tablet I was addicted to it, I played the games and apps on it for at least five hours a day. Now, I am only playing on the tablet for two hours because I got used to it and I beat most of the games (Mark, Grade 6 Student, Interview).

The students and teachers commented that the immediate feedback encouraged many students to keep working on difficult problems, 'There's immediate feedback and they can see where it is that they're struggling or what they need to do to correct it, and normally it gets rectified immediately' (Natasha, Grade 6 Teacher, Interview).

Performance feedback and score keeping also allowed the students to track their progress towards desired goals (to reach the end level of a game), which seemed to stimulate their intrinsic motivation (Malone & Lepper, 1987).

> The games and apps on the tablet push you a lot. For word games like Whirly Word or Bluster, the words get harder and harder and harder on each turn, and that's good because you can't just have easier words all of the time. On some of the apps, we get points and rewards which pushes you a lot. Every time you get a word you have a bar and your bar goes up. You have to try to reach your goal. (John, Grade 6, Interview)

Similarly, Natasha shared an incident where her students had about 10 min to spare before lunch, and they were each given tablets. Much to Natasha's surprise, all of her students chose the common word games they played in class during language arts instruction, including Whirly Word and Bluster. Natasha thought her students would select a non-educational game like Angry Birds. When she asked her students to explain their rationale for choosing the vocabulary games over Angry Birds, they informed her that it was 'more fun playing games they keep getting better at and learning from' (Natasha's blog entry, January 24, 2012).

Control

Multimedia presentations are more effective when the learner has the ability to interact with the presentation and work at their own pace; when learners are able to control the pace of the presentation, they learn more (Mayer, 2005). Interviews with students confirmed these findings. One of the reported benefits of using tablets in the classroom was that it allowed the students to do tasks at their own pace and placed the locus of control in their own hands. Personalization increased learner's choice where students had greater locus of control (Rudd, 2008). The following quotes illustrate that when students used mobile technology, they were able to personalize their learning experiences in many more ways than would be allowed by paper and pencil or possible during teacher-direction instruction.

> We used a How to Draw app in class, which had numerous artistic videos that anyone can use to draw a face, cartoon character, etc. This app taught us how to draw. I am not one of those people that really enjoy actually drawing, but I liked this because no matter how little your creativity and drawing skills were, it was possible that your drawing could look the exact same as the artist's in the video. So, it really helped. I think it was so easy for me because it gave you step-by-step simple instructions, you could pause the video and there was no rush. Whereas sometimes when the teacher is teaching it, you have some kids falling behind others . . . this was at your own pace. (Samantha, Grade 5 Student, Interview)

> You can read on the tablet at your own pace. You don't have to listen to the teacher talking too fast or too slow for you because if it's too fast, you won't learn anything. If it's too slow you'll just stand there listening to the stuff you already learned. I think people should read at their own pace. They don't have to let teachers read for them. Plus, we're only in grade 6 now . . . we can read by ourselves. (James, Grade 6 Student, Interview)

The smallness of the devices did not seem intimidating to the students. As evident in the following excerpt, the smallness of the technology seems to make some students feel in control, less overwhelmed and more empowered, and thus they were willing to take more risks and expend more energy (Looi *et al*., 2009). According to Natasha, she would not see this same focused activity if the students were all working on other types of technology such as a SMART Board™ which was located at the front of her class.

> My students are not all that impressed with any activity done on the SmartBoard™. I'll have the same game up there that I think they'll really like. But they don't want to do it; they want to do it on their own tablets because everyone else can see what they're doing on the Smartboard™. That's why they don't use our computers in the classroom. But when they're given that opportunity to do that game on the tablet they're happy to do it. They get very shy when it's broadcasted to the whole class, but when they're by themselves, they're ok to give it a try (Natasha, Grade 6 Teacher, Interview)

Nicole also reported that his study reported that the use of the iPod Touch catapulted students into the role of 'teacher' or 'expert' and Nicole into the role of 'novice learner' in immediate and obvious ways. The students' own perceptions of their social experience took on a new dimension when they were the technology experts in the classroom, and some of the teachers were positioned in a novice role. According to Nicole, mobile learning technologies seems to afford learners with more control over their own learning and the ability to access, create and share information across different settings (e.g., home and school; van't Hooft, 2008). As illustrated below, by making personal and relevant connections and bringing her students' own knowledge and experiences to the learning table, Nicole was able to spark her grade 6 students' interest and engagement.

> There is their (the students') world and then there is the world of school. I thought I knew a little bit about their world, but then you say to a group of grade 6 students, 'Do what you would do if you had this at home', and the whole place just lights up they're showing you things that you're going 'Oh my!' (Natasha, Grade 6 Teacher, Interview)

Sensory curiosity

The multimedia learning principle states that people learn better from pictures and words than words alone; the combination encourages active cognitive

processing and cognitive load reduction to promote deeper learning (Mayer, 2005). Unlike traditional teaching methods that may miss some of these preferred learning styles, the multimedia and interactive capabilities afforded by such mobile devices as the tablet allowed teachers to create multidimensional learning environments which catered to multiple learning styles (e.g., visual, auditory, kinaesthetic) at the same time. The tablet can assist teachers in their effort to personalize instruction according to students' preferred learning styles. The students demonstrated a stronger desire to learn when using the tablet, as they were given the opportunity to interact with information in a way that made sense to them.

> In math, we were learning about patterns and our teacher told us to try a few questions in our math books. Some of us didn't get it. But when we gave us a tablet and we used a Math app with an example of a pattern- a circle, square, circle, square, and it asked 'What's the next pattern?' We'd easily get it. It was really colourful, there were lots of charts, tables, and you could draw what's going on inside of your head on the app. And then it would give you hints and help you figure out the best way. Sometimes technology helps more than someone explaining it to you. (James, Grade 6 Student, Interview)

> There was this one app we used in language arts class.. . . some of us weren't enjoying it because it was too much information. There weren't many pictures. The appearance of the app itself was actually kind of bland. We were looking at a text on our tablets for an hour and a half. The information needs to be a little appealing and not just a plain page . . . a picture or a simple diagram with words and sounds will sometimes help a little more. (Stephanie, Grade 6 Student, Interview)

Touch screen-based devices such as the tablet seemed to provide more freedom to the students in terms of control as compared to print-based texts. The haptic/tactile technology delivered a differentiated, more interactive and personalized user experience, which enhanced their attention (Wong *et al*., 2010).

> One really big difference between reading print texts and reading on the tablet is that on the tablet screen you can increase or decrease the font size, so it's easier for you to read on the tablet. For some people with glasses, they either have to put the book far away or really up close. But instead of using all that arm power, you could just easily zoom in or zoom out on the screen. (Stephanie, Grade 6 Student, Interview)

> I did find that when I would ask my students after they worked on the iBrainstorm app, they found it better than just putting it down on paper. Using the virtual sticky notes, changing the colour of the sticky notes and moving and rearranging the notes on the screen . . . who knew such a subtle little thing could make a reading and writing activity that much more exciting?! (Natasha, Grade 6 Teacher, Interview)

Cognitive curiosity

By enabling learners to learn 'anytime, anywhere', mobile technology augments the propensity for students to engage in self-directed, informal learning beyond the classroom walls (Sharples, Taylor, & Vavoula, 2007). As shown in the excerpt below, Kathy viewed mobile technology as a tool for bridging school learning and home learning as she engaged in learning that was both spontaneous and deliberate (Sharples *et al*., 2009).

> I was just searching the app store on my own time, because I like doing that every other week to see what new games are on the charts. We were doing an advertising unit in class, and I found an app under the school section in the app store. It was a logo quiz. What you had to do was you have logos and then you had to match them with the name and then you had to say what they sell/services, and it really made me think cause it kind of helped me with my homework. I enjoyed it; I went through all the levels, too. (Kathy, Grade 6 Student, Interview)

According to Sarah, the tablet was her preferred device because of its convenience and ease-of-use.

> It's very easy to get from one app to another, and it's easier to start an Internet browser. The tablet is very convenient and very easy to understand because sometimes when you're on a laptop, there are certain things that pop up like advertisements. There's nothing that pops up on a tablet. And I really like that on the tablet, if there's a word that you don't know then you can figure out really easily what that word means. Whereas with the book, you have to read it all over again, and still don't know what it is, so you have to get up and get a dictionary, look it up, which takes a while and it's harder. (Sarah, Grade 6 Student, Interview)

Similarly, Jeremy preferred the tablet to printed materials because of the speed of access to updated and current information, which may have increased his intellectual curiosity.

> The textbooks I've noticed are not updated, they're old and most of the time our teachers say 'Oh don't look at

> that graph or diagram, it's behind 10 years' so I'd rather be reading stuff on the tablet that's recent and not stuff that's older and getting wrong information. (Jeremy, Grade 5 Student, Interview)

Students reportedly enjoyed the fact that the tablet presented them with a wealth of media choices and an instantaneous wealth of information available to them at their fingertips (Sharples, 2007). 'It's really unlimited what we can do with the tablet. For example, I really enjoy the iBooks on the tablet because you can find any book; whereas in our library, there isn't as much variety and choices' (Sarah, Grade 6 Student, Interview).

Thus, mobile technology may have played a role in cultivating students' curiosity by providing greater and easier access to a wealth of new information (Jenkins, Clinton, Purushotma, Robison, & Weigel, 2006).

Competition

Consistent with previous research (e.g., Deci & Ryan, 1985; Deci *et al.*, 1981), Natasha and her grade 6 students perceived the tablet as supporting mastery-oriented evaluation rather than performance-oriented evaluation. As illustrated below, Natasha strived to create a mobile learning environment that emphasized mastery over performance, where success was defined as increasing one's own competence rather than outperforming others (Dweck & Leggett, 1988). In this case, competition was perceived as enhancing intrinsic motivation by providing optimal challenge and ongoing feedback (Malone & Lepper, 1987).

> I directed them to Math Edge where they completed 50 questions at a time. The app gave them the amount of time it took to complete all questions, so I suggested they see if they could better their time, and all of a sudden the challenge was the reward. Every student felt successful and was eager to beat their best time. We played for 20 min and every pair improved their score by half. There was a real excitement because something hard-learning multiplication facts-became achievable and the kids recognized immediately that something important had just happened. The satisfaction of learning was much more rewarding than the Ninjas (from the Math Ninja app), and they were able to articulate that! (Natasha, Grade 6 Teacher, Blog Entry, April 15, 2012)

Cooperation

The tablet can be instrumental in creating inclusive learning environments that engage all students regardless of ability, disability, background or learning style (Wellings & Levine, 2009). This was confirmed by Natasha, who believed that the integration of mobile technology into her classroom fostered inclusion. The tablets seemed to remove the barriers to learning, put all children on a level playing field and engaged diverse learners in activities that otherwise may have been impossible or even avoided using traditional methods (Looi *et al.*, 2009).

> Before the Grade 6's started using the tablets, they couldn't get Lisa (a Grade 6 student) to work with other people because she was just refusing and wouldn't do it. She was always very anti-social. She hated all sorts of technology. But she loves the tablet, just from what she's found she can do with it. She likes it because it's more user-friendly, its smaller, she can carry it around. She recently worked on a Tunetastic video with one of the higher students in the class using the tablet. She's one of the lower students, and you wouldn't know it just from the work that they created. Her peers were like, 'Wow, you're really good at acting!' So that built her confidence, and they were showcased it in front of her class. Now wants to use the tablet for a lot of things and is looking at technology. She also hated reading and writing, she can't read very well . . . but now she has the tablet read to her, and she knows how to highlight text. So, she's really learned a lot from the other student as well, and has also taught the other student how to use the tablet and knows the features more than the brighter student. (Natasha, Grade 6 Teacher, Interview)

The use of mobile technology markedly improved learning outcomes and promoted greater motivation to persist on tasks. Students in cooperative learning groups engaged in more positive, task-oriented interaction with each other. The following quote highlights the affordances of a technology-enriched classroom where such practices as (cross-age) peer mentoring and reciprocal teaching is fostered.

> I will put a grade 5 with a grade 6 or a higher or lower level and they get so excited when they get to do that, because very rarely do they get to do things together and they want to be together all the time, so it's a good comradery. (Natasha, Grade 6 Teacher Interview)

The use of mobile technology in Natasha's classroom markedly improved student learning outcomes and promoted greater motivation to persist on tasks.

The students in cooperative learning groups seemed to engage in more positive, task-oriented interaction with each other.

Recognition

Satisfaction is necessary for learners to have positive feelings about their learning experiences and to develop continuing motivation to learn (Maehr, 1976). This means that extrinsic reinforcements, such as rewards and recognition, must be used in accordance with established principles of behaviour management (Skinner, 1968). As mentioned earlier, in order for an environment to engage and motivate the student, the results of one's activities must also be visible to other people (Malone & Lepper, 1987). The following quote highlights what transpires in Vygotsky's (1978) zone of proximal development, as students were teaching each other how to use the tablet and showcasing their creative work. According to Natasha, the mobile technology was also an outlet for some of her quieter students to overcome their shyness, become engaged, which led to improved participation.

> Yesterday my students were using this app for the first time in preparation of creating story boards for an upcoming digital comic strip they will be creating. Once completed, students presented their ideas to the class and were overjoyed to share them with everyone. I am still so pleased and surprised to see how excited and confident they are to use the tablet and its capabilities. It allows my quieter kids to have a platform to shine. (Natasha, Grade 6 Teacher, Blog Entry, April 11, 2012)

Discussion

This paper extends previous work by Malone and Lepper (1987) and applies their theoretical approach in a new context for learning in terms of how mobile devices motivate students to learn. We identified six key aspects of successful (mobile) learning systems: challenge, control, curiosity, recognition, cooperation and competition.

The importance of appropriate challenge cannot be overstated. In accordance with Malone and Lepper's (1987) taxonomy of intrinsic motivations for learning, this study discovered that such motivational aspects as optimal challenge (against oneself) and immediate feedback were incorporated into the mobile apps used in Natasha's grade 5 and 6 classroom. This type of learning honours choice in activities, allows for self-paced learning and publicly acknowledges achievement by providing almost instant feedback. In traditional classrooms where quizzes and assignments are graded by hand, students may not find out how they have done until long after a concept has been taught; consequently, some students may lose interest and have little incentive to complete these activities (Brophy, 2010). On the contrary, the quizzes and games available on the mobile apps provided opportunities for repeated student self-assessment and instant feedback (correct or incorrect answer along with their completion time). Natasha and a majority of student participants found the instant feedback to student responses was useful and an especially appealing form of incentive for the students which encouraged many of them to keep working on progressively more difficult problems and scaffolded challenges provided by the mobile apps. Natasha welcomed the fact that she could track each student's progress, understand the strengths and weaknesses of individual students, and refine their teaching. The interactivity and automatic feedback features of these tablet apps may have also contributed to heightened cognitive curiosity and students' voluntary use of these same educational apps at home.

According to Howard Gardner (1999), seven kinds of intelligence allow seven ways to teach, rather than one. Natasha's mobile multimodal classroom built on Gardner's insight by letting students learn at their own pace (learner control) and enhanced their sensory curiosity (Malone & Lepper, 1987). Alongside this, the tablets provided built-in means of differentiated learning experiences that otherwise may have been impossible using monomodal, traditional methods (Looi *et al*., 2009). Videos and iBooks available on the tablet gave students control over aspects of their learning where they can listen and view the instructional information repeatedly at their own pace (McKinney, Dyck, & Luber, 2009).

Traxler (2007) states that 'mobile learning delivers learning to the learner when and where they want it' (p. 7). In other words, mobile device use augments the propensity for students to engage in self-directed learning and stimulate their cognitive curiosity beyond the classroom walls (Traxler, 2007). Consistent with previous research (e.g., Sharples *et al*., 2009; Traxler, 2007), the findings of this research illustrate how mobile devices were viewed as a tool for bridging school learning and home learning. The portability and convenience of mobile devices emerged as determining

factors in students' decisions to use them actively for leisure and social networking purposes, as well as formally and informally to support their schoolwork (Low & O'Connell, 2006; Malone & Lepper, 1987).

Numerous studies on computer-supported cooperative learning have also demonstrated positive effects on the amount and quality of social interaction (e.g., Fishman & Gomez, 1997). The concept of cooperative learning is based on a social learning theory that students are more likely to possess high self-efficacy, confidence, and have higher motivation to complete a task when they know they will have assistance from their peers (Cheng & Ku, 2009). Cooperative learning was enhanced in Natasha's grade 5 and 6 classroom by the use of mobile devices. Consistent with previous research, Natasha's students worked cooperatively with technology and some even held more positive attitudes, improved intergroup relations and increased acceptance of academically challenged peers (Cheng & Ku, 2009). Cooperative learning resulted in supportiveness for partners and increase in helping behaviours. These practices also helped many students overcome their shyness and led to improved participation. The use of the tablet removed the barriers to learning, put all children on a level playing field and engaged these diverse learners in activities that otherwise may have been impossible or even avoided using traditional methods (Looi *et al*., 2009).

Vygotsky's (1978) zone of proximal development (ZPD) can also be used to explain this finding. According to Vygotsky's social constructivist theory, learning is a socio-culturally mediated and collaborative process that occurs through interactions and sharing with others, including teachers, parents and other learners (Vygotsky, 1978). More specifically, Vygotsky's theory of the ZPD which accentuates the supportive guidance of mentors and 'experts' (usually but not exclusively teachers), as they enable the novice learner to achieve successively more complex skill, understanding, and ultimately independent competence. However, rather than focus on the adult as the more capable other who mentors the younger student, this finding looks at the ZPD from the vantage point of the student being the more capable other. Natasha reported that the use of the tablets in her classroom catapulted students into the role of 'teacher' or 'expert' and teachers into the role of 'novice learner' in immediate and obvious ways. The students' own perceptions of their social experience took on a new dimension when they were the technology experts in the classroom, and Natasha was positioned in a novice role. Natasha was pleasantly relieved to find that she was participating in reciprocal teaching methods as her students were teaching her about the capabilities of the tablet and some apps. These findings highlight the shifting dynamics in a technology-enriched classroom where such practices as (cross-age) peer mentoring, reciprocal teaching and student-teacher role reversals are fostered.

Whether competing for grades in classrooms or trophies in athletic contests, individuals may view their behaviour as externally controlled and experience pressure to win (direct competition; Malone & Lepper, 1987). On the other hand, indirect competition can lead individuals to view activities as challenging and opportunities for immediate feedback, making competition attractive to some individuals. Consistent with earlier findings (e.g., Reeve & Deci, 1996), the latter form of competition seemed to play a greater role in influencing students' motivation to learn with the tablet. The mobile apps used in Natasha's grade 5 and 6 classroom provided a personalized learning experience and fostered indirect competition where students strived to beat their own previous best performance (mastery orientation) by completing a series of self-selected quests.

Implications

Although much has been said about the inherent motivating qualities of mobile technology; generally, there is a paucity of research that directly reflects the connection between mobile technology use and the role of motivation in learning with mobile technology. What are the motivational affordances of using these mobile devices for learning? This is the pivotal question we sought to address in this paper. The results of this study indicate that Malone and Lepper's (1987) taxonomy of intrinsic motivations for learning can be applied to mobile learning. The taxonomy described in this paper provides a model of how to design intrinsically motivating mobile (and non-mobile) learning environments.

In this study, students identified teamwork and opportunities to work with other students as important motivational factors. It is in these venues that individuals can share thoughts and ideas and become active participants in a digital society and develop the skills of cooperation and collaboration. As with any learning experience, providing a scaffold experience can help

develop the individual. Teachers and mobile developers may be able to facilitate the development of the interpersonal skills required for teamwork by grouping students heterogeneously – mixing students of different ability levels and grades – and suggesting roles for group members (e.g., typer, recorder, reader, time keeper, checker for understanding), as well as providing students time to analyse and discuss how effectively they are working together and how they may work more effectively together in the future.

Especially for struggling learners, there is a need for a 360-degree approach to learning, in which the experiences that underlie in-school learning are aligned with those in afterschool and home settings (Shore, 2008). Anchoring instructional lessons for the 21st century classroom can be viewed as an intimidating task, especially considering that many students are more technologically engaged outside of school than inside their classrooms (Walker & Shepard, 2011). One way to harness student motivation is by allowing and encouraging students to utilize their technical knowledge and experiences and allow them to engage in self-directed learning activities.

When compared to the wide range of technologies at our disposal, the highly personalized nature of mobile devices provides an excellent platform for the development of personalized, learner-centric educational experiences marked by flexibility, customization, collaboration, active participation and co-creation (Looi *et al*., 2009). Most importantly, mobile learning gels with constructivist principles where multiple learning pathways and scaffolding activities can be constructed, and knowledge can be explored in multiple ways and in multiple contexts that best resonates with the needs of the users (Looi *et al*., 2009). When learning with mobile devices is carefully designed, it is possible to create more collaborative and participatory learning experiences while increasing pupil engagement and mastery of important concepts (West, 2012).

The theoretical approaches that appear to be most relevant to mobile learning are those that involve learner control and challenge by setting an appropriate level of complexity, provoke their user's curiosity, and allow them to engage in active learning conversations. Teachers should also help children develop a mastery orientation to learning, which holds great promise for improving academic achievement. To influence children's mastery orientation towards learning, teachers should provide tasks that are meaningful to children, given their interests and environments; place the emphasis on mastery of the skill, rather than performance; and focusing on the value of learning (and what can be gained) in formal and informal evaluations. Building these concepts into mobile apps and activities designed for mobile learning will support and motivate future learners.

Limitations of the study

This case study has some limitations that need to be considered when interpreting the findings above. Limitations from this study stem from its scope, particularly the size and composition of the sample population and lack of a control group. There is a need for future empirical research with a larger and more varied sample to clarify the present findings and examine the connection between motivation and learning outcomes.

As with any technology that is introduced in a learning environment, there is always a novelty effect (Krendl & Clark, 1994). Students tend to be more motivated to use a new piece of technology for learning because it is new. The implication of this criticism is that the positive outcomes learning from the new medium, having more positive attitudes about learning will tend to decline as the technology becomes more familiar and its novelty wears off. Future studies on mobile technology should include more longitudinal research to determine whether motivation to use mobile technology and levels of mobile technology use change over a longer period of time. Designers and teachers will also need to explore what steps can be taken to combat the 'novelty effect' in order to achieve sustained motivation, one of which may include creating difficulty levels that will keep the player in flow for as long as possible.

In sum, a better understanding of the nature of intrinsic and extrinsic motivation and the ability to gauge students' motivation while interacting with mobile technology-supported learning environments promises to contribute to the design of more effective educational programs and thus ultimately to higher educational performance.

Acknowledgement

We wish to acknowledge the Social Science Humanities Research Centre of Canada (SSHRC) whose funding support has made this international study possible.

Note

[1]SSHRC is a Social Science and Humanities Grant funded by the Canadian government conducting with an international group: Dr. Jennifer Rowsell, Tiffany Gallagher, Ruther McQuirter-Scott, Mary Saudelli, Katia Ciampa (Brock University); Doug Fisher and Diane Lapp (San Diego State University); Alyson Simpson (University of Sydney) and Maureen Walsh (Australian Catholic University).

References

Beck, J. C., & Wade, M. (2006). *The kids are alright: How the gamer generation is changing the workplace*. Boston, MA: Harvard Business School.

Beishuizen, J. (2008). Does a community of learners foster self-regulated learning? *Technology, Pedagogy and Education*, *17*, 183–193.

Brophy, J. (2010). *Motivating students to learn* (3rd ed.). New York, NY: Taylor & Francis.

Cheng, Y. C., & Ku, H. Y. (2009). An investigation of the effects of reciprocal peer tutoring. *Computers in Human Behavior*, *25*, 40–49.

Chou, C. C., Block, L., & Jesness, R. (2012). A case study of mobile learning pilot project in K-12 schools. *Journal of Educational Technology Development and Exchange*, *5*(2), 11–26.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson Education.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper Collins.

Deci, E. L., Betley, G., Kahle, J., Abrams, L., & Porac, J. (1981). When trying to win: Competition and intrinsic motivation. *Personality and Social Psychology Bulletin*, *7*, 79–83.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.

Dornyei, Z. (2000). Motivation in action: Towards a process-oriented conceptualization of student motivation. *British Journal of Educational Psychology*, *70*, 519–538.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.

Evans, M., & Johri, A. (2008). Facilitating guided participation through mobile technologies. *Journal of Computing in Higher Education*, *20*, 92–105.

Fishman, B., & Gomez, L. (1997). How activities foster CMC tool use in classrooms. In R. Hall, N. Miyake, & N. Enyedy (Eds.), *Computer support for collaborative learning* (pp. 37–44). Toronto, CA: Erlbaum.

Gambrell, L. B. (1996). Creating classroom cultures that foster reading motivation. *The Reading Teacher*, *50*(1), 14–25.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York, NY: Basic Books.

van't Hooft, M. (2008). Envisioning the future of education: Learning while mobile. *Learning and Leading with Technology*, *35*(6), 12–16.

Jenkins, H., Clinton, K., Purushotma, R., Robison, A. J., & Weigel, M. (2006). *Confronting the challenges of participatory culture: Media education of the 21st century*. Chicago, IL: The MacArthur Foundation.

Johnson, D. W., & Johnson, R. (2003). Student motivation in cooperative groups: Social interdependence theory. In R. Gillies & A. Ashman (Eds.), *Cooperative learning: The social and intellectual outcomes of learning in groups* (pp. 136–176). New York, NY: Routledge Falmer.

Joiner, R., Nethercott, J., Hull, R., & Reid, J. (2006). Designing educational experiences using ubiquitous technology. *Computers in Human Behavior*, *22*, 67–76.

Klein, J. D., & Keller, J. M. (1990). Influence of student ability, locus of control, and type of instructional control on performance and confidence. *Journal of Educational Research*, *83*(3), 140–146.

Krendl, K. A., & Clark, G. (1994). The impact of computers on learning: Research on in-school and out-of-school settings. *Journal of Computing in Higher Education*, *5*, 85–112.

Lepper, M. R., & Cordova, D. I. (1992). A desire to be taught: Instructional consequences of intrinsic motivation. *Motivation & Emotion*, *16*, 187–208.

Liu, M., Toprac, P., & Yuen, T. (2009). What factors make a multimedia learning environment engaging: A case study. In R. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 173–192). Hershey, PA: Idea Group Inc.

Looi, C. K., Wong, L., Hyo-Jeong, S., Seow, P., Toh, Y., Chen, W., . . . Norris, C. (2009). Anatomy of a mobilized lesson: Learning my way. *Computers & Education*, *53*, 1120–1132.

Low, L., & O'Connell, M. (2006). *Learner-centric design of digital mobile learning*. Paper presented at Learning on the Move, Brisbane, Australia.

Maehr, M. L. (1976). Continuing motivation: An analysis of a seldom considered educational outcome. *Review of Educational Research*, *46*, 443–462.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: III. Conative and affective process analyses* (pp. 223–253). Hillsdale, NJ: Erlbaum.

Mayer, R. E. (2005). *Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.

McKinney, D., Dyck, J. L., & Luber, E. S. (2009). iTunes university and the classroom: Can podcasts replace professors? *Computers and Education*, *52*, 617–623.

Merriam, S. B. (2001). *Qualitative research and case study applications in education: Revised and expanded from case study research in education* (2nd ed.). San Francisco, CA: Jossey-Bass.

Norris, C., & Soloway, E. (2008). An instructional model that exploits pervasive computing. In J. Voogt & G. Knezek (Eds.), *International handbook of information technology in primary and secondary education* (pp. 849–860). New York, NY: Springer.

Quinn, C. N. (1994). Designing educational computer games. In K. Beattie, C. McNaught, & S. Wills (Eds.), *Interactive multimedia in university education: Designing for change in teaching and learning* (pp. 45–57). Amsterdam, The Netherland: Elsevier.

Reeve, J., & Deci, E. L. (1996). Elements of the competitive situation that affect intrinsic motivation. *Personality and Social Psychology Bulletin*, *22*, 24–33.

Reynolds, R., & Harel Caperton, I. (2011). Contrasts in student engagement, meaning-making, dislikes, and challenges in a discovery-based program of game design learning. *Journal of Educational Technology Research and Development*, *59*, 267–289.

Rowsell, J., McQuirter-Scott, R., & Bishop, A. (2013). Tablets as placed resources: Community literacies in global domains. *Language Arts*, *90*, 351–360. Special issue edited by R. Rogers.

Rudd, T. (2008). *Learning spaces and personalisation workshop outcomes*. Bristol, UK: NESTA Futurelab. Retrieved from http://www.futurelab.org.uk/resources/documents/event_presentations/Learning_Spaces_and_Personalisation_workshop.pdf

Shaffer, D. W. (2006). *How computer games help children learn*. New York, NY: Palgrave Macmillan.

Sharples, M. (Ed.). (2007). *Big issues in mobile learning: Report of a workshop by the Kaleidoscope Network of Excellence Mobile Learning Initiative*. Nottingham, United Kingdom: Univeristy of Nottingham, Learning Sciences Research Institute.

Sharples, M., Arnedillo-Sánchez, I., Milrad, M., & Vavoula, G. (2009). Mobile learning: Small devices, big issues. In S. Ludvigsen, N. Balacheff, T. D. Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning: Principles and products* (pp. 233–249). Berlin, Germany: Springer-Verlag.

Sharples, M., Taylor, J., & Vavoula, G. (2007). A theory of learning for the mobile age. In R. Andrews & C. Haythornthwaite (Eds.), *The Sage handbook of e-learning research* (pp. 221–247). London, UK: Sage.

Shore, R. (2008). *The power of Pow! Wham!: Children, digital media & our nation's future. Three challenges for the coming decade*. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.

Shuler, C. (2009). *Pockets of potential: Using mobile technologies to promote children's learning*. New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.

Skinner, B. (1968). *The technology of teaching*. New York, NY: Appleton-Century-Crofts.

Stake, R. E. (1995). *The art of case study research*. London, UK: Sage.

Traxler, J. (2007). Defining, discussing, and evaluating mobile learning: The moving finger writes and having writ. . . . *The International Review of Research in Open and Distance Learning*, *8*(2), 1–12.

Turner, J. C. (1995). The influence of classroom contexts on young children's motivation for literacy. *Reading Research Quarterly*, *30*, 410–441.

Valk, J. H., Rashid, A., & Elder, L. (2010). Using mobile phones to improve educational outcomes: An analysis of evidence from Asia. *The International Review of Research in Open and Distance Learning*, *11*, 117–140.

Van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. London, ON: Althouse.

Vogel, D., Kennedy, D. M., & Kwok, R. (2009). Does using mobile device applications lead to learning? *Journal of Interactive Learning Research*, *20*, 469–485.

Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Walker, L. R., & Shepard, M. F. (2011). Phenomenological investigation of elementary school teachers who successfully integrated instructional technology into the curriculum. *Journal of Educational Research and Practice*, *1*(1), 23–35.

Wellings, J., & Levine, M. H. (2009). *The digital promise: Transforming learning with innovative uses of technology*. New York, NY: Joan Ganz Cooney Center at Sesame Workshop.

West, D. M. (2012). *Digital schools: How technology can transform education*. Washington, DC: Brookings Institution Press.

Wong, C. Y., Chu, K., Khong, C. W., & Lim, T. Y. (2010). *Evaluating playability on haptic user interface for mobile gaming*. Paper presented at the Information Technology International Symposium, 15–17 June, Kuala Lumpur, Malaysia.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Worksheet 2- Research Design and Methodology

1. Research Paradigm Choice

You need to consider the view of 'reality' which best suits your research topic, i.e. decide on whether you are taking a 'objective' or 'subjective' ontological assumption. From there you can determine whether you would be following either a Positivistic or Phenomenological approach.

Select ONE from Below that best fits your Research

	Worldview- Subjective	**Worldview-Objective**
Ontology	Phenomena and their meanings are continually being accomplished by social actors.	Phenomena and their meanings have an existence that is independent of social actors i.e. beyond their influence.
Epistemology	Believes that knowledge is based on the perception of the individuals	Phenomena which are observable and measurable can be regarded as valid knowledge
Paradigm/Philosophy Choice	Phenomenology	Positivism
Research Approach	Qualitative	Quantitative
Role of Theory	Inductive	Deductive
YOUR CHOICE	☐	☐

2. Research Methodology

Your choice will be largely determined by the research situation/context, area of investigation and the research paradigm chosen earlier. You have an extensive list to choose from: experiment, case study, action research, survey, grounded theory, ethnography.

Select ONE from Below that best fits your Research

Methodology	**Description**	**Choice**	**Idea for Action** **(only for selected methodology)**
Case Study	Commonly used to illustrate or understand a problem or indicate good practice. It is an extensive examination of a single instance of a phenomenon of interest. It focuses on understanding the dynamics present within a single setting	☐	

Action Research	An approach which assumes the social world is constantly changing and the researcher and the research itself are part of this change. The research will make a change and measure the results	☐	
Survey	Provides a quantitative or numeric description of trends, attitudes or opinions. leads to general inferences about a population from a sample of the population	☐	
Grounded Theory	To move beyond description of a phenomenon, to generate or discover a theory, The development of the theory might help explain practice.	☐	

3. Data Collection

You need to give details of the techniques that will be used for actually collecting the data. Choices include: questionnaire, observation, interview, focus group etc. What is important is that for the methods chosen, you must specify exactly **how it will be applied**.

Fill out table below with details of YOUR data collection

Objectives	Data Required (What)	Source (Who)	Method to Be used (HOW)	Procedure (Detailed HOW)
1.				
2.				
3.				

GYM Questionnaire

1. Are you male or female (please tick)

☐ Male ☐ Female

2. How old are you?

-------- Years

3. Which of the following best describes your main reason for going to the Gym? (please tick one only)

☐ Relaxation ☐ Maintain or Improve Fitness ☐ Lose Weight ☐ Meet Others ☐ Build Strengths

4. When you go to the gym, how often do you use the cardiovascular equipment? (please tick one only)

☐ Always ☐ Usually ☐ Rarely ☐ Never

5. When you go to the gym, how often do you use the weights machine? (please tick one only)

☐ Always ☐ Usually ☐ Rarely ☐ Never

6. How frequently do you usually do to the gym?

☐ Every day ☐ 4-6 days a week ☐ 2 or 3 days a week ☐ Once a week ☐ 2 or 3 times a month

☐ Once a month ☐ Less than once a month

7. Are you usually accompanied when you go to the gym or do you usually go on your own?

☐ On my own ☐ With a friend ☐ With a partner/spouse

8. Do you have sources of regular exercise other than the gym?

☐ Yes ☐ No

9. If you have replied YES to question 8, please indicate the main source of regular exercise in the last six months from the list below.

☐ Sport ☐ Cycling ☐ Jogging ☐ Long Walks

10. During the last visit to the gym, how many minutes did you spend on the cardiovascular equipment?

------ minutes

11. During the last visit to the gym, how many minutes did you spend on the Weights Machine?

------ minutes

12. During the last visit to the gym, how many minutes did you spend on other activities?

------ minutes