An analysis of the teaching of introductory statistics at university in “context”

by

Helen Harth

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University

September 2018

Abstract

In this research study I explore the teaching of introductory inferential statistics to non-statistics undergraduates. My main aim in this work is a characterisation of teaching practice in the context of two introductory statistics university modules, one aimed at Psychology students and another at Engineering students from the perspective of the lecturers.

In the pilot study, I investigated lecturers' beliefs about intended statistics curricula at university. The study used repertory grid interviews with twenty statistical methods lecturers. Qualitative and quantitative data analysis revealed that lecturers conceptualised the intended curricula around three themes: (1) teaching of statistics with "context", (2) teaching the statistical process components, and (3) student learning.

The main study focused on the teaching of statistics on two introductory modules. Observational and interview data was interpreted at the macro and micro levels of analysis using sociocultural theory as a theoretical lens and applying a grounded analytical approach. Introductory statistics modules are taught in a range of disciplines, including Psychology and Engineering. Previous research shows that some students find statistics very difficult and challenging. The two lecturers, although approached the teaching of statistics very differently, had a deep concern for their students' learning. The first lecturer, a Psychologist, approached the teaching of statistics in a 'philosophical' way meaning that the explanations were non-mathematical and there was a sequence of cases or "contexts" which the lecturer taught in different ways throughout the module. The second lecturer, a Mathematician, taught a 'typical' statistics module consisting of the mathematical underpinnings of statistical models through a sequence of statistical theory and calculations.

Through this research, I provided representations into the lecturers' beliefs, intentions and strategies in relation to their teaching. The application of the sociocultural lens with a grounded analytical approach enabled me to conceptualise the lecturers' teaching actions and present a model of teaching statistics in *context*.

Acknowledgements

Throughout my PhD research studies many people have helped me on my journey. I would particularly like to thank my supervisors, Professor Barbara Jaworski and Professor Carol Robinson who with wise words and encouragement have ensured that I continued when I thought of giving up. I would also like to thank Dr Ian Jones and Dr Irene Biza for their guidance and encouragement with my pilot study. I would like to thank Dr Paul Hernandez-Martinez and Dr Stepanie Treffert-Thomas for stimulating conversations about teaching and learning. I would like to thank staff and PhD students in the Mathematics Education Centre for invaluable feedback, comments and advice over the years. Most of all I would like thank the twenty lecturers of statistics for their openness, kindness and time without which this study would not have been possible. Last but not least I would like to thank my family and friends for their patience and affection. I could not have done this without you.

Contents

List of Figures

List of Tables

1 Introduction

The present study is an investigation into the teaching of statistics at university to non-statistician students, i.e. as a 'service' subject. The research is based on two lecturers' teaching of two statistics modules to large groups of psychology and engineering undergraduate students at a UK university. My focus is on the teaching of statistics to include lecturers' teaching outcomes, planning of lectures, face-to-face delivery of teaching and reflections about their teaching.

This introductory chapter summarises the main aspects of the thesis, including the main topic of my study, statistics and statistics education. I also define the meaning of the term 'context' contained in the title *An analysis of the teaching of introductory statistics at university in 'context'*. Next, I identify the purpose of the research and the research questions, the research process and the university of the study. I conclude with an overview of the structure of the thesis.

Chapter contents

1.1 Statistics and statistics education

Statistics is seen as a 'a general intellectual method that applies wherever data, variation, and chance appear' (Moore, 1998, p. 1254). Moreover, statistical 'data are not just numbers, they are numbers with a context' (Cobb and Moore, 1997, p. 801). In statistical data analysis, problems gain intellectual substance from the interplay between mathematical content, pattern and *context* (or story). For this reason, Cobb and Moore likened statistics to poetry 'where pattern and context are inseparable' (p. 803), setting it apart from mathematics and probability.

As a discipline, important changes have taken place in statistics from 1977, when Tukey introduced the systematic treatment of data in *Exploratory Data Analysis* (EDA) (Tukey, 1977), to 1997 when the term "Big Data" first appeared in print (Champkin, 2014). The growing importance of statistics to students' future lives and professions, aided by the increased availability of statistical software, justifies the continuous increase in students learning statistics in 'service' courses at university (Garfield and Ben-Zvi, 2008).

Changes in statistics have lead to changes in statistics education. Defining the key statistical skills and goals for learning statistics and identifying teaching practices that stimulate students' learning are key issues in statistics education. Statistical literacy, reasoning and thinking are considered key goals of teaching statistics. Gal (2002) defined statistical literacy to refer to people's ability to interpret, critically evaluate and communicate statistical information, data-related arguments and stochastic (statistical and probabilistic) phenomena. Statistical reasoning is about how people reason with 'big statistical ideas' (e.g. data, statistical models, distribution, centre, variation and so forth), interpret and represent data, connect concepts and combine ideas about data and chance and finally make inferences and interpret statistical results (Garfield and Chance, 2000; Garfield and Ben-Zvi, 2008). Statistical thinking is considered a type of expert thinking, but its core concepts have evolved and expanded since the beginning of the twentieth century (Marriott, 2014). Kahneman (2012) includes core concepts such as expectation, variance, distribution, probability, risk, correlation, data, visualisation and cognition in defining statistical thinking.

The statistics education literature however is less clear about the differences between statistical literacy, reasoning and thinking (Ben-Zvi and Garfield, 2004). These key goals for statistics education also require considerable effort to teach effectively, i.e. lead to student learning (Garfield and Ahlgren, 1988; Ben-Zvi and Garfield, 2004). New curricula and effective pedagogies to develop statistical literacy, reasoning or thinking have been considered to be key issues in introductory statistics in higher education (Moore, 1997). At tertiary level, 'statistics' includes descriptive statistics, exploratory data analysis and inferential statistics. As described earlier, 'statistics' may also be about statistical thinking, reasoning and literacy. Research in statistics education has focused on students' difficulties with statistics, the effects of implementing new types of teaching on students' learning and evaluations of curriculum designs. Relatively little research exists however on the practical implementation of teaching from a lecturer's perspective.

1.2 Macro-micro context

My main aim in this study is to characterise the teaching of statistics in *context*, i.e. the situation within which the teaching of undergraduate statistics happens, and that can help explain it. I gain insights into the teaching of statistics by embedding my study in a socio-cultural theory perspective and studying the *context* in which the teaching activities took place alongside two dimensions, *macro* and *micro* within a research design that includes a Pilot and a Main study.

This study follows on from previous work within mathematics education which sought to identify relationships and connections between the conceptual understanding of mathematics and broader issues with educational systems (curriculum or educational practices), social spaces, national economic and political systems (Jaworski and Potari, 2009). Rather than considering the macro and micro as dichotomies, in this study I provide an account of pedagogic practice in which large-scale macro factors are integrated with micro-levels of analysis (Popkewitz, 1998). Macro necessarily includes micro and is also a continuum between macro and micro (Hammersley, 1993b; Jaworski and Potari, 2009). Similarly, Lerman (2001) considers that "an integrated account, one that brings the macro and micro together, one that enables us to examine how social forces such as a liberal-progressive position, affect the development of particular forms of mathematical thinking" (p.89). My research therefore aims to focus on integrating wider socio-cultural factors with aspects of the teaching activity, the setting, the individuals and groups in the setting.

In order to locate the teaching activities contextually, within a range of macro and micro conditions in which activity is embedded, and the connections and relationships between them, as practical means for helping my analysis, I used two similar depictions of context. First, the conditional/consequential matrix developed by Corbin and Strauss (1990, 2008) as a coding device to show the interections of macro and micro conditions/consequences on actions. Second, I was guided by the sociocultural representations of various interpersonal-relationships in proximal "contexts" as nested and weaving together (Wertsch, 2000; Cole and Gajdamaschko, 2010), as shown in Figure 1.1.

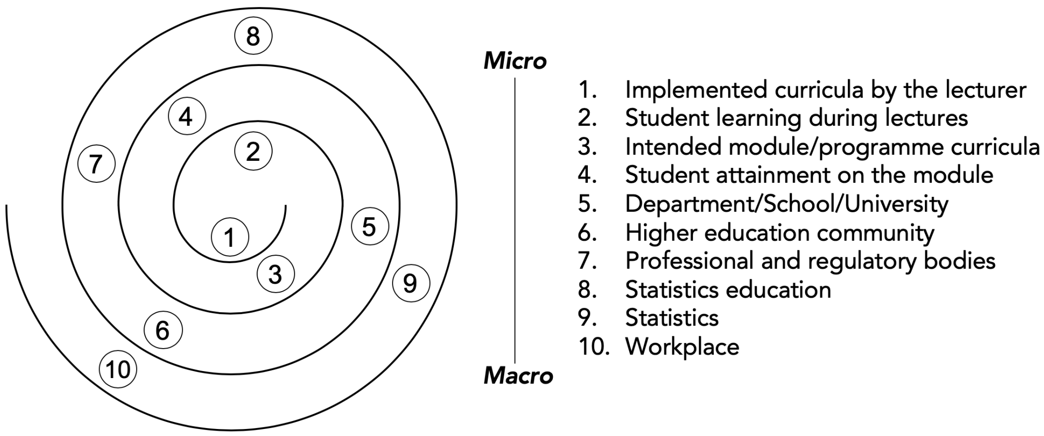


Figure 1.1: A definition of context

At the *macro*-level, I considered the ways in which statistics lecturers view the nature of statistics, the statistics curricula, their students, and the teaching and learning of statistics. In my Pilot study therefore, I carried out interviews with twenty statistics lecturers about "what" they would like their students to learn and "why". In addition, to be able to characterise the teaching of statistics in context in further depth, in my Main study, I observed two lecturers' teaching of statistics on two modules. To characterise the two lecturers' teaching at a *macro*-level in further depth, I used the two lecturers' reflections of their intentions for teaching statistics in interviews, as well as documentation of broader institutional practices.

At a *micro*-level, using observations of teaching on the two modules and interview data with the two lecturers (in the Pilot and Main studies), I investigated how the lecturers implement the statistics curricula in their modules. In my analysis, I extended the focus on lecturing tasks to address wider complexities of lecturing of statistics in macro-social-cultural setting.

In this study, I apply a grounded analytical approach to data collection and analysis and use a sociocultural theoretical lens to further refine and integrate the macro-micro analyses.

1.3 Purpose and research questions

The purpose of my study was to *characterise practices and processes in the teaching of non-specialist ('in service') introductory statistics modules or courses at university from the perspective of the lecturers.*

My general research question is as follows.

> *What insights can I gain about the teaching of statistics at university by studying the teaching activity at two levels of context, the macro and micro?*

The macro-analysis focused on intended curricula, "what" lecturers plan for their students to learn and "why". The micro-analysis focused on implemented curricula, "how" lecturers use learning resources (examples, problems or exercises) in lectures and "why".

1.4 The research process

In my Pilot study, I aimed to explore the teaching and learning of statistics at university in a range of disciplines: Engineering, Economics, Psychology, or Geography. I wished to explore what content lecturers focused on in their modules, what were the challenges and barriers to learning statistics at university and what teaching approaches lecturers preferred. In these initial stages, I planned to carry out interviews with lecturers in statistics across a variety of departments and institutions. To further understand the teaching and learning of statistics, I also carried out preliminary observations of lectures and laboratory sessions so that I could define in more detail what I wished to achieve.

This Pilot study phase helped me select the methods of enquiry for my Main study, observations and interviews. Since I was guided by an *interpretative paradigm* that involved the development of rich descriptions of the teaching I was observing, my data sources needed to be of a qualitative nature. I considered that using sociocultural theory as a theoretical lens and applying a grounded analytical approach was appropriate for building interpretations about teaching statistics at two levels of context, macro and micro (Charmaz, 2006; Corbin and Strauss, 2008; Charmaz, 2014).

My research process consisted of the following four main ***stages***:

1. *Identify a specific challenge.*
   1.1. Literature review.
   1.2. Plan the Pilot study.
2. *Pilot studies.*
   2.1. *Pilot 1:* repertory grid interviews (design 1) with nine statistics lecturers.
   2.2. *Pilot 2:* repertory grid interviews (design 2) with eleven statistics lecturers.
   2.3. *Pilot 3:* observations of teaching (two lectures and two laboratory sessions) of one module (memo-writing).
   2.4. Data analysis of interviews (produce 'rough' conceptual framework).
3. *Main study.*
   3.1. *Methodology.*
      3.1.1. Research paradigm, theoretical perspectives and research design.
      3.1.2. Select modules, collect observational data and initial data analysis (concepts and categories, observation notes and memo-writing after each observation).
      3.1.3. Post-module interviews with three lecturers and continue data analysis.
   3.2. *Data analysis 1.*
      3.2.1. Use grounded analytical approach.
      3.2.2. Coding and comparing (two modules of the three modules investigated; interviews with lecturers in the pilot and main studies).
      3.2.3. Refining concepts.
      3.2.4. Analysing data for micro-macro context, process and theoretical integration.
      3.2.5. Memo-writing.
   3.3. *Data analysis 2.*
      3.3.1. Use sociocultural theory as a theoretical lens and apply grounded analytical approach to data analysis.
      3.3.2. Merge the results from the pilot and main study analyses.
4. *Interpret and integrate the findings.*
   4.1. Consider how the characterisation of teaching of statistics at university in context contributes to new knowledge and understanding of teaching.

1.5 Study participants

I carried out the data collection at a university in the UK that has a tradition in providing courses that focus on transferable knowledge and skills. The university ranks highly in engineering and technology and it offers its students state-of-the-art sporting facilities and training. Around a quarter of its annual income comes from research grants and contracts. Statistics modules are taught by lecturers from across disciples, including the Mathematics department.

In my pilot study, I interviewed twenty lecturers that offered statistical modules in a variety of departments. I used the repertory-grid technique to focus on the lecturers' learning intentions for their students. Of these twenty lecturers, two taught at different universities in the UK.

Three of the lecturers who participated in the Pilot study also agreed to take part in my main study: a mathematician teaching statistics on an engineering course, a psychologist teaching statistics to psychologists and an engineer teaching on an engineering course. Following initial data analysis, I included two lecturers in my main data analysis, the mathematician and the psychologist. The two modules included in the Main Study analysis had broadly similar content in inferential statistics. The third module excluded from the data analysis was a second module in inferential statistics and thus taught at a more advanced level. The observations of teaching in the main study took place over one term (11 weeks).

The psychology lecturer owned the module and had been teaching it for five years. During the data collection, it became clear that the psychology lecturer had a deep understanding of the type of student cohort he was teaching and had considered carefully their prior mathematical and statistical background as well as the more general aims of the psychology course. A level mathematics was not a requirement for admission on this psychology course, although it was recommended.

At the time of the study, the mathematician was teaching the statistics module for the first time and was using teaching resources produced by a previous lecturer. The engineering students were expected to have completed A level mathematics prior to entry onto the course.

My focus on two different modules and the pilot interviews data allowed me to compare and contrast different contexts in order to arrive at a characterisation of the teaching of statistics in context.

1.6 The structure of the thesis

To characterise the teaching of statistics, I used data from three main sources: the pilot interviews, the main study and the post-module interviews. In Table 1.1, I summarise the interdependence between the four research processes and the three main data sources. I used the pilot study to select the modules that were part of my main study and to refine the coding system taking a grounded analytical approach. In this section, I provide a summary of the structure of the thesis.

In this chapter, **Chapter 1**, I introduce and provide background for my study, including a definition of statistics and statistics eduction, the research questions, purpose and structure of the thesis.

In **Chapter 2** (page 9), I summarise the research literature relevant to the teaching of inferential statistics at university, the main topic of my study. Based on my literature review, I define the general topic area for my study and research questions for the Pilot study.

The theoretical basis, methodology and methods for data analysis of the Pilot study are depicted in **Chapter 3** (page 68).

My main study consisted of observations of teaching and interviews with lecturers. **Chapter 4** (page 102) explains the methodology for the Main study. This included my research paradigm, theoretical perspectives and the research design.

The data analysis and findings of the main study are in **Chapter 5** (page 149). Within the interpretative paradigm, I use Vygotskian sociocultural theory of teaching and learning as a theoretical lens and apply a grounded analytical approach to analysing and interpreting data from two introductory statistics modules taught by two statistics lecturers. My data analysis looks at both macro and micro levels of context with the aim to characterise two lecturers' teaching of statistics in context.

Finally, in **Chapter 6** (page 216), I integrate the findings from my three studies to offer a characterisation of the teaching of statistics in context, with possible implications for statistics lecturers and research of university teaching. I also discuss some of the opportunities offered by the use of a qualitative enquiry in characterising teaching and areas which deserve further research.

Table 1.1: Data collection and analysis schedule

Research process	**A. Pilot study**	→	**B. Main study**	→	**C. Post-module interviews**
1. Data collection	Repertory grid interviews (20) Designs 1 and 2	→	Observations of statistical modules (3) Lectures, tutorials, field notes, memos, conversations with lecturers and students	→	Interviews with lecturers (3) Semi-structured
	↓		↓		↓
2. Data analysis - phase 1	Quantitative (factor analysis); Qualitative (repertory grid content)	→	Grounded analysis (coding system)	→	Grounded analysis (refine coding system)
	↓			↓	
3. Data analysis - phase 2	Content analysis	→	Refine coding system (macro-micro analyses of context using sociocultural theory)		
4. Interpret and integrate the findings			Characterise the teaching of statistics in context		

2 | Literature Review

Chapter contents

In this chapter, I review the literature concerned with the teaching of introductory statistics at university that informed my study. Zieffler et al. (2008) contend that the main goal of statistics education research is to enhance the teaching of statistics, which in turn leads to improved student learning. The Research Advisory Board of the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE) in the USA supports this view of statistics education, adding that research in statistics education needs to formulate implications for teaching in the classroom context as well as the generation of new research questions. In this literature review, I take a broad view of statistics education to include studies that can inform my present research into the teaching of statistics at university.

The studies I identified for this review come from diverse fields of inquiry, including mathematics education, statistics education, psychology and educational psychology. The research questions, methodologies and participants are very different and vary in their usefulness for my research project. As a result, I chose to focus on research relevant to the teaching of statistics in introductory university modules (or courses) for non-statisticians. Since I found that there is a lack of research directly focusing on observations of teaching in this context, I also considered the few studies that have investigated observations of mathematics teaching at university.

My research design (Section 1.4, p. 4) involved two main phases to my literature review. In the initial stages of my study, I consulted the statistics education literature more broadly to inform my Pilot study (Chapter 3), formulate terminology used in statistics education and gain insights into the learning and teaching of statistics at university. In my Main study (Chapter 4), the literature review provided theoretical insights into my study using sociocultural theories of teaching and learning and was part of my grounded analytical approach.

This chapter begins with (1) an overview of my approach to the literature review (Section 2.1), followed by (2) an overview of research in statistics education (Section 2.2), (3) a definition of curricular planning (Section 2.3), followed by (4) research into intended curricula (Section 2.4), (5) the teaching of statistics at university (Section 2.5), and (6) into the impact of teaching methods and approaches on student learning (Section 2.6). Finally, I discuss the implications for teaching statistics and for my research study in Section 2.7.

2.1 Approach to the literature review

My approach to the literature review was synthetic, focusing on providing a critical outline and analysis of current and recent research relevant to the teaching of statistics at university level. I did not seek to carry out a systematic review which quantifies the extent of the literature written over a particular time period, and/or tests particular hypotheses. With this review therefore, I describe the scope and scale of the statistics education research into teaching practices, adopt specific terminology relevant to researching the teaching of statistics at university that I use in this thesis and provide a knowledge base to compare and contrast my research outputs.

To ensure breadth and depth in the statistics education literature, my first step in the literature review was to identify keywords based on my research aims which I constructed into search strings. For example, the following search string

> su(teaching statistics OR lecturing statistics) AND su(higher education OR tertiary OR university) AND su(statistical reasoning OR statistical thinking OR statistical literacy).

on the database ProQuest Social Sciences Premium Collection produced 1,465 results in English between 1970 and 2018, including peer-reviewed articles, dissertations and theses, or reports.

Next, I identified bibliographic databases (APA, JSTOR, EBSCO, BHI, ASSIA, COPAC, ERIC, IBSS, Informa, Ingenda Connect, SciVerse ScienceDirect Elsevier, Scopus Elsevier, SpringerLink, SpringerLink Books, Taylor & Francis Cross Ref, Wiley Online, CAUSE website), specific journals, conferences and organisations (e.g. Journal of Statistics Education, Statistical Education Research Journal SERJ) and authors (e.g. Cobb, Garfield, Ben-Zvi, Bakker).

Three, I reviewed the references of each selected article to identify any potentially relevant sources and four I searched relevant citing literature. Relevant articles were considered to be concerned with

- *Curricular planning at university*, including the nature of statistical skills, undergraduate statistics curricular movements and strategies for specifying module content.
- *Teaching of statistics at university*, including teacher characteristics (knowledge, beliefs, planning), observational studies of teaching and pedagogic strategies (e.g. explanations, motivating students, questioning, teaching discourse).
- *Teaching and student learning*, including teaching methods (e.g. lecturing, active learning, teaching resources, technological tools) and students' difficulties with learning statistics (e.g. misconceptions and non-cognitive skills such as anxiety, attitudes and motivation for statistics).

As the fifth step, I carried out the searches using Primo® ExLibris interface `www.exlibrisgroup.com` to identify and download resources, such as books, journal articles, digital objects, reviews, conferences proceedings and so on using a range of databases, as described above. The search term used posed a challenge since 'statistics' and 'education' are used in a variety of disciplines. This meant that the high volume of results had to be screened further to identify relevant sources. To do this, I used the title, keywords, abstracts and source.

Six, using bibliographic software (Papers 3), I saved the sources and categorised them into one or several broad purposes, summarised in Figure 2.1 on p. 12.

Using the same software, the next step, seven, involved writing up a brief critical synopsis for each source, including educational level (school-based, university), purpose of study, data, analysis techniques, findings, implications, inconsistencies. Further, the sources identified could be further categorised into research, professional, and pedagogical literature (see Treffert-Thomas and Jaworski, 2015).

Finally, the eighth step was to compile a table to map the different elements of the sources and identify key themes in statistics education relevant to my research. The table contained the following columns: study, purpose, education level (school, university), methodology, summary of study, relevance (1-not relevant to 3-include).

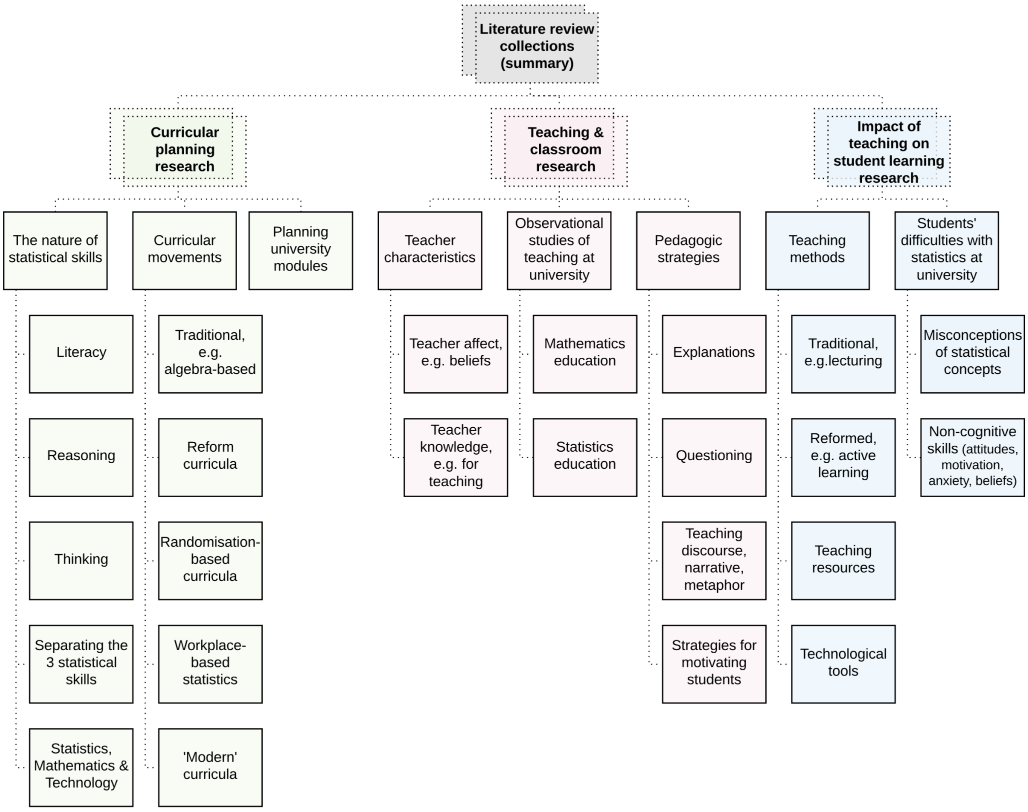


Figure 2.1: Literature review collections

2.2 Overview of statistics education research

Statistics emerged as a subject taught at university level with a focus on probability in the 1930s (Ottaviani, 1989). Since then, statistics has developed into a broad subject, applied to a range of real-world contexts and disciplines, including the social sciences, engineering, the sciences and even the humanities (Nisbett et al., 2009; ACME, 2011; Davies et al., 2012; Zieffler et al., 2008). As a result, the contexts of teaching and learning statistics are also diverse. Researchers in the field of statistics education might also work as educational researchers, statisticians, psychologists, pedagogists, scientists, economists and so on (Zieffler et al., 2008). Despite such diversity of backgrounds and perspectives, the ultimate goal of research in statistics education, as with any other educational research, is on how to teach to improve the educational outcomes of statistics learners at all levels (Ben-Zvi and Garfield, 2008; Petocz et al., 2017).

Research in statistics education has emerged as a field of inquiry in the 1940s and has focused on informal adult reasoning with statistics, formal teaching of statistics and students' difficulties with formal learning of statistics in the classroom (Garfield and Ben-Zvi, 2008). In Table 2.1, I summarise the key research strands chronologically. In the first column, I depict different decades of developments in the field of statistics education. In the second column, I summarise the main research focus or strand with examples of studies I considered to be representative for that particular strand in the third column. In deciding which study to include in the third column, I used the table mapping the sources and key themes discussed in my approach to the literature review in Section 2.1, although the studies included in this column are used as examples I considered to be important and are not intended as an exhaustive list.

Table 2.1: Key developments in statistics and statistics education

Year	**Research focus**	**Examples of studies**
1700s	*Bayes' Theorem*	1763
1800s	*Foundation of the Royal Statistical Society in London*	1834 (UK)
	Foundation of the American Statistical Association	1839 (USA)
1900s	*Chi-square*	Pearson (1900)
	T-test	Gosset, 1908
	Theory of statistical estimation	Fisher, 1925
	Hypothesis testing	Neyman and Pearson, 1933
	Confidence intervals	Neyman, 1937
1930s	Statistics courses offered at university	USA
1940s	Training of statisticians	Garfield and Ben-Zvi (2008)(USA)
	Teaching in universities	Royal Statistical Society, 1947 (UK) in Holmes (2003)

Continued on next page

Table 2.1 – continued from previous page

Year	Research focus	Examples of studies
1950s	*Decision Theory*	Wald, 1950
	Advent of electronic computations of statistical applications	Efron and Hastie (2016)
	Teaching at school level	Royal Statistical Society (1952, UK)
1960s	How to teach secondary-level students how to use and analyse data	Curriculum in Statistics and Probability (USA); statistics at A level (UK); Nuffield primary mathematics (UK); School Mathematics Project (SMP, UK)
1970s	The development of interesting and engaging teaching resources at school level	Project on Statistical Education (POSE), Schools Council and Committee on Statistical Education (UK)
	Students' difficulties with learning statistics at university level	Urquhart (1971); Kalton (1973); Jolliffe (1976)
	The cause and nature of errors in intuitive statistical reasoning	Kahneman and Tversky (1972); Tversky and Kahneman (1975)
	Processes involved in solving statistical problems	Chervany et al. (1977)
	Probabilistic thinking and intuitions	Piaget and Inhelder (1975)
1980s	How to teach statistics at school and university	Kapadia (1980); Kempthorne (1980)
	Student misconceptions and difficulties with statistical reasoning	Garfield and Ahlgren (1988)
1990s	*The advent of fast computation*	
	Reform movement: new guidelines for introductory statistics modules	Cobb (1992, 1993)
	Research into statisticians' practices and comparing expert and novice thinking	Wild and Pfannkuch (1999)
	Defining statistical skills (statistical literacy, reasoning and thinking)	Moore (1990), Gal (2002), Chance (2002)
	Improving statistical reasoning at all levels of education	Konold et al. (1993a); DelMas et al. (1999)
	Assessment at all levels	Garfield (2003); Hirsch and O'Donnell (2001); Schau and Mattern (1997); Groth and Bergner (2005); Mathews and Clark (2007)
	The role of non-cognitive factors of adult learners	Gordon (1995, 1993)
	Pedagogical methods at university	Magel (1996); Quilici and Mayer (2002)
	The use of technology in teaching	Garfield et al. (2012b)
2000s	*Big data, machine learning and data mining, large-scale prediction algorithms*	
	Student learning	Derry et al. (2000)
	Lecturers' conceptions of teaching statistics	Gordon et al. (2007)

Continued on next page

Table 2.1 – continued from previous page

Year	Research focus	Examples of studies
2010s	The nature, development and teaching of statistical skills in school and workplace	Bakker and Akkerman (2013); Dierdorp et al. (2011); Garfield et al. (2012a);
	Student informal reasoning as a precursor to formal inferential reasoning	Ben-Zvi et al. (2012)
	European Conference on Data Analysis (ECDA)	Luxembourg (2013)
	Statistical Learning and Data Mining Section	American Statistical Association (2014)
	Data Science Section	the Royal Statistical Society (2017)

Psychological research in the 1970s and 1980s was dominated by quantitative methodological approaches, while educational research attempted to use a mixed methodology. The focus on adult statistical reasoning using undergraduates as participants in the 1970 had shifted to school children's development of statistical skills in 1980s and 1990s and more recently to university/mature students' learning and using statistics.

An important focus in statistical education research in the 1990s was on understanding and defining the nature of statistics as a discipline. This was considered to be important in formulating learning intentions for the students, planning and implementing curricula, teaching resources and assessments of students' learning outcomes (Garfield and Gal, 1999). As such, considerable effort has been made to define different types of statistical processes and knowledge, i.e. statistical literacy, reasoning and thinking, mathematical knowledge, understanding of the real-world context and critical thinking (Zieffler et al., 2008).

In the 2000s, research has focused on knowledge for teaching (e.g. Makar and Confrey, 2004; Pfannkuch, 2006; Gordon et al., 2007) and on identifying pedagogies and didactics that promote these types of knowledge (e.g. Wild and Pfannkuch, 1999; Pfannkuch and Wild, 2000, 2004; Cobb and Moore, 1997; Allan, 1996; Ben-Zvi and Garfield, 2004; Chance, 2002). This body of research envisioned the study of curriculum and teaching methods as a means to improve teaching practices and consequently students' attainment, values and attitudes. Regarding student learning, statistics education research has investigated how students develop understanding of statistical concepts, has sought to understand why introductory statistics is challenging, how to make it more accessible for students and which curricula leads to enhanced student outcomes, both in terms of exam results and using statistics after the completion of a module (Magel, 1996).

The research strands depicted in Table 2.1 show the diversity of research interests in statistics education, which I regard as informative for my general understanding of teaching practices in statistics education. Since most of the research in statistics education in recent years has been based at school level, I only include in this literature review studies of direct relevance to my research questions and context such as defining the nature of statistical skills, curricular planning research and research into the teaching of statistics, including teacher characteristics, teaching methods and approaches, pedagogic strategies and teaching interventions.

Further, I considered that the contexts of research into teaching statistics in compulsory (school-based) education differed in important ways from research in post-compulsory education. For example, teacher beliefs about statistics curricula, about the nature of statistics, about its relationship with mathematics and about teaching and learning might be different according to the age and stage of the student group and teachers' academic background. Teachers in compulsory education are generally mathematics teachers whereas in post-compulsory education, statistics lecturers may be statisticians or users of statistical methods within other disciplines (e.g. Psychology). Facets such as teachers' academic background, the place of statistics within the curriculum, e.g. as part of mathematics or as a standalone subject taught as part of a student's general academic education, learning outcomes or students' age might impact on the transferability of findings across contexts (Kapadia, 1980; Reading, 2011). The literature reviews relating to my methodological approaches for the pilot and main studies are included in Chapters 3 and 4 respectively.

2.3 Curriculum

Eckstein et al. propose that a curriculum is more than a syllabus or module specification. Instead, the authors suggest that curriculum is 'the collection of organised learning experiences' that a school or university provides (Eckstein et al., 1982, p. 11). In order to understand the *context* in which introductory statistics is taught and learnt, in my thesis, I focus on official curricula and on the factors that influence how curricula is implemented, the teaching of statistics.

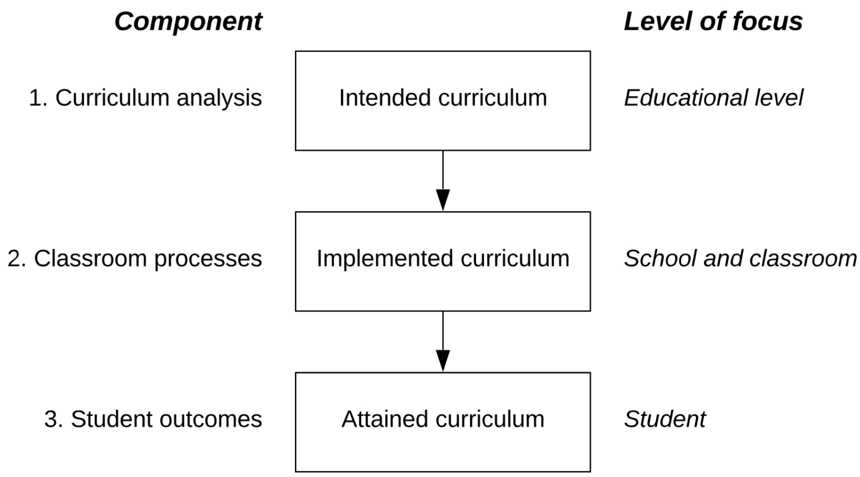


Figure 2.2: The 'tri-partite' view of curriculum (from Eckstein et al., 1982, p. 12)

Eckstein et al. (1982) and Westbury and Travers (1990) define three components of a curriculum. The first component is the **intended curriculum** which includes a set of expectations for learning a subject which are held in official documentation, e.g. benchmarking statements, professional body standards, programme and module specifications (terminology explained in Section 2.4.3). The second component is about **implemented curriculum** which focuses on how much statistics is taught and how curricular guidelines, the intended curriculum, are taught by the teacher in the classroom, including for example teaching styles. The third component is student achievement, the **attained curriculum**, which can include knowledge about the subject and also values and attitudes. Eckstein et al. (1982) conceptualises the three aspects of the curriculum - intended, implemented and attained and relates to each of these components curriculum analysis, classroom processes and student outcomes with levels of focus - educational system, school and classroom and student (Figure 2.2).

In my research, I found this 'tri-partite' perspective of the curriculum useful for conceptualising the different aspects of my research study. However, I considered that the elements in the diagrammatic representation in Figure 2.2 were inter-connected such that students' attained curricula can influence both the intended and the implemented curricula by the teacher. For instance, in her teaching triad, Jaworski (1994, 2002) proposed that the teaching of mathematics involves three related yet distinct domains: management of learning, sensitivity to students and mathematical challenge. Management of learning describes the teacher's role in creating a learning environment such as classroom grouping, curriculum decisions, fostering ways of working and establishing values and expectations. Sensitivity to students involves developing a knowledge of students, their individual characteristics and needs and an approach to working with students, consistent with these needs. Finally, mathematical challenge involves stimulating mathematical thought and enquiry, and motivating students to become engaged in mathematical thinking. As a tool for making sense of the practice of teaching mathematics, Jaworski's teaching triad suggests to me that all three domains influence the design of activities (the intended curricula), the style in which they are presented (the implemented curricula) and ultimately students' mathematical thinking and learning activity (the attained curricula), as drawn in Figure 2.3.

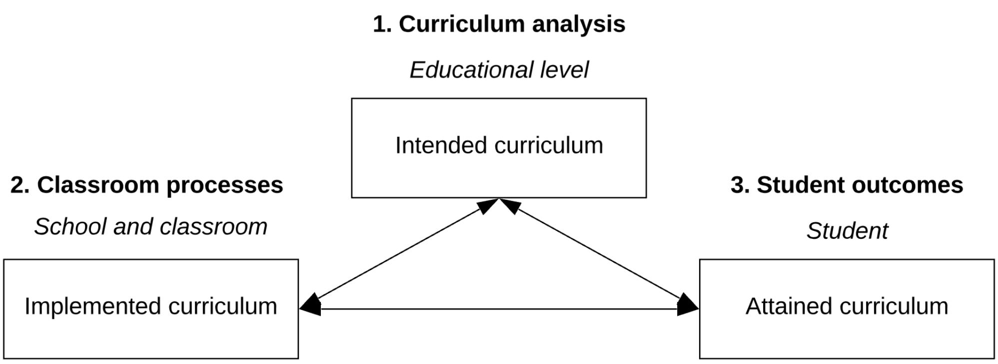


Figure 2.3: A revised view of curriculum

In this chapter, I review studies in statistics education concerned with understanding statistics as a discipline and defining the intended, implemented and attained statistical curricula relevant at university level. I discuss the literature relating to intended curricula in Section 2.4, implemented curricula in Section 2.5 and to impact of teaching methods on students' attainment in Section 2.6. In the initial stages that included the pilot study (Chapter 3), I focused on the intended curriculum. My main study (in Chapters 4 and 5) was concerned with implemented curricula and the links between planned and implemented curricula. Although the student as a focus was not part of my inquiry, I was sensitive to student outcomes in interviews with lecturers in the pilot study and when observing classroom processes in my main study.

2.4 Research into curricular planning

In a technical sense, statistics is a summary of data, involving statistical analyses and methods that apply wherever data, variation, and chance appear since they are 'omnipresent' in modern life (Moore, 1998; Gould, 2010). Further, Abelson promoted the idea that 'the purpose of statistics is to organise a useful argument from quantitative evidence, based on a form of principled rhetoric (Abelson, 1995, p. xiii). More recently, statistical software has increasingly replaced hand computations with the need for more conceptual understanding of statistics (Moore, 1998; Gould, 2010; Malone et al., 2012).

Table 2.1 shows that early statistical education work at university focused on defining intended statistics curricula and designing interesting and engaging teaching resources to specify the content of statistics modules in more detail.

In order to understand the circumstances in which statistics is taught, I examined the literature concerned with the learning intentions, concepts and procedures proposed to be important for students to learn at tertiary level. Learning intentions (goals or objectives) are what the statistics education literature recommends for students to learn as a result of engaging in a particular sequence of learning activities, e.g. over the course of a lecture. Learning intentions are separate from the context of learning, which is defined by specific applications or success criteria. For example, the learning intention 'by the end of their learning students should be able to select appropriate statistical methods to analyse data' is separate from its context (or range), e.g. 'analysing numerical parametric data'.

2.4.1 The nature of statistics

Teaching statistical concepts of statistical literacy, reasoning and thinking and/or statistical methods (computations) are considered key learning goals in introductory statistics across disciplines (Garfield, 1994; Bradstreet, 1996; Cobb and Moore, 1997; Broers, 2006).

2.4.1.1 Statistical literacy

In information societies that rely on computers and data, statistical or quantitative literacy is considered a key 21st century skill for efficient citizenship, similar to the ability to read and write (Watson and Callingham, 2003). Gal (2002) suggests a model of statistical literacy composed of five knowledge bases and a cluster of beliefs, attitudes and a critical stance. The knowledge elements include literacy skills, statistical knowledge, e.g. number sense, understanding variables, interpreting tables and graphs, experimental design, data analysis, sampling and inferential reasoning, mathematical knowledge which is kept to a minimum since computers replace computations, context/world knowledge to understand potential sources of variation and error.

Gal (2004) further proposed that statistical literacy is the ability to interpret, critically evaluate, and communicate about statistical information and messages. In this model, statistics is viewed as *numbers with a context*. Importantly, Gal (2002) considered that statistical literacy is not only a skill or an ability but also as a set of cultural practices that people engage in. World knowledge and literacy skills are considered to be prerequisites to enable 'critical reflection about statistical messages and for understanding the implications of the findings or numbers reported' (p. 15).

Watson and colleagues (Watson, 1997; Watson and Callingham, 2003) examined the levels of thinking required to interpret statistical information presented in society (e.g. as reported in media articles). In their view, these levels of thinking are characterised as basic understanding of statistical terminology, language and concepts when they are embedded in a wider context. Further, higher order analysis skills are represented in a questioning attitude (e.g. critiquing sampling methods), the distribution of raw data, the appropriate use of statistics, and the validity of an argument.

2.4.1.2 Statistical reasoning

A number of studies have studied university students statistical reasoning, in the presence or absence of formal teaching in statistics and using both qualitative and quantitative methods (Zieffler et al., 2008). Statistical reasoning research has been highly influenced by Biggs and Collis Structure of the Observed Learning Outcomes taxonomy (Biggs and Collis, 1982) which specifies that a cycle of learning has five hierarchical levels: prestructural, structural, multistructural, rational and extended abstract. These five levels have been used to characterise students' development of statistical reasoning. For example, Chervany et al. (1977) proposed a model of statistical reasoning that included three stages. The first stage, comprehension of a problem, involves the formulation of the research question as an instance of a first version of the problem. The second stage, planning and execution of a statistical analysis, is about applying the methods to the specific case of the problem. Finally, the third stage, evaluation and interpretation, is concerned with the validity of the inference (outcome) from the initial problem (research question).

Ben-Zvi and Garfield (2004) considered that statistical reasoning is about how people make sense of statistical information, which may involve representing and summarising data, interpreting an analysis, and making connections amongst concepts. Further, statistical reasoning is about understanding statistical processes and also being able to explain and interpret them. Garfield (1998, p. 78) further contends that statistical reasoning may be about ‘the way people reason with statistical ideas and make sense of statistical information’, e.g. interpreting data representations and outputs.

Nisbett et al. (2009) define statistical reasoning as the need for clarity of the sample space and sampling process and add a third factor, cultural prescriptions to reason statistically about events of a given kind: ‘statistical reasoning is the culturally prescribed way to think about randomizing devices in our culture’ (p. 346). This definition of statistical thinking highlights that statistics connects everyday life and mathematics. So, a key difference between mathematics and statistics is the role of “context” (scenario) in problem solving. Understanding and using the context of a problem is also considered key in the statistical investigative process (i.e. formulating the research question, data collection, selecting methods of analysis, testing assumptions).

2.4.1.3 Statistical thinking

Cobb and Moore (1997) considered the core elements of statistical thinking to be theory of experimental design, data exploration (which is helped by statistical software) and interpretation. The emphasis in Cobb and Moore’s (1997) model is on context and the non-mathematical substance of statistical thinking that includes the authentic interplay between data pattern and context.

In a study that aimed to define statistical thinking in empirical enquiry, Wild and Pfannkuch (1999) used empirical data from interviews with statisticians and students doing statistical projects. Wild and Pfannkuch’s model suggests that there are *five* fundamental modes of statistical thinking: 1) recognising of the need for data, 2) accepting that variation exists, 3) translating the real world system into a quantitative model, in order to arrive at a better understanding of the world, 4) building and reasoning with models using logical thinking, common sense and number sense, and 5) integrating and synthesising contextual knowledge with statistical knowledge (e.g. drawing inferences from data, communicating a statistical argument about a situation).

Wild and Pfannkuch (1999) also provide a comprehensive description of the processes involved in the statistical investigative enquiry, from formulating research questions to interpretation and conclusions. The model suggests that statisticians operate (sometimes simultaneously) along four dimensions: the investigative cycle (define a problem, plan to solve the problem, collect data, analyse it, and formulate conclusions), types of thinking using statistics, strategies for problem solving, and dispositions (logical thinking or imagination). Ben-Zvi and Garfield (2004) also recognise the importance of non-statistical skills such as critical thinking to evaluate the results of a statistical study.

Cobb and Moore (1997) suggest that 'data are not just numbers they are number with a context' . The context motivates statistical procedures and is quintessential in interpreting the results of such an analysis. Pimenta (2006) posit that statistical thinking can only materialise as part of a context (problem situation) and that it is not possible to apply statistical methods without considering the context. Fundamentally, statistical thinking is about noticing and seeking to explain the differences in the numbers using context knowledge (Pfannkuch and Wild, 2004).

In a review of research into defining key goals for statistics education, Ben-Zvi and Garfield (2004) contend that statistical thinking involves an understanding of 'why and how statistical investigations are conducted and the "big ideas" that underline statistical investigations' (p .7) and should be included in statistics instruction: data, distribution, trend, variability, models, association, samples and sampling, and inference. These '*big statistical ideas*' relate to Wild and Pfannkuch's (1999) fundamental modes of statistical thinking that are part of their four-dimensional framework for statistical thinking in empirical enquiry.

2.4.1.4 Separating the three statistical skills

In summary, statistical literacy is the ability to interpret, critically evaluate and communicate about statistical information (Gal, 2002). Statistical reasoning is about conceptual understanding, how people make sense of the '*big statistical ideas*' such as data, statistical models, variability, distribution, centre, sampling, uncertainty, significance, experimental design, informal statistical inference as foundation to formal statistical inference. Statistical reasoning is also about connecting these concepts to each other and combining ideas in order to understand and be able to explain statistical processes and interpret the results (Garfield and Ben-Zvi, 2004). Finally, statistical thinking is about procedural knowledge, processes involved in the statisticians' practice of data-based inquiry from problem formulation to conclusions (Wild and Pfannkuch, 1999).

DelMas (2004) distinguished between statistical thinking and statistical reasoning on the basis that different cognitive processes lead to different types of student errors. DelMas suggested that 'a person who knows when and how to apply statistical knowledge and procedures' demonstrates statistical thinking. By contrast, a person who can 'explain why results were produced' or why a conclusion is justified demonstrates statistical reasoning' (p. 85).

However, differences between statistical thinking and reasoning are not always clear or mutually exclusive, which can lead to issues with evaluating and conducting research in this area (e.g. Pimenta, 2006; Gould, 2010). In my understanding, statistical reasoning and thinking involve application of statistical methods, interpretation and critical thinking. By contrast, statistical literacy seems to involve more basic yet important levels of understanding, such as being able to understand symbols, concepts, terminology, statistical information or research results, organise and represent data, e.g. tables or graphical displays (Ben-Zvi and Garfield, 2004). Further, although Nisbett et al.'s (2009) focus on statistical reasoning, the difference between statistical thinking and reasoning is not considered. For instance, the subjects' failure to complete a task is explained by a 'lack of concrete experience in thinking about samples', seen as 'an appreciation of statistical principles'. Inductive reasoning appears to be about generalization from instances as *selective statistical intuitions*.

The conceptualisations of statistical thinking, reasoning and literacy proposed in the literature emphasise the key role played by both statistical and non-statistical factors and components. Statistical processes reflect the broad and multi-faceted nature of the contexts or situations in which statistics is used or applied. The three processes, statistical thinking, reasoning and literacy, are particularly indistinct in some of the statistics education literature. For instance, statistical literacy and quantitative literacy. Similarly, statistical thinking and reasoning are used in some of the literature to also define quantitative skills (Ben-Zvi and Garfield, 2004; Dwyer et al., 2003). In general, these processes represent statistical concepts rather than numerical manipulation and graphing of data.

As the focus of my study is on identifying practices and processes in the context of teaching of introductory inferential statistics at university, next I provide a brief overview of some of the research aiming to define the nature of statistical skills. Statistical thinking, reasoning and literacy models are influenced by Piagetian and neo-piagetian domain-specific cognitive perspectives, which historically have evolved from models of development in probability (Jones et al., 2004). Subsequently, statistical reasoning models have been based on structured interviews with students or statisticians, clinical studies, and teaching experiments. Studies using teaching experiments since they rely on students' individual and collective development in statistical reasoning during teaching are particularly relevant to teachers and researchers looking to design and plan teaching. Cognitive models also have important implications for curriculum design, teaching, and assessment. For example, Zieffler et al. (2008) claim that such models of statistical knowledge and skills are better able to explain the nature of conceptual understanding of learning goals such as statistical reasoning and thinking and also students' development of such concepts.

2.4.1.5 Mathematics, technology and statistics

The focus on content of university statistics in the 1960s changed to a focus on defining statistical skills and processes in the 1990s. These shifts in how learning goals are defined are motivated by shifts in how statistics has evolved from probability theory in the 17th century to a mathematical tool for analysing experimental and observational data in the 20th century (Porter, 1986). However, although mathematical concepts and procedures are used in applications of computations and procedures, Moore (1998) considered statistics to be an independent discipline with its own core ideas rather than, for example, a branch of mathematics. Moore further considered statistics to be 'fundamentally different from mathematics, and can be taught appropriately only by someone who recognizes and understands that difference' (Moore, 1988).

Technological tools (such as computers, graphing calculators, software and internet) have changed the way data and data analysis are conceptualised in statistics and can help students develop statistical literacy and reasoning (Garfield and Ben-Zvi, 2007; Petocz et al., 2017). Statistical software has been available for doing statistical analysis since the 1950s, although its role in teaching and learning the subject is still evolving (Garfield et al., 2002; Ragasa, 2008). Despite this, the role of technology in teaching and learning statistics has had an impact on the nature of statistics being taught, curricula and the knowledge required for teaching (King et al., 2002). For example, the assumption is that an increased use of statistical software applications for doing statistics such as spreadsheets, IBM SPSS or R or for teaching statistics (such as TinkerPlots or web applets) has made the subject more accessible for students who can focus on learning statistical concepts and techniques rather than mathematical computations (Mills, 2002).

However, Lane and Peres (2006) found that the effectiveness of computer simulations (software developed for learning statistics) depends on how they are used in the classroom. For instance, simulations can lead to passive learning when teachers use them to demonstrate characteristics of a particular concept. Instead, Lane and Peres recommend "guided discovery learning" in which teachers provide enough structure for students through carefully crafted questions to support students in their discovery of statistical principles. Research into the impact of technology tools on learning points to the importance of the context in which they are used. Further, Garfield and Ben-Zvi (2007) and Pfannkuch (2017) argue that technological tools should be used to help students visualise and explore data for conceptual development, immerse in data-rich environments, probability modelling and statistical reasoning rather than just to follow algorithms to particular ends. The inclusion of EDA and CDA (Cobb, 2005) accompanied by increased access to technological tools have provided opportunities for students to develop conceptual approaches rather than procedural (Langrall et al., 2017).

Some studies have argued that the gap between software for doing statistics and for teaching statistics can hinder the progress of students of statistics, especially at the introductory level or it can lead to a content depth versus breadth, i.e. a reduction in the statistical content that can be covered (Nicholl, 2001; McNamara, 2015). For this reason, Hogg (1991) calls to continually review course content and delivery to meet the demands of society. More recently, McNamara (2015) argues that a future ideal tool for learning statistics needs to allow learners to transition from "a supportive tool for learning to an expressive tool for doing" since "it would allow students to be true 'creators' of statistical products".

Although progress has been made in defining statistical thinking and reasoning skills and although there was agreement that these statistical skills are key outcomes for statistics education, the models do not specify how such skills should be elicited or how teachers should plan for developing these skills in the classroom context using technology. The literature suggests that in order for students to develop statistical literacy, new curricula, teaching methods and teacher training were required that would promote statistical literacy, reasoning or thinking. This curricula would be based on strong connections among content, pedagogy and technology (Garfield and Ahlgren, 1988; Garfield, 1995; Moore, 1997; Gal, 2002). In the next section, I discuss a range of approaches to curricular planning at university proposed over the decades.

2.4.2 Statistics curricular movements

Chervany et al. (1977) elaborated on their model of statistical reasoning based on problem-solving processes in order to evaluate the teaching of introductory statistics. The three main processes of statistical reasoning are (1) comprehension of the problem's context and concepts as an instance of a *prototype*, (2) planning and execution of a solution to the problem and (3) evaluation and interpretation broken down into verifying the validity of the solution to the problem against the initial problem. The emphasis in Chervany et al.'s (1977) framework therefore was both on statistical content and problem solving skills for the introductory statistics course. Chervany et al. were interested in designing evaluation tools for evaluating large-class statistics courses taught as a one semester course and without a calculus prerequisite. The authors proposed that since statistical reasoning is not directly observable, frameworks of statistical reasoning need to specify the reasoning processes to guide the development of learning outcomes and guide the planning and evaluation of content for teaching and for assessment.

Bessant (1992) also suggested that problem-solving skills should be at the core of a statistics curriculum, instead of a focus on statistical procedures. Bessat further considered that curricula should emphasise how and when to use particular procedures given a new context, e.g. a problem or a new data set.

Using findings from the psychology literature, Beyth-Marom and Dekel (1983) propose a curriculum aimed at improving students' inductive reasoning, i.e. thinking under uncertainty that emphasised 'existing intuitions'. The emphasis in this curriculum is on the process of solving a problem rather than on the solution to the problem, thus making students' implicit thought processes explicit by getting them to talk about their own beliefs. In this way, Beyth-Marom and Dekel proposed a curriculum that does not emphasise probability. Instead, they emphasise students' existing intuitions through introspection (e.g. by asking questions such as *how do people think?*, *why did I reach that conclusion?*).

This early curricular work, although fundamental to the changes that took place over the following four decades in statistics education, did not seek to evaluate whether a different approach to teaching estimation, sampling and other concepts related to uncertainty improved students' thinking skills.

2.4.2.1 Reform curriculum

The reform movement at the introductory statistics university level in the 1990s recommended that the statistics curriculum should change to respond to students' needs. Several authors began to suggest new and reformed curricula that can lead to students developing statistical thinking, reasoning and literacy. It is perhaps relevant to note that this new curricula followed nearly three decades after statistics emerged as a new subject taught at university level focusing on probability, two decades after Tukey's (1969) reinterpretation of statistics as exploratory data analysis (EDA) and confirmatory data analysis (CDA, inferential statistics) and a decade after cognitive psychology started to test theories of adult statistical cognition.

The content of *reformed curricula* intended to cover exploratory data analysis (using software), theory of experimental design and key ideas in statistical inference (understanding the logic of statistical inference, significance tests and hypothesis testing) but not mathematics or formal probability theory (Cobb, 2015; Cobb and Moore, 1997; Moore, 1998, 2001; Gould, 2010). Bradstreet (1996) proposed that on statistical modules, statistical reasoning should precede statistical methods. Further, Moore (1998) viewed statistical reasoning as a distinct intellectual skill and recommended that students should have the opportunity to experience the process of data collection and data exploration. Watson (1997) also emphasised the need for teachers to be able to teach students statistical skills such as interpreting statistical information in context and questioning arguments in order to achieve the proposed targets of achievement such as being able to have a basic understanding of terminology, embed language and concepts in the wider context, and importantly question the claims. Watson considered that skills such as computation are not required, since understanding media articles involves the understanding of concepts, the communication of ideas as well as the recognition of meaning in context rather than computation.

At this time, statistical thinking, a focus on more data and concepts, less theory and 'recipes' and active learning became prominent ideas in the teaching of undergraduate statistics. Cobb and Moore (1997) proposed that traditional ways of teaching statistics with an emphasis on hand computations, formulas and procedures should be replaced with a focus on statistical ideas, data, and statistical concepts. In order to 'teach statistics as statistics', introductory statistics curriculum needs to abandon theoretical presentations of basic statistics and instead encourage students to reason from uncertain empirical data. This view of the statistics curriculum is justified by the fact that statistics, as opposed to mathematics, is linked to a large body of professional practice, which has been highly influenced by the advent of available statistical software. In their view, 'theory makes much more sense when its setting in practice is clear' (p. 821). Similarly, Gal and Garfield (1997) compared curricular goals across educational levels and suggested that suitable cognitive models are those that focus on specific processes (e.g. inference) as they occur when students apply statistical processes, such as collecting and analysing data.

Rossman and Chance (1999) offered a top ten list of recommendations for teaching the reasoning about statistical inference. The authors' main points included issues with investigating and discovering inferential reasoning, emphasising appropriate interpretation of inference results, warning students about overuse and misuse of inference procedures (i.e. understand Box's statement that "all models are wrong, but some are useful," Box and Draper, 1986), building connections and communicating results. Based on the statistics education literature, the authors further contend that students should discover basic ideas of inference (e.g. variability) by performing calculations using own empirical data and software.

Another principle was for students to learn confidence intervals without the use of formulas and instead use graphical representations.

In an opinion article, Hoerl (1997) echoed the view in the statistics education profession that the introductory course should focus on students' future employment needs and on statistical thinking rather than numerical calculations. Moore (1998) contended that the statistics curriculum needs to equip students to think about data, variation, and chance and prioritise discussing broad ideas over technical content (e.g. calculations).

These authors emphasise that statistics should involve students learning from experience, i.e. inductive rather than deductive inferencing. The inferences expected of students (e.g. predictions about a population from a sample) might not be characterised as 'right' or 'wrong'. Instead statistical inferences might need to be evaluated in terms of 'quality of reasoning, adequacy of methods employed, and nature of data and evidence used' (Garfield and Gal, 1999, p. 3).

Research and professional literature in statistics education in the 1980s to late 1990s investigated the nature of statistics and made recommendations regarding the content areas and curricular intentions for the students. The picture presented in these studies was of a context-free, universal content which aimed to justify the content of statistics as 'basic skills' that needed to be learnt by any student taking an introductory statistics module. This lead to the proposal of 'classic' statistical modules that took for granted the components of statistics, from sample, variance and mean to statistical models. Understandably, curricular research aimed to move from a *decontextualised* view of schooled knowledge and discourse (Derry, 2013) and present forms of curricula that made references to the context in which the learning took place or of the processes involved in actualising statistical concepts. For example, in order to overcome perceived challenges with student learning and motivation (Section 2.6), some proposed solutions for improving statistical modules. Such solutions for module improvement were based on professional judgement, on experience, on learning theory and on teaching interventions (Moore, 1988; Cobb, 1993; Snee, 1993; Gal and Ginsburg, 1994; Garfield, 1995; Gordon, 1995; Konold, 1995; Sowey, 1995; Aberson et al., 2003).

Despite a growing number of papers calling for statistics to be taught as 'a decision-making science with an empirical orientation', using electronic computations and real-world problems that emphasised active learning (Swanson and McKibben, 1998), undergraduate modules continued to employ the so-called 'algebra-based' introductory statistics module which included topics in descriptive statistics and experimental design, probability and sampling distributions and statistical inference (Moore et al., 2003).

2.4.2.2 Randomisation-based curriculum

In 2005, George Cobb challenged the so-called *consensus curricula* of the 1990s that put the normal distribution and the central limit theorem at the center of all inference. Cobb (2007) stated that 'we may be living in the early twenty-first century, but our curriculum is still preparing students for applied work typical of the first half of the twentieth century' (p. 7).

Cobb proposed a new randomisation-based introductory statistics curricula designed around the core logic of inference of randomisation approaches: permutation tests, bootstrapping and simulation (Cobb, 2007). Repeatedly throughout the course, the curricular goals were to focus on the statistical investigative process, using technology simulations to help students develop an understanding of the concepts of statistical significance and p-values.

Cobb (2007) recommended the teaching of statistical inference centred around the logic of inference rather than the normal distribution, referred to as the 3Rs: (1) randomize data production to provide a basis for inference, (2) repeat by simulation to see what outcomes are typical (and which ones are not), (3) reject any model that puts the data in its tail. Grounded in the history of statistics, Cobb's (2007) proposal was that the more conventional approach based on calculations from normal-based probability distributions was less effective than when the logic of statistical inference was placed at the centre of the teaching. Cobb's curricula included exploratory data analysis and also randomized data production, both sampling schemes and experimental design, inference by way of the permutation test for randomized experiments.

In a research-based undergraduate statistics module inpired by Cobb (2005, 2007), Park et al. (2011) designed a sequence of activities to enable students to develop an understanding of chance models and simulation, models for comparing groups and authentic models of statistical thinking. This *new* curriculum used open-ended, authentic, real-world problems and was based on the work by Schoenfeld (1992) on developing "expert" ways of thinking statistically" and (Himmelberger and Schwartz, 2007) creating prior knowledge to build new learning. The main aims were to promote statistical thinking through modelling and simulation methods and students' positive attitudes about statistics rather than procedural approaches.

2.4.2.3 Workplace-based statistics

Other studies however have identified a gap between the practice of statistics and the teaching of introductory statistics. As statistics and software were becoming more sophisticated in the 2000s, the focus shifted towards methods required in statistical practice such as multivariate statistics which traditionally did not form the focus of introductory statistics. In an article pertinently titled *What your doctor should know about statistics (but perhaps doesn't. . .)*, Switzer and Horton (2007) reviewed the statistical methods used in medical research published in *The New England Journal of Medicine*. The analysis showed a continuous trend towards using more sophisticated statistical methods that can describe complex phenomena, well beyond the content of introductory statistics, thus posing issues for readers (consumers) of this literature. While the reform movement in statistics education placed an emphasis on statistical concepts, Switzer and Horton (2007) considered that statistical modules need to include the statistical concepts students encounter in their future professional practices (e.g. reading and conducting research).

Wild and Pfannkuch (1999, Section 2.4.1.3) base their conceptualisation of statistical thinking on an empirically-based description of the processes involved in statistical practices of both expert statisticians and novice students working in statistics projects (in Section 2.4.1.3). In later studies, Bakker and colleagues began to recognise the importance of involving employers as sources of knowledge for defining the nature of statistics and defining curricula (Bakker and Akkerman, 2013).

For example, Bakker et al. (2008) investigate a group of vocational students carrying out workplace-based statistics projects to characterise statistical inference in the workplace rather than college. The authors observe that an important difference between school and workplace knowledge bases is that “at school, contexts are often used to learn about statistics, whereas in the workplace, statistics is more likely to be used to learn about the context” (p. 132). Further, Bakker et al. (2008) note that “workplace makes sophisticated use of elementary statistics whereas in the classroom we encounter elementary use of sophisticated statistics”. In the workplace, statistical inferences require statistical and contextual grounds, considering real-world constraints, actions and responsibility. Based on this comparison between formal teaching of hypothesis testing and workplace use of statistical inference, Bakker et al. (2008) recommended that the formal teaching of statistical inference needs to include how it was used in practice and the understanding of statistical processes.

Similarly, in Malone et al.’s (2012) view, teaching should be more closely aligned to what a scientist/statistician does. The focus in his view should be on the “context” (situations) of the examples used in teaching captured in a sequence of cases through the module. Malone et al. (2012) consider that probability is not a necessary curricular outcome on introductory statistical modules and proposed an alternative sequence in which the topics were presented. Each case should cover a consideration of the research questions, experimental design, data collection, descriptive statistics, assumptions and inferential processes. In this way, statistical inference is introduced earlier in the module while students go through the process of statistical analysis several times and in multiple contexts. For example, Pfannkuch and Wild (2004) considered that students should be *acculturated* into ‘how statisticians reason and work within the statistics discipline and developing new ways for them to view the world’ (p. 43).

These studies highlight that there are opposing or divergent views of what a statistician’s practice might involve and suggest that the relevance of curricula might be subject-specific, e.g. curricula designed for medical or vocational modules. It is also likely that differences amongst applications of statistics in the workplace present a challenge when designing curricula for a heterogeneous student group since an understanding of the problem context is important. For example, in a recent book pertinently titled *Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things*, Marr (2017) considered that *business skills* are most important for data science, ahead of analytical skills, computer science, statistics, mathematics or creativity. In essence, teaching “big data” needs to involve a “variety” of data, which involves missing data, outlying observations, text data or aggregated information (Khachatryan and Karst, 2017).

2.4.2.4 Modern curriculum

Gould (2010) recommends a curriculum for the 'modern student' who is both a producer of data and of statistics and also a consumer of statistical summaries, e.g. when reading about statistics in the media. Broadly based on definitions of statistical literacy and thinking, Gould (2010) proposes that the statistics curricula should focus on an inclusive view of data that include a context but also create a context when analysed. Further, since statistical activity continues beyond the module or university education, Gould considers that the focus of teaching should be on creating citizens statisticians who can access data, perform analyses and see data summaries in their proper substantive context. To achieve these aims, the challenges are to emphasise technological literacy, i.e. knowing how to use technology and how to create it. Finally, concepts such as central tendency and variation leading for example to the t-test, fundamental in introductory statistics, need to be adapted to new contexts and types of data, e.g. social networks, music, video (Cobb and Moore, 1997; Gould, 2010).

GAISE (2005) and GAISE (2016) also promote the learning of statistics as a scientific enquiry consisting of problem-solving and decision-making processes, rather than a collection of unrelated formulae and methods. In this view, the goal and desired result of all introductory statistics courses is *statistically educated* students who have the ability to think statistically, including understanding the variables have distributions. The recommendations for teaching include the teaching of statistics using “contexts” from students' “own fields of study and everyday lives” in order for them to “appreciate the value of statistical thinking and methods” as an investigative process using active learning strategies. Based on advances in school-based statistics education, Makar (2018) proposed three principles for curriculum design, such as a focus on holistic learning through informal statistical inference, tasks which use an inquiry-based approach and a *layered curriculum.* In a layered curriculum, statistical ideas are introduced first at an informal level, then students develop statistical reasoning and build on concepts multiple times over the years until concepts are formalised. However, such an approach might need to be refined in the context of introductory statistics which expects students to develop statistical thinking over much shorter periods of time.

Important developments in statistics as a scientific field of inquiry took place towards the 2000s. As 20th century statistics was dominated by the Bayesian-Fisherian inferential methodologies (Table 2.1), in the 21st century 'big-data' era, large-scale prediction algorithms, a simple use of regression theory, become prominent, aided by their applications in prediction and also the advent of technology which enable the analysis of huge amounts of data. Cross-validation, permutations and bootstrap are required instead of probability models (Efron and Hastie, 2016).

In the next decade, it is reasonable to believe that 'traditional statistical methods' will continue alongside and probably connect to 'modern prediction algorithms' (Baumer, 2015). Developments in statistics have implications for students' learning of statistics, with a focus on applications within specialist fields of enquiry. The advent of technology (Section 2.4.1.5) in the 20th century has made it possible to carry out complex statistical analyses. In the 21st century, employers are beginning to take advantage on the vast increase in data and the new big data technologies becoming available for analysis to tackle all sorts of problems. It is likely that key statistical skills discussed in Section 2.4.1 as well as mathematics, business skills, statistical programming and dispositions or personal attributes will continue to be the building blocks of statistics and data science education. Ultimately however, the content of statistical modules needs to be related to the students' degree subject, context, interest, prior learning of mathematics and statistics within an inter-disciplinary approach.

2.4.3 Strategies for specifying module content

Few studies during 1990s and early 2000s directly explored implemented curricula within the classroom context or with different teaching approaches. Bradstreet (1996) and Garfield (1998) showed that in practice, statistical skills were not always emphasised in statistics modules or textbooks, which instead, in their view, focused primarily on teaching concepts and procedures. In this section, I discuss strategies for specifying module content, including benchmarking statements, learning outcomes, sequencing of content in statistics modules and challenges with planning or defining such curricula in higher education. So my focus here is on what is done in practice to design statistical modules.

In the UK higher education, the intended curriculum is contained in a set of expectations for learning a subject such as statistics. Such expectations are held at the programme level, but also by professional bodies and regulatory bodies such as the Quality Assurance Agency for Higher Education in the UK (QAA). *Learning outcomes*, what the students should know and be able to do by the end of learning, are specified for every higher education programme and module in the UK (Otter, 1994). Institutions are required to align module and programme learning outcomes with relevant qualification descriptors or benchmarking statements in the national framework for higher education qualifications (QAA, 2013). Standards are therefore specified on a number of levels, from generic in terms of the benchmarking statements, professional body requirements, programme and module learning outcomes, assessment criteria, level and grade descriptors, to very specific in terms of the assessment tasks and associated marking schemes (Krathwohl and Payne, 1971).

For example, in the case of statistical modules, benchmarking statements for psychology require any graduate with such an honours degree to 'reason statistically and demonstrate competence in a range of statistical methods' (QAA, 2010). In the case of economics programmes, in addition to understanding the relevant mathematical and statistical techniques, students are required to have had exposure to the use of such techniques on actual economic, financial or social data, using suitable statistical or econometric software (QAA, 2007b). These examples are broad since they need to be suitable in a wide range of contexts across diverse institutional contexts but may also be evidence of the difficulty in setting standards in a way that is useful and meaningful to users at national level (O'Donovan et al., 2004). Price and Rust (1999) further argue that attempts to establish common minimum standards can fail if they are too generic and open to subjective interpretation. Such standards then become aspirational rather than threshold academic standards.

In some cases, professional body requirements for accreditation are expected to provide narrower and more specific objectives. For example, the UK Standards for Professional Engineering Competence (The Engineering Council, 2013) require registered professionals to 'use a sound evidence-based approach to problem-solving' by conducting statistical data analysis, which is also broad. The British Psychological Society provides more detail by requiring accredited graduates to be able to 'analyse data as specified by research protocols', 'interpret the results of data analysis', 'evaluate research findings and make recommendations based on research findings' and 'write up and report research findings' (The British Psychological Society, 2012, p. 30).

An institution develops programme and module specifications that include specific learning outcomes at programme and module level[1] linked to benchmark statements, professional body requirements or institutional objectives. For example, an introductory statistical methods module for psychologists interpreted the above professional body standards in the module learning outcomes as 'compute inferential statistics for data using SPSS', 'apply inferential statistics to a range of data' and 'communicate effectively with social scientists about statistical issues in hypothesis testing'.

The higher education framework (QAA, 2011) emphasises the importance of specifying the learning outcomes in each lesson and in the assessment tasks used in the module (Gibbs and Dunbar-Goddet, 2007; Elton and Johnston, 2002). Allan (1996) identifies three types of module learning outcomes in higher education: subject-based outcomes (e.g. related to the academic content of the specific module), personal transferable outcomes (e.g. communicate effectively, numeracy, work in teams) and generic outcomes (e.g. critical thinking, analysing and synthesising ideas). Other types may include knowledge, comprehension, the ability to apply knowledge in different situations, cognitive strategies and processing skills (Otter, 1994).

To be useful, the module learning outcomes, content and level descriptors that are translated into assessment criteria, grade descriptors and mark schemes are supposed to be transparent and understandable to staff and students since it is what is formally assessed and credited to students (Allan, 1996). Learning outcomes considered to be 'useful' include what the student will be required to do that demonstrates learning, the context within which the student will demonstrate learning and how well (Otter, 1994). For example, 'apply knowledge of basic statistics when analysing data' might be considered too broad since the verb 'apply' does not denote a specific action or may have different meanings for different people (Raymond and Neustel, 2006). Instead, 'calculate mean, mode and standard deviations using a statistical software such as SPSS when evaluating data' or 'read and critically evaluate psychological research reports containing statistics (e.g. bi-variate methods, graphical representations)' might be considered to be more specific.

[1]In the UK, there are nine qualification levels, from Entry level qualifications to Level 8 at doctorate level. An introductory module is benchmarked at Level 4 (Ofqual, 2012).

In a report on the use of learning outcomes in higher education in the UK, the audit teams were unable to 'find an explicit linkage between learning outcomes and assessment' or 'the links between subject benchmarks statements and learning outcomes at both programme and module level' (QAA, 2007a, p. 1). Although favoured by regulatory bodies such as the QAA, it is a challenge for universities to interpret generic standards such as QAA's benchmarking statements and to formulate module and assessment objectives with sufficient specificity to make them useful for teaching, learning and assessment (Krathwohl and Payne, 1971; Elton and Johnston, 2002; Price, 2005; Hussey and Smith, 2003). Elton and Johnston (2002, p. 14) summarised the challenge faced by departments 'to formulate such objectives so that they are informative and yet not constraining is often an art and never a science' and even that formulating precise, clear and objective learning objectives is 'fatuous or impossible' (Hussey and Smith, 2002, p. 230).

Learning outcomes have been criticised since, given the challenges in setting specific, transparent and measurable learning outcomes, lecturers might find them to be irrelevant to classroom activities and practices or students might find it difficult to understand or achieve them (Hussey and Smith, 2002; Maher, 2004). Lecturers may design their modules not by considering the learning outcomes first, but the content of the syllabus, the time allocated to the different teaching and assessment activities, the year of study (i.e. level), the textbooks used and modes of assessment. There seems to be anecdotal rather than empirical evidence that supports this claim (Hussey and Smith, 2002; Maher, 2004).

Tyler (1949) further proposed that learning outcomes must be different from content, topics or concepts covered in the module because they cannot specify what the student is expected to do by the end of their learning with such content. Good quality learning outcomes, although a challenge to set, seem critical for defining the statistical practice domain students should learn (together with content and theories of statistical reasoning and thinking) and in determining whether particular teaching and assessment methods are appropriate (Garfield et al., 2011). At the same time, there is the notion that expressing very precise learning outcomes in module specifications before the learning takes place can *entrap* both the student and the lecturer into particular teaching experiences which might not be relevant to the students on the module (Hussey and Smith, 2003).

Whilst accepting the notion of constructive alignment, in which learning outcomes, teaching, learning and assessment activities are meant to be connected and related (Biggs and Collis, 1982; Biggs and Tang, 2011; Biggs, 2003), I believe that the reliance in some cases on a set of generic descriptors is too simplistic. For example, using Bloom's taxonomy for framing learning outcomes (Bloom, 1956; Krathwohl and Payne, 1971), especially using particular action verbs in the learning outcomes at particular qualification levels is a rigid interpretation of Bloom's original work and not representative of the range of tasks and activities that should be achieved at each level (Hussey and Smith, 2002). It may well be that the module objectives can be identified in the learning outcomes, but the actual learning outcomes are what the lecturer teaches and the student achieves in the module (Hussey and Smith, 2003).

At the school level, Newton et al. (2011) for example have looked at the extent to which official curricula or standards have a requirement for statistical thinking and reasoning. The authors analysed 1,711 mathematics standards from forty-one states in the USA. This analysis aimed to capture the representation of statistical reasoning in the state standards (curricula) when compared to statistical procedures. The authors categorised the learning standards into the process components of the statistical investigative process, i.e. formulate research questions, collect data, analyse data, and interpret results (Franklin et al., 2007). Further, the analysis looked into whether there was an expectation of statistical reasoning, conceptualised as evaluating or reflecting on statistical procedures, interpreting data sets, connecting concepts, critical thinking, problem solving and decision making (Garfield and Ben-Zvi, 2004). Although the study focused on school-level standards, I could adapt Newton et al.'s (2011) approach in my study in Chapter 3 to map curricula in the context of my study in higher education.

The study found that 28% of standards encouraged statistical reasoning, which was considered low. The study made an important point that, although there is agreement in the statistics education community regarding key curricular aims, in practice it seems less clear how in different contexts such curricula should be formulated, taught or assessed (Zieffler et al., 2008). Such an approach however can only describe the intended curricula, the nature and number of curricular objectives, and not whether the implemented or attained curricula also continue to emphasise statistical procedures rather than statistical reasoning.

Research has also looked into the structure and sequencing of content in statistics modules (Friedrich et al., 2000; Barron and Apple, 2014). For example, Friedrich et al. (2000) were concerned with whether statistics and research methods should be taught simultaneously or one after the other. The authors reported that introductory modules devote an hour or less of class time to each of confidence interval estimation, power analysis, effect size estimation, graphical analysis of data, and statistical reporting yet no more than a class hour to discussion of a general linear model approach and the equivalence of ANOVA and regression. Only a handful of modules in Friedrich et al.'s (2000) survey devoted time to statistical power, effect size estimation, general linear model approaches, and assessment of assumption violations. Statistics modules outside psychology in this data spent more time covering confidence intervals. Overall, the statistical tests covered in introductory statistics emphasised correlation, between-subjects t tests and one-way ANOVA (not part of the general linear model approach) and focused less on regression and contingency tables. When compared to other programmes, psychology modules emphasised ANOVA, including factorial and repeated measures designs, and data analytic procedures rather than probability or mathematics. However, this research documented particular curricular approaches but did not compare these approaches to students' learning outcomes.

Stoloff et al. (2012) conducted a survey of curricular structures of undergraduate psychology departments to compare psychology programmes in 1995 to 2005. The aim of the study was to evaluate the curriculum design recommendations for psychology courses, including statistics, in terms of students outcomes as measured using standardised psychology tests (the Area Concentration Achievement Test [ACAT] and Major Fields Test). The study found no significant differences between methodological approaches to curriculum design.

Barron and Apple (2014) however, based on a study comparing student achievement in different curriculum sequences and structures, found that students in the integrated sequence (statistics with research methods) achieved higher grades than students on non-integrated sequences (statistics before research methods). The students in the integrated sequence also had better long-term retention of methodology concepts based on students' ACAT scores.

Garfield and Ben-Zvi (2008) in a recent book aimed to fill the gap between defining key learning goals in statistics and planning teaching that develops these skills. The book discusses a large body of literature focusing primarily on student understanding of statistics across all levels of education. The authors use this research to make practical recommendation for teachers of introductory statistics in the form of lesson plans and resources. The authors claim to have used and evaluated these resources in their own teaching (Ben-Zvi at school level and Garfield at university), although there are no published observational studies emerging from this work. The book however highlights the challenge with providing *hard evidence* that the materials and approaches are *effective*. In their view, the "context" of the educational interactions such as a teacher's implementation of the materials, teacher's beliefs about teaching and learning statistics, the student cohort and culture in the classroom can influence the effectiveness of curricular planning in specific situations.

So far, I have presented the nature of curriculum and statistics as a subject and some of the challenges in defining curricula that can support 'good' teaching practices. The professional and research into curricular planning shows that statistics curricula are delivered in many ways and that, in general, there is consensus regarding the content of teaching and the use of particular teaching approaches such as active learning that focuses on conceptual understanding and key statistical skills (Section 2.4.1) using real or simulated data in authentic problems rather than statistical procedures. In the next section, I discuss the professional and research literature that is concerned with the teaching of statistics at university.

2.5 Research into the teaching of statistics at university

Historically, teaching processes have been divided between "student-centred" and "teacher-centred" teaching processes, "authoritarian" versus "democratic" or "project-based" versus "lecture" methods. The teaching process believed to be most commonly used in schools until recently is the teacher-centred approach which involves the teacher directing and structuring most student activities (Cuban, 1984). Characteristics of teacher-centred approaches include teacher talking more than the students during class, the teacher frequently instructing the whole class, with less group-work or individual attention. The teacher determines how the class time is used and the classroom layout is generally in rows with students facing the board. It might not mean however that if the classroom layout is different, e.g. students sat at round tables, the teaching style is necessarily student-centred. However, there might be the opportunity for other aspects of a student-centred approach to happen.

Recently, Barak Rosenshine (Rosenshine, 1995, 2010, 2012) proposed seventeen *Principles of Instruction* which I think capture the essence of a modern teacher-centred approach. The teacher is expected to start the lesson with a short review of previous learning (a *plenary*) and learning outcomes, present new material in small steps with student practise at each step, give clear and detailed instructions, ask a large number of questions and check for understanding. The expectation is to use fewer concepts, but with many examples and teacher explanations that can lead to independent learning. This model of *direct* teaching involves a range of beliefs about what good teaching is, including seeing the teacher as guiding the students towards independent practice. However, it describes the teaching as 'explicit'.

At the turn of the 20th century however, educationalists such as Dewey started to challenge the dominance of traditional classroom teaching and proposed a more progressive view of education. Ahead of Piaget (e.g. Inhelder and Piaget, 1958) and Vygotsky's theories of learning (Section 4.2.1), Dewey recognised the importance of the social context in teaching and learning (Dewey, 2011, /1916). Dewey put forward that children should not be *passive* recipients of knowledge. Instead, children learn best through *experiencing* practice, *actively* taking part in the process of learning and interacting with the curriculum. In *Democracy and Education*, Dewey (2011) believes that the teaching and learning are pedagogical. The goal of the educational process is 'one of continual reorganizing, reconstructing, transforming'. Development itself represents a transformation, reorganisation and reconstruction of experience or as Vygotsky conceptualised teaching as transition from preparatory stages to mastery.

In Dewey's view, the curriculum alone does not lead to learning and it should not be imposed on the students. Rather, the curriculum needs to be planned by an adult in effective ways in order to connect the subject matter to the student and taking into account students' prior knowledge, cognitive development and beliefs. Student learning in his view is as important as the teacher and the curriculum. Vygotsky later developed a similar view about the role of teachers as guides or facilitators of learning, enabling and engaging students to acquire knowledge.

Dewey's perspectives on teaching and learning supported *experiential* learning, a student-centred approach which allows the students hands-on, collaborative learning experiences. In essence, Dewey proposals of progressive education, were in sharp contrast to the conventional teacher-centred teaching. In progressive education, there is an expectation of the presence for group-work, joint pupil-teacher planning, evident link between classroom learning tasks and life outside the school and a focus on the mental health of students (commitment to the *whole* child). Alternatives, such as "teacher-centred progressivism" were also developed where some forms of progressive education were integrated within a more traditional approach (Cuban, 1984). I could see possible reasons why teachers would prefer hybrid approaches, such as the need to fit in with wider school cultures, behaviour management and curricular pressures on teaching. Dewey's progressive education was followed by discovery, sociocultural and constructivist perspectives on teaching and learning (e.g. Section 4.2.1)

Towards the 1980s, statistics education research mainly focused on adult reasoning and on probability (Tversky and Kahneman, 1975; Pfannkuch and Wild, 2004) and in the 1990s and 2000s into recommending reformed curricula. Research into teachers' characteristics, such as teachers' thinking and beliefs, teaching processes and practices suggest that such factors can have an impact on students' learning outcomes and beliefs (Calderhead, 2004). On the other hand, effective pedagogies are believed to be influenced by what teachers do, teacher content and pedagogical knowledge and beliefs (Shulman, 1986, 1987; Ball et al., 2008; Callingham and Watson, 2011). However, there is limited evidence available in the literature about the teaching strategies and processes employed by lecturers as they go about their teaching.

Some studies also explored implemented curricula in the university context or with different teaching approaches (Bakker, 2004). Bradstreet (1996) and Garfield (1998) showed that in practice, statistical skills were not always emphasised in statistics modules or textbooks, which instead, in their view, focused primarily on teaching concepts and procedures. The image of statistics teaching that was emerging at the time was that students were able to carry out calculations but encountered difficulties with using statistical concepts, e.g. sampling, variation, hypothesis testing, confidence intervals, statistical graphs (Pimenta, 2006). One solution to students' apparent under-achievement were new methods of teaching and learning that matched new and reformed curricula. For this reason, research into the teaching of statistics at university has been concerned with promoting new styles of teaching.

In addition, the teaching of statistics looked at how students' knowledge could be created through experience, considering that statistical knowledge was a process of transformation, continuously created and recreated rather than acquired or transmitted (Kolb, 1984). Moore (1997, 1998) and Cobb and Moore (1997) proposed that a new pedagogy which stipulated that statistics should be learnt using hands-on practice of analysing data using software. The problems in statistics modules designed for these studies were meant to answer real-world questions using authentic datasets. Further, computer simulations were used to teach various statistical phenomena. So, an interpretation of experiential learning in statistics education is that students should be encouraged to apply statistical skills in a relevant context (scenario). Instead of the 'information transfer' model, a new constructivist[2] view of learning was being promoted which did not view students as 'empty vessels to be filled with knowledge poured in by teachers' but rather able to 'construct their own knowledge by combining their present experiences with their existing conceptions' (Moore, 1997, p. 125). In particular, the focus in these studies was on the planned curricula (e.g. tasks or teaching strategies) and students' achievement as a result of going through particular types of learning. In the statistics education literature, limited information is available about the teaching strategies and processes employed by lecturers as they go about their teaching.

In this section therefore, I review the literature concerned with teacher beliefs and knowl-

[2]Socio-constructivist research in mathematics and statistics education focuses on social and linguistic influences on learning and in particular the relationship between the teacher's instruction and the learners' conceptual understanding (Schmittau, 2004).

edge, teaching methods and approaches and pedagogic strategies. A paucity of research concerned with these issues at university level meant that I also summarised school-based studies yet being mindful of extrapolating the findings to the tertiary level context.

2.5.1 Teacher characteristics

In addressing some of the issues of statistics education, some research studies promoted a pedagogical ‘reform’ (Section 2.4.2.1, p. 24) that involved changes in content, methods of teaching, assessment and attitudes (Garfield and Gal, 1999; Gal et al., 1997). Many of these papers investigated teachers or students’ perspectives, suggesting that changes and developments in teaching practice result in some sort of change in students’ learning (Petocz and Reid, 2003; Groth and Meletiou-Mavrotheris, 2017).

In statistics education, a small number of empirical studies have investigated *teacher characteristics*, such as teachers’ conceptions about statistics, teachers’ beliefs about teaching goals and teachers’ planning of statistics teaching at the elementary (primary) or secondary school (Shaughnessy, 2007; Eichler, 2008b, 2007, 2008a; Pierce and Chick, 2011) and tertiary levels (Petocz and Reid, 2003; Hassad, 2011; Zieffler et al., 2012). In my present review, I focused specifically on studies which investigated teacher characteristics. Since I was interested in investigating lecturers’ beliefs (Chapter 3) and and how lecturers’ beliefs influence teaching practices and activities (Chapters 4 and 5), I also included studies carried out with primary or secondary school teachers where the methodological approach was relevant to my study.

In mathematics education, Philipp defined beliefs as ‘lenses through which one looks when interpreting the world’, or in other words ‘psychologically held understandings, premises or propositions about the world that are thought to be true” (Philipp, 2007, p. 257-259). From this perspective, beliefs are different from attitudes which are to do with emotions. However, Pajares (1992) recognises the challenge in defining *beliefs*, which can be disguised as ‘attitudes, values, judgments, axioms, opinions, ideology, perceptions, conceptions, conceptual systems, preconceptions, dispositions, implicit theories, explicit theories, personal theories, internal mental processes, action strategies, rules of practice, practical principles, perspectives, repertories of understanding, and social strategy” (Pajares, 1992, p. 309). Instead, Pajares defines *educational beliefs* as

> “beliefs about confidence to affect students’ performance (teacher efficacy), about the nature of knowledge (epistemological beliefs), about causes of teachers’ or students’ performance (attributions, locus of control, motivation, writing apprehension, math anxiety), about perceptions of self and feelings of self-worth (self-concept, self- esteem), about confidence to perform specific tasks (self-efficacy), [...] about specific subjects or disciplines (reading instruction, the nature of reading, whole language)” (p. 316)

In statistics, beliefs are identified as one of the dispositional elements of statistical literacy, which comprise

> "...a broad cluster not only of factual knowledge and certain formal and informal skills, but also of desired beliefs, habits of mind, or attitudes, as well as general awareness and a critical perspective." (Gal, 2004, p. 48)

Studies in statistics education have been interested in students or teachers' beliefs about curricula, statistics, the place of statistics in the curriculum, the relationships between mathematics and statistics, the teaching and learning of statistics and the impact of beliefs on teaching activities or students' learning (Gal et al., 1997; Cobb and Moore, 1997; Wild and Pfannkuch, 1999; Begg and Edwards, 1999; Gordon et al., 2007; Sedlmeier and Wassner, 2008; Eichler, 2011). Further, González (2011) regards beliefs held by teachers on statistics, teaching and learning of statistical concepts as fundamental in his model of statistical knowledge for teaching. Studies in tertiary statistics teaching investigated lecturers' conceptions and beliefs about their teaching and about their students, defining components of statistical reasoning and thinking and the development of statistics curriculum content (Chance, 2002; Cobb, 2007; Gordon et al., 2007; Tintle et al., 2011; Newton et al., 2011). Other studies further considered that lecturers' knowledge about curriculum, a combination of subject matter expertise, knowledge about the context of learning, teaching, assessment and the cohort, lecturers' personal, cultural and political values, is often inaccessible.

Subject matter knowledge and expertise is therefore critical for defining tertiary level learning outcomes and content (Section 2.4.3). Sadler (1989) uses 'guild knowledge' to refer to teachers' "ability to make sound qualitative judgements" (p.126) and similar to *craft* or *tacit* knowledge (as used in the context of the Pilot study, Chapter 3). Sadler (1989) argued however that teachers' beliefs about learning outcomes, assessment criteria and grade descriptors are *tacit* and therefore difficult to examine. *Tacit* knowledge is a combination of subject matter expertise, knowledge about the context of learning, teaching, assessment and the student cohort, lecturer's personal, cultural and political values, is often inaccessible to the student and, as I showed in the case of standards at different levels of specificity, difficult to capture in a set of brief statements (Section 2.4.3).

Polanyi defines *tacit knowledge* as "that which we know but cannot tell" (Polanyi, 2009, p. 8) and suggests that "we can know more than we can tell", meaning that it might be difficult to identify and represent tacit knowledge. For Polanyi (1974), knowledge is 'a system of true justified beliefs'. For instance, a statistician is able to interpret an analysis using a perceptual and conceptual skills that lead to even more knowledge. So many aspects of skilled performance, such as that of a statistician, depend on 'knowledge' that is difficult to articulate in a list of statements (e.g. learning outcomes). The polar opposite of *tacit* knowledge is *explicit* knowledge, yet explicit knowledge may be used tacitly. Eraut (2000) identifies several types of situations and several types of knowledge in which tacit knowledge is present. It may be the tacit nature of this knowledge and beliefs about the goals and purposes of higher education modules and programmes leads to the difficulties inherent in articulating assessment standards and criteria (O'Donovan et al., 2004).

Lecturers' beliefs about statsitics and its role in the curriculum might influence statistics educators' curricular planning, for example by conceptualising statistics as a set of procedures or computations instead of a set of investigative processes used in a range of "contexts" or situations (Cobb and Moore, 1997; Wild and Pfannkuch, 1999). Pfannkuch (2008) highlighted that in order to help students develop statistical thinking, lecturers or teachers need to develop their own knowledge of statistical thinking and experience in empirical enquiry cycle. In Pfannkuch's (2008) view, teachers should allocate teaching time *transnumeration* (Wild and Pfannkuch, 1999), i.e. thinking about the different aspects of the story within the data and the decision-making based on representations of data rather than on constructing graphs. A focus on 'desirable' content such as transnumeration thinking, reasoning with statistical models and consideration of variation requires teachers' substantive knowledge of statistics that values the key statistical skills of thinking, reasoning and literacy (Section 2.4.1) and the belief that statistics need to be valued as a distinct discipline.

Gordon et al. (2007) focused on lecturers' conceptions of teaching of statistics to non-statisticians. The authors found that the participants had a range of beliefs about teaching statistics, from a focus on the teacher, the subject matter and course content to a broader focus on the student and their future profession, which were influenced by the contexts, cultures, values, resources and constraints surrounding their teaching. Other studies sought to identify lecturers' views about topics in undergraduate statistics. For instance, Gardner and Hudson (1999) asked lecturers to rank thirty-four topics in order of importance. These research studies either focused on lecturers' situated experiences of teaching statistics or on itemising content while ignoring the context and other curricular elements.

Hassad (2011) developed a 10-item scale to empirically assess and describe the teaching practice of introductory university tutors from the health and behavioural sciences. The resulting two-dimensional scale characterised the teaching either as *concept-based* (constructivist, reform-oriented) or *behaviourist* (less reform-oriented). A higher proportion of tutors with academic degrees in mathematics and engineering appeared to be characterised as *behaviourist*. By contrast, tutors who declared to be members of professional organisations were characterised as constructivists. Hassad explains these two complementary dimension in terms of tutors' beliefs relating to contextual factors (e.g. students' preparedness for statistics, amount of teaching, or class size) which can influence their perception of self-efficacy and affect decision-making regarding the use of particular teaching strategies. Using the survey approach however may not explain why the lecturers were characterised as either concept-based or behaviourist. In another quantitative study that sought to assess statistics lecturers' teaching practices and beliefs about teaching and learning statistics, Zieffler et al. (2012) developed a six-part, 50-item scale. Based on qualitative data from interviews with lecturers to validate the survey items, the study found that the interviewees had very different conceptions about the end points of the Reform–Traditional scale when asked to rate themselves on the four areas of content, technology, teaching and assessment. Zieffler et al. (2012) recognise that observational data collected in the classroom might present a different picture of lecturers' actual teaching practice compared to self-reports. Further, the credibility of survey data depends upon the chosen sample since these two studies used purposive samples.

At the school-level, a few studies have offered empirical data concerning the intended statistics content and reasons for teaching that content from the primary and secondary teachers' perspective (Begg and Edwards, 1999; Watson, 2001). For instance, some studies have focused on the link between intended curricula and teacher belief systems. Using interviews with 34 teachers, surveys and concept maps, Begg and Edwards expected to find a link between teachers' beliefs and attitudes towards and about statistics, their content knowledge of statistics, and their pedagogical knowledge. As the participants in this study, similar to other primary teachers around the world, did not contribute towards setting the official (written) curricula (part of the mathematics curricula), they were unfamiliar with new curricular developments or some statistical content and terminology. For example, in general, the results suggested that these participants focused predominantly on data collection, graphs, data interpretation, and probability, which seemed to fit the official curriculum. However, the participants were also less familiar with the terms mean, median and mode and that they had different conceptions of mathematical and everyday measures of central tendency. Despite apparent issues with teachers' content knowledge, the participants considered statistics to be useful, e.g. since it helps make sense of the world. Although findings from research in the school context may not necessarily apply to university introductory statistics given the differences in the way curricula are set and taught (Sections 2.2 and 2.4.3), these studies highlight the importance of investigating the teaching of statistics from the lecturers' perspective, focusing on their belief systems and context of teaching.

In a series of studies, Eichler (2006, 2007, 2008a,b, 2009) used a theoretical framework to analyse school teachers' intended and implemented curricula (Figure 2.2) and the impact of these practices on students' learning (IAEEA, 1979; Stein et al., 2007). Eichler (2011) focused on the relationships between school teachers' beliefs about statistics, the content statistics teachers intend to teach and the learning outcomes for their every-day classroom practice. In Eichler's framework of belief systems, a characterisation of intended curricula might include 'beliefs about specific content, teaching goals linked to this content, the best way to teach mathematics or statistics, and the way students learn mathematics or statistics' (Eichler, 2011, p. 177). The implemented curricula might involve 'the observable part of the teacher's intended curriculum, transformed by the interaction of the teacher, the students involved and the content within the classroom practice' and finally the attained curricula represents 'the students' belief systems concerning mathematics or statistics that are strongly determined by the classroom practice'.

For example, in a study with 27 pre-service primary teachers Chick and Pierce (2008) found that when asked to plan a lesson for a hypothetical class on a particular topic ('the environment') to teach some aspect of statistics, the resulting written responses and lesson plans matched the participants' reported beliefs about statistics. Participants who did not value data also struggled to plan lessons that emphasised significant statistical concepts. Instead, the lesson plans focused on data collection, rules and procedures (e.g. how to design a graph). Further, although the participants were asked to plan a lesson based on the same contextual information, the lesson plans included a range of content and teaching methods, highlighting that the same task could result in very different teaching practices, which might depend on teachers' beliefs about statistics, students or learning.

For my study, I found Eichler's framework also relevant in the context of tertiary statistics education. A characterisation of teaching statistics at university might involve for instance lecturers' beliefs about curricula, statistics and students' learning as well as a focus on observations of teaching in order to capture the interactions between curricula, lecturers' planning, teaching and students' learning.

Using interviews with school teachers, Eichler (2007) proposed five aspects that shaped teachers' curricular planning: (1) the content of statistics teaching, (2) the teachers' objectives of statistics instruction, (3) the teachers' objectives of mathematics instruction, (4) the teachers' beliefs about the students' benefits of statistics instruction, and (5) the teachers' beliefs about effective teaching of mathematics. Such aspects of teachers' beliefs could also be explored in the context of tertiary statistics to characterise the links between curricular planning and beliefs.

Further, this research identified four types of teacher individual stochastics (mathematics, probability and statistics) curricula: *the traditionalists*, *the application-preparers*, *the everyday-life-preparers* and *the structuralists* (Eichler, 2006, 2007). At one end of the scale, *the traditionalist* teachers focused on establishing a theoretical base for mathematics comprising of algorithmic skills and insights into the abstract structure of stochastics in the absence of applications. *The application-preparers* focused on students grasping the interplay between theory and applications. *The everyday-life-preparers* aimed to prepare students to deal with real-world situations in a variety of contexts by developing methods from real-life applications, with a focus on statistics rather than probability. Finally, *the structuralists* used applications to exemplify statistics concepts. The teachers' subjective knowledge and conceptions about stochastics, about students' learning and teaching stochastics were considered to be part of teachers' belief system. Although this study was carried out in the context of school-based teaching where statistics is taught as part of mathematics curricula, it highlights that teaching statistics might involve specific teaching styles characterised by the mathematical content, statistical theory and the "context" of problems and examples included in teachers' planning and delivery of teaching.

In a study that investigated secondary teachers' intended (individual) curricula in terms of content and learning objectives, Eichler (2008b) found that although the content across the four participating teachers was similar and that matched the official (written) curricula, their learning objectives (intentions) were different. It further appeared that the learning objectives influenced the participants' teaching practices. For instance, one of the participants in the study was keen to promote the role of "context" in statistics through both a process-oriented or dynamic view of statistics and real statistical problems, while the other three participating teachers focused more on pseudo-realistic contexts. The participants' focus on "context" was also associated with the teachers' beliefs about statistics as well as institutional demands concerning the final examinations. This study highlights the influence of context of teaching and learning and the need to consider these issues together in order to understand behaviours in the "everyday" classroom (Jacob, 1997).

In mathematics education, Thompson (1984) for example considered that the first step in characterising school teachers' classroom practices was to understand the teachers' beliefs and conceptions and how these were related to their teaching practices. So information about teacher's belief systems and curricular goals is considered important in understanding teaching practices as well as improving students' attainment and attitudes in mathematics (Thompson, 1984; Hiebert and Grouws, 2007). In my study, this approach was particularly relevant since I was interested in depicting a model of the teaching of statistics Also at secondary level, the Second International Study of Mathematics or SIMS considered that analyses of intended curricula (official curricular goals) need to be conducted in order to interpret the findings of the classroom processes (implemented or enacted curriculum) and student outcomes (attained curriculum) of a study and the relationships between these three elements (IAEEA, 1979). In the framework proposed in SIMS, teachers' beliefs and attitudes included a variety of factors which were considered to influence the implemented curriculum, understood as the ways teachers interpret teaching situations, including internal representations of the mathematical subject matter, beliefs about specific content, about teaching and about student learning of mathematics or statistics.

Other studies have attempted to find a link between teachers' content knowledge (the intended curricula) and student achievement. In a quantitative study using measures of teachers' knowledge, Callingham et al. (2015) used multilevel modeling of scores from 789 secondary school students on three tests and scores from 36 of their teachers on one instrument designed to measure pedagogical content knowledge. Although the student data used was naturally occurring and diminished in size over time, the study supported the conclusion that teachers' pedagogical content knowledge in statistics was associated with their students' learning outcomes but there was a negative effect over time in that the effect of teacher quality diminished as students progressed through school.

Another research strand investigated statistics teachers' intended curricula in relation to their students' learning difficulties (see Section 2.6.4). Cardona (2008) for instance looked into secondary school statistics teachers' understanding of students' difficulties in statistical reasoning with respect to *representativeness* and the teaching approaches teachers believed to help students' statistical reasoning. Using interviews, the participants were asked to anticipate students' answers and difficulties in reasoning on five tasks based on the conjunction fallacy, the law of small numbers, and the gambler's fallacy (Tversky and Kahneman, 1983; Kahneman and Tversky, 1972; Tversky and Kahneman, 1975). The data showed that participating teachers underestimated students' difficulties and considered the context of the task and students' prior knowledge to be important to students' understanding. Importantly, experienced teachers seemed to 'describe more articulated and more integrated interventions' than novice teachers. While an inexperienced teacher only suggested one strategy for helping students' learning, make a tree diagram, the experienced teacher proposed more actions - tactile simulations, technology-based simulations, conceptual explanations and comparing the experimental probability with the theoretical probability. Finally, experienced teachers were more able to use heuristics of representativeness than inexperienced ones.

In summary, there seems to be fairly strong consensus regarding the relationships between teachers' beliefs and curricula (Figure 2.2). The next challenge in researching statistics teaching however is to provide evidence, as documented in education research studies, of improvements in student learning outcomes and of the connection amongst intended curricular, implemented teaching, context and culture (Hiebert and Grouws, 2007; Eichler, 2011). The focus in my research is on this latter challenge. A study investigating statistics lecturers beliefs about their intended curricula might shed light into the way curricula are understood by stakeholders. A qualitative study using interviews with lecturers and observations of teaching can help identify teacher characteristics and can explain how such characteristics manifest in different contexts (Groth and Meletiou-Mavrotheris, 2017). In my pilot study in Chapter 3 therefore, I aim to investigate statistics lecturers' beliefs and conceptions about the content and outcomes of statistics courses. Since my focus ultimately is on characterising teaching practice, another issue which gained in importance following this literature review and my own observations was the implications of "context" for teaching practice.

2.5.2 Observational studies of teaching processes

A small number of studies focused directly on the *process of teaching* and the roles and behaviours of teachers. Classroom observations while teaching is taking place seem to be the best way to provide evidence for teaching processes (Gage, 2009). At university level, there are few observational studies of teaching process and practice (Weber, 2004; Speer et al., 2010; Viirman, 2015). Here, I specifically discuss studies carried out at university level which are relevant in the context of my study.

In mathematics education research, a number of studies sought to understand the teaching of mathematics at university, focusing on the relationships between mathematics and pedagogy (Nardi, 1998; Jaworski, 2002; Nardi et al., 2005; Thomas, 2011). For example, Jaworski (2002, Section 2.3) focused on tutors' activity (processes), tutor-student interactions and tutors' thinking about their teaching in first year mathematics tutorials. Using data from tutorial observations, teaching resources and interviews with tutors, the study was able to offer interpretations of tutor knowledge (of students, of mathematics, of pedagogy). By applying a teaching triad as an analytical tool to categorise the teaching in undergraduate tutorials in relation to three elements, management of learning, sensitivity to students and mathematical challenge (Jaworski, 1994), the study characterised what the teaching of university mathematics can or could involve rather than evaluate it.

In an observational study of mathematics teaching at university level and adopting an activity theory perspective with a grounded analytical approach, Thomas (2011) and Jaworski et al. (2009) investigated the teaching practice to connect it to specific module content. The studies looked at the nature of linear algebra to present a model of the teaching process. The emerging hierarchical model comprises five *goals* or aims, from the initial engagement with mathematics, to intuitive understanding, acquisition of mathematical language, conceptual understanding and mathematical competence. Each of these hierarchical goals is linked to teaching *actions* or strategies, which include for example presenting an example, providing notes or allowing for breaks in order to engage the students. Thomas (2011) further identifies four *operations* or processes that the lecturer in her case-study performed in lectures: providing mathematical content, maintaining a physical presence (being in the lecture theatre), writing and talking. Taking a different approach to my research design and within statistics education, in my study I sought to also look at lecturers' aims for teaching and processes of teaching.

Some studies have characterised lecturing (discussed in more detail in Section 2.6.1), rather than research its effectiveness. Taking a grounded theory approach to data analysis and focusing on the mathematical content of one lecturer, Weber (2004) described three *teaching styles* used by one mathematics lecturer to construct proofs. Weber justifies the lecturer's action in the classroom as the result of the lecturer's beliefs about mathematics, students and education and knowledge of the subject matter being covered in lectures. The lecturer's actions in the classroom can therefore be interpreted and are dependent on complex relationships between knowledge, skills, learning goals, and beliefs.

Adopting Sfard's commognitive framework (Sfard, 2015), Viirman (2015) investigated the discursive practices of seven university mathematics lecturers. Viirman (p. 1177) identified patterns of discourse as a categorisation of three types of didactical routines and sub-routines:

1. *Explanation routines*: known mathematical facts, summary and repetition, different representations, everyday language, and concretisation and metaphor.
2. *Motivation routines*: reference to utility, the nature of mathematics, humour and result focus.
3. *Question posing routines*: control questions, asking for facts, enquiries and rhetorical questions.

To some extent, the 'routines' proposed in Viirman's studies correspond to Bellack et al.'s cycles of 'moves' or exchanges between the teacher and the student. For example, structuring 'moves' are similar to explanation routines since they both involve explaining, teaching to the whole class known facts, or summarising for students. However, structuring is also about the organisation of content. Soliciting, responding and reacting are similar to the question posing routines. Viirman's analysis also shows how the lecturers' teaching practices and processes present similarities as well as differences.

Further, in a large-scale observational study involving 1,017 classrooms, Sirotnik (1983) identified a similar uniformity of cycles of teaching. The use of cycles of teaching in Bellack et al.'s study is also an opportunity to link the content of teaching to the process of teaching, part of a teacher's actions and behaviours and style of 'lecture' and 'lecturing'. Given the generality of the routines included within the classification of discursive routines and the similarities to Bellack et al.'s studies on school-based teaching, one might expect to identify such routines or moves in other subjects, including statistics lectures. Brown et al. (1984) differentiates between six types of *lectures* and five types of *lecturing*. The study shows preferences for particular styles of lecturing, defined as a "person's habitual mode of responding to a similarly perceived task" across four broad subject areas. In a later study, Saroyan and Snell (1997) aimed to describe characteristics of lectures at a more detailed level. Using data from observations of three lecturers in the medical sciences, Saroyan and Snell (1997) characterised the teacher-centred lectures as content-driven and student-centred lectures as context-driven and pedagogy-driven. The study found that the pedagogy-driven type of lectures are highly rated by students.

School-based observational studies have investigated the process of teaching since the 1960s. In an analysis of fifteen teachers' teaching practice in fifteen schools, Bellack et al. (1966) analysed teaching in terms of *cycles* or sets of exchanges between the teacher and the students. Each cycle consists of various combinations of "moves" during the lesson: structuring, soliciting, responding and reacting. *Structuring* moves set the context for subsequent learning. *Soliciting* moves include verbal exchanges and cognitive and physical responses. Responding are directly related to soliciting. Finally, *reacting* moves involve clarifying, summarising, expanding or giving feedback in relation to responding and soliciting. Cycles of moves are repeated in a variety of ways and different combinations of moves (one or more moves can be repeated or omitted within a cycle). Importantly, Bellack's research found applicability in a range of studies and across subjects (Gage, 1978; Sirotnik, 1983). During the teaching observations of my study, I became interested in looking at the exchanges that take place in the lecture theatre and see how such an approach that looked at "moves" could be applied in a different context. However, the nature of interactions between lecturers and students might be slightly different in lectures at tertiary level.

2.5.3 Pedagogical strategies

In his model of statistical literacy, (Gal, 2002) proposed two broad yet interrelated components: the knowledge component and the dispositional component (Section 2.4.1.1). However, one of the challenges in statistics education is to address students' lack of motivation for and difficulties with learning statistics (Garfield, 1995; Garfield and Ben-Zvi, 2007). Teachers of statistics need to account for students' challenges with statistics both while planning and while implementing curricula. One way to address challenges students encounter with learning statistics is to use a range of pedagogic strategies such as lecturers/teachers' explanations, questioning, teaching discourse and strategies for motivating students to learn statistics.

2.5.3.1 Explanations

Explanations and context are considered to be key tools for understanding a discipline such as statistics (Makar and Rubin, 2009). Explanations in statistics involve teachers and students' experiences and knowledge of the context and of the statistical tools and ideas to support statistical reasoning and thinking (Wild and Pfannkuch, 1999; Gil and Ben-Zvi, 2011). Further, statistical reasoning involves understanding and explaining statistical processes and also interpreting statistical results (Ben-Zvi and Garfield, 2004). In effect, teaching statistics “in a memorable way” is about “being able to identify related elements in the discipline and to explain the elements in a way that highlights the relationships” (Sowey, 1995). Sowey suggested that teacher’s explanations of statistical concepts or of the usefulness of statistics can aid students’ retention of statistical ideas in the long term. In the case of lecturing (Section 2.6.1), Bligh (1972) considered that “understanding consists of relating ideas to ideas we already possess” (p. 112). In the case of lecturing, Bligh focused on the importance of making links across the curriculum to students’ understanding of the subject matter by using reasons as explanations. Bligh recorded eight types of explanations used by lecturers to build up explaining links, including *regulative* explanations which explain links according to rules, *analytical* which explain the constituent parts and how they link together, *spacial* which involve visual aids, *temporal* which are represented spacially on a time dimension, *kinematic* which assumes that object(s) have moved over time, *causal* involving before and after explanations, *functional* which emphasise relationships and generalisations about effects, and *mental* explanations which are types of metaphors.

On reviewing textbooks on factor analysis, Chong et al. (2012) suggests that misconceptions may arise when explanations of terms such as “eigenvalue” are omitted. For example, using “rotating factors to better interpret the data” may lead students to develop a non-technical understanding using everyday, non-statistical meanings. For example, vectors representing factors are rotated in subject space. Without understanding of vectors and subject space, students might represent this rotation as spinning a plot to get a better perspective. At the same time, using technical terms to explain a term such as “factor” might overwhelm students. In their animated tutorial, Chong et al. propose that renaming certain terms, expanding on the explanations of those terms and visualising eigenvectors in subject space can help students conceptualise factor analysis. This approach to teaching factor analysis suggests that teachers’ explanations, including choice of terminology and visual tools, are fundamental to students’ statistical meaning-making.

In a study investigating in-service undergraduate statistics students’ self-explanations, Hall and Vance (2010) found that students who made higher gains in statistical problem solving from pre-test to post-test also gave more thoughtful and elaborated responses than students with lower gains. This might mean that encouraging students to explain their responses might help with developing statistical reasoning. Similarly, in her statistical literacy module, Chance (1997) grade students’ explanations of what they have learned both in writing (by grading students’ journals) and to peers. Such studies emphasise the importance of explanations in making sense of the statistics and in improving communication skills in statistics (Tempelaar et al., 2017; Khachatryan and Karst, 2017).

2.5.3.2 Questioning

Despite the demonstrated importance of students' explaining their thinking, teacher-centred instruction consitutes an important approach in undergraduate statistics teaching (Zieffler et al., 2008). Teachers' questions can be used for a number of purposes, including getting students' attention, initiating dialogue, reviewing content, organising a specific task or encouraging students' self-explanation. For instance, Pedrosa de Jesus and da Silva Lopes (2011) carried out a discourse analysis focusing on the frequency, the cognitive level of teachers and students' questioning practices, the way teachers dealt with the absence of a solicited student answer and the nature of teachers' reaction to a student intervention (question or answer). In this way, the study characterised the nature of lecturers' self-answers and dialogic attitudes using observational and interview data.

In mathematics lectures, the majority of lecturer questioning consists of short-answer, low-level questions that require students to recall facts, rules and procedures instead of high-level questions that require students to draw inferences and explain ideas. For instance, in an investigation of mathematical lecturing discourse using Sfard's (2015) commognitive framework, Viirman (2015), investigated question posing routines, including control questions, asking for facts, enquiries and rhetorical questions. In this way, Viirman was able to characterise the teaching of function at university level and offered a model for identifying patterns in university teaching discourse. In statistics education however, less is known about the nature of teachers' questioning to elicit students' statistical reasoning (Groth, 2017). However, questioning during lecturing may encourage students to develop a *questioning attitude*, considered essential in statistical literacy and reasoning (Watson, 1997; Gal, 2002).

2.5.3.3 Teaching discourse

Part of teaching discourse, narratives are means of representation (see Section 4.2.1.5, p. 125), including telling evocative stories that intrinsically involve statistical thinking. Representations of practice have to be learnt when becoming a member a particular professional community of practice and used correctly (Goodwin, 1994). For instance, statistical stories considered to be effective are interesting, topical, and have a moral, e.g. use of statistics often achieves desirable outcomes or correct decisions (Martin, 2003). So, an understanding of how teachers use narratives in their teaching is considered important for characterising their teaching (Preskill, 1998). Since humans are storytellers, studying teachers' narratives is about how they experience the world and represent the *context* for making meaning of teaching situations (Connelly and Clandinin, 1990). In this view, curriculum includes teachers' narratives as metaphors for teaching-learning relationships (Pulvermacher and Lefstein, 2016).

In statistics education, meaning might not represent anything real or concrete (Martin, 2003). Instead, the way data behaves lead to invention of statistical terminology. Teacher narratives in statistics education have been used as teaching devices to enhance students' learning (DelMas, 2004; Gentner and Holyoak, 1997). Narratives, which include analogies (explicit comparisons), can be used for a range of purposes, including constructing explanations and building arguments (Gentner and Holyoak, 1997; Martin, 2003).

For example, Groth and Bergner (2005) investigated statistics trainee teachers' use of metaphors to characterise the structure of their content knowledge. A metaphor is a type of narrative which links one domain of experience with another seemingly disparate domain and creates meaning from the connection (Kovecses, 2004). Groth and Bergner (2005) also expected that teachers' ability to construct metaphors can provide some insight into the role of teachers' content knowledge and pedagogical content knowledge in teaching statistics (Shulman, 1987). In this qualitative study, fifty-four trainee teachers were asked to write a metaphor for the concept of statistical sample. This exploratory study found that the trainee teachers' ability to create metaphors varied from metaphors which captured a number of important attributes of statistical sample to those which did not. Groth and Bergner thus concluded that teachers' metaphors are indicative of their statistical thinking. However, the research studies reviewed focused on school-based teaching and less is known about lecturers' explanations of statistical concepts, questioning or teaching discourse (e.g. narratives and metaphoric thinking) using empirical evidence from 'natural' lecturing contexts, i.e. *during* teaching (Makar and Confrey, 2004).

2.5.3.4 Motivating students

Garfield (2017) recommended the use of small group activities rather than individual activities to motivate students to complete work and promote cooperation amongst group participants. In addition, examples that have recently appeared in the media, government reports or news are used to promote students' engagement with statistics (Garfield, 2017; Watson, 1997; Gal, 2002). Tishkovskaya and Lancaster (2010) developed a statistics curriculum to address motivation by including media reports, examples of misleading statistics and other interesting material about real-world situations. In this way, Tishkovskaya and Lancaster aimed to demonstrate to students the usefulness of statistical knowledge in understanding the world around them.

Another strategy used to ensure students' engagement with the module content is to develop tasks within different contexts so that students can apply the statistical content of a module in a variety of ways and be able to consider real world contexts and applications. For example, Chance (1997) suggested that computer-based exercises, projects with presentations and peer reviews, take-home final exam questions and student journals enables students to experience the statistical investigative cycle (from posing the research question to presentation of results and interpretation). Similar to Tishkovskaya and Lancaster, Chance (1997) argued that the tasks are effective as learning tools as well as motivational tools.

Chance (1997) and Allen et al. (2010) further argued for changing the focus of statistics curricula from mathematical calculations to using real data from the students' programme of study. Authentic tasks based on real data and contexts may help students understand why a good understanding and knowledge of statistics is important for their discipline, engage them and improve their attainment. However, in the context of mathematical modelling, Vos (2011) shows that what is perceived as "authentic" is constructed as an interplay between the authenticity of the non-mathematical aspects of the task as well as the mathematical model. In Vos's (2011) interpretation of authenticity, some aspects of a task may be authentic, while other aspects may be included for educational purposes. In statistics education, a similar situation may arise where the data originates from the 'real' world, but some aspects are *adapted* for pedagogical purposes. Further, a dataset may be simulated for pedagogical purposes (e.g. changing the number of variables or data points) within a realistic context. The literature reviewed for this study was unclear on which types of tasks or aspects of a task are most effective in improving students' motivation and engagement in various contexts.

2.6 Research into the impact of teaching methods on student learning

In addition to curricular issues, the impact of pedagogic resources and approaches on student learning, i.e. how lecturers help students learn statistics at university, and cognitive and non-cognitive factors that impact on student learning of statistics are important research strands in statistics education. This research highlights the importance placed in statistics classrooms on teaching methods, teaching resources (problems), problem contexts and technology used in the statistical analysis process.

2.6.1 Lecturing versus other teaching methods

In general, research has described the process of teaching as traditional teacher-centred and "reform" student-centred approaches as defined by variables such as classroom organisation (e.g. desks), grouping of students, activities or classroom talk (Cuban, 1984). Traditional ways of teaching statistics at undergraduate level normally combine lectures, also discussed in Section 2.5.2, with laboratory sessions in the form of lecture-discussion-homework-test. Traditional curricula covered in lectures often includes a set of techniques and definitions focusing on formal inference based on probability theory (Tintle et al., 2011). Perceived as more teacher-centred, lecturing as a teaching method is however still prevalent in higher education institutions (Rodd, 2003; Huxham, 2005). Lecturing has been defined as "continuous periods of exposition by a speaker who wants the audience to learn something. That learning may be of different kinds" (Bligh, 1972, p. 6).

Based on an overview of research, Bligh proposes that lectures, if used appropriately, can be effective at achieving four objectives: 1) the acquisition of information or transfer of facts, 2) the promotion of thought and the development of critical thinking, 3) changes in attitudes such as the acquisition of values, adjusting to a professional role, professional ethics and scientific integrity and 4) behavioural skills such as practical work. Bligh's book however shows how different studies came to opposing conclusions regarding the effectiveness of lecturing.

Some studies have looked at how best to deliver lectures to improve student learning. Using action research, Wulff and Wulff (2004) investigated how to use communication to enhance teaching and learning in statistics, i.e. to achieve the learning outcomes through lectures, problem-solving activities, the use of office hours, and students' written reflections about their learning on an introductory statistics module. Using student feedback, interviews and email questionnaires, the students in this study considered important that lectures were clearly structured and that the lecturer moved away from the traditional style and instead used interactive examples with the students. These participants also considered that the problem-solving activities contributed most to their learning, especially the session following the lectures when students were able to practise with the material, although the study found that students were still relying on the lecturer to demonstrate the techniques. Wulff and Wulff (2004) concluded with the recommendation that lecturers should "no longer spend time [in lectures] trying to cover the content, but rather trying to uncover the material" (p. 98).

In the context of mathematics education, Rodd (2003) investigated the use of lectures in undergraduate mathematics from the students' perspective. In her study, Rodd suggested that students feel inspired by lectures. However, student enjoyment of a learning activity such as lectures does not necessarily mean that it translates in improved student learning when compared to other methods. Rodd sees the role of lectures to be to enthuse, motivate, challenge and lead by example. In this view of lecturing, the lecture theatre is a place of 'mathematical awe and wonder' which contributes to students' enculturation within a mathematics community and can serve a 'spiritual' purpose. As 'participatory witnesses', students are kept 'rapt' by the mathematical ideas presented and are 'lifted' to reflect, apply the content and review afterwards (Rodd, 2003, p. 18). Rodd challenges the view of lectures as relying on unsound transmission of knowledge.

Critics of the lecture method however claim that lecturing is teacher-centred and provides an authoritarian social context compared with student-centred discussion methods, perceived to be more democratic (Gage, 2009). The 'reform movement' in statistics education, that has led to the focus on statistical literacy, reasoning and thinking and the randomisation-based curriculum promoted less statistical theory and probability and more focus on data analysis (often without mathematics), technology and situated or active learning, i.e. fewer lectures and more collaborative problem solving (Ben-Zvi and Garfield, 2004; Garfield and Ben-Zvi, 2007). Bloom (1953) for example contrasted students' thought processes during lectures and during small group tutorials (discussions). Bloom found that students reported sixty-four percent of the thoughts to relate to the ideas presented during lectures while fifty-five percent did so in discussion groups. Only one percent of the students in this sample reported to attempt to solve problems, synthesise or inter-relate information.

The pedagogy proposed in the consensus curriculum of the 1990s or the randomisation-based curriculum of the 2000s was based on theories of learning that emphasise the need for students to construct their own understanding and expressed negative views on the traditional, teacher-lead approach that relies on lecturing as a teaching method (Cobb, 2007;

Tintle et al., 2011, 2012; Rossman and Chance, 2014; GAISE, 2016). Some studies proposed that courses which aim for students to develop statistical thinking and improve students' attitudes towards statistics need to lecture less and instead include active learning (Dixon and Judd, 1977; DelMas et al., 1999; Cobb, 2000). Further, in her teaching of university statistics, Garfield stated that

> "In my classes, I do not lecture at all, which takes a while for students to adjust to. Instead, students are required to read the textbook before coming to class, guided by a study guide/student handbook I have written containing study questions, sample problems, etc. When students come to class each day we first discuss the study questions, often arguing about issues such as which is the best measure of center to use, which type of plot gives the most information, etc. They rapidly learn that there is often not one right answer nor one way to solve a problem." (Cobb, 1992, p. 16)

In this depiction of her teaching, Garfield appears to suggest an active learning teaching method as an alternative to lecturing. Active learning has been defined as a teaching method that "engages students in the process of learning through activities and/or discussion in class, as opposed to passively listening to an expert. It emphasizes higher-order thinking and often involves group work" (Freeman et al., 2014).

Hogg (1990) highlighted that some of the problems with introductory statistics modules were the gap between the teaching of statistics, 'often stagnant' relying on traditional lecture/discussion, the teachers' reluctance to change their approaches and the rapid progress of statistics due to the advent of technology (Ben-Zvi, 2000). Hogg (1991, 1990) proposed that undergraduate teaching of statistics should encourage teamwork, build a 'sense of community' and allow interaction among students and teachers and connect statistical methods to the underlying philosophy and real life applications. As a group of thirty-nine statisticians, they recommended four steps to teaching the subject: (1) state the goal(s), (2) analyse data and do projects, (3) use the computer and (4) lecture less, teach more.

Within university mathematics for example, Iannone and Nardi (2005) used focus group interviews with lecturers to gain insights into their thinking regarding the teaching of mathematics. The mathematics lecturers emphasised the value of learning through *interaction*, and recognised that lecturing is not a method generally conducive to interaction. Instead, the participants recognise the value of seminars, tutorials and student feedback which are more prone to interaction.

Research comparing lecturing with active learning suggests that students learn better, i.e. obtain better test scores at the end of a module, if the teaching involves interactive, collaborative tasks (Freeman et al., 2014). In a systematic review of 225 studies, Freeman et al. (2014) found that on average, student performance increases by just under half a standard deviation with active learning compared with lecturing across a range of science, engineering, and mathematics course types. Other quantitative studies found that students took longer to graduate and achieved less credits when the lecturing time increased due to the interactions between student/teacher and time spent on independent learning (Van Den Berg and Hofman, 2005; Schmidt et al., 2009).

In order to simulate a teaching and learning environment more closely related to how statisticians work in practice, Holcomb et al.'s (2010b) randomisation-based curriculum proposed that the learning process should consist of five elements: (1) research study and data, (2) tactile simulation (e.g. flipping a coin) and class discussion of results, (3) simulation using a tailored applet, (4) understanding statistical significance, what an p-value is and recognise factors that affect p-values (5) discussion and conclusion in context. A task in the randomisation approach would involve the whole class taking part in a classroom experiment to answer a research question (e.g. how likely is a particular result under the null hypothesis model). The group is then asked to record and combine data (e.g. originating from flipping a coin) to simulate and determine the proportion of repetitions from the experiment (Holcomb et al., 2010a,b).

For example, Tintle et al. (2011) took such an active-learning approach and implemented the GAISE pedagogy (also in Section 2.4.2.4) with a randomisation-based introduction to logic of inferencing, i.e. using simulation, bootstrapping and permutation tests to learn statistical inference instead of the *consensus curriculum* of the 1990s that focused on asymptotic sampling distributions (central limit theorem) (Cobb, 2007; Rossman and Chance, 2014; GAISE, 2016). On this algebra-based introductory statistics module (240 students), all classes took place in the laboratory, with students working on large-scale group research projects. The module materials included real research studies, articles and data, while de-emphasising the use of symbolic notation and mathematical formulae. The authors evaluated the module using the CAOS test (DelMas et al., 2007; Garfield et al., 2012b). When comparing the CAOS scores of two cohorts, one following a traditional curriculum and the other a randomisation-based curriculum, Tintle et al. (2011) suggested that the new curriculum improved student learning compared to the *consensus curriculum* and that these gains were retained after the module (Tintle et al., 2012).

Further, Tintle et al. (2014) found significant yet small increases in students' affect, cognitive competence and perceived difficulty when comparing a radomisation-based curriculum group (425 students) to another large group of students (2,200 students) on a consensus curriculum . However, the change in scores for both cognitive competence and difficulty was not statistically significant.

It seems that so far the results of these studies looking to evaluate the value of particular curricula are inconclusive, relying on small sample sizes and using basic statistical analyses. With opportunity samples which are not randomly selected from existing populations and without randomly assigned treatment groups, it is not possible to draw causal inferences. Further, the studies reviewed so far did not take into account potential covariates between samples that might affect students' non-cognitive and cognitive performance (Section 2.6.4.1). For example, teacher characteristics such as experience, pedagogical approaches, institutional cultures might influence students' performance and enjoyment of the subject (Gage, 2009, also in Section 4.2.2).

The work of Ben-Zvi and colleagues is an example of a research programme that investigated how the learning environment can promote statistical reasoning about particular concepts, e.g. data analysis or data representations (Ben-Zvi and Arcavi, 2001; Ben-Zvi, 2004; Abrahamson, 2009; Ben-Zvi et al., 2012). By using classroom investigations, video recordings, in-depth interviews with teachers, tutorial clinical interviews with students, review of students' notebooks and research projects, the authors focused on how students develop an understanding of statistical concepts over time by solving statistical tasks in dialogue with teachers. In this way, the authors provide a description of what it means to learn to reason about statistical concepts that have implications for curriculum design and teaching. For instance, challenges for teaching and assessment of students' understanding would stem from the existence of multiple goals for students, the mismatch between real-world contexts and statistical contexts, the role of the technology in learning statistics and the need to work collaboratively (i.e. in groups) versus individually. The authors suggested that group projects using extended tasks provide an authentic experience when learning statistics (Ben-Zvi, 2004). This work was however carried out with secondary school children and it appears that less is known about how learning is constructed at more advanced educational levels.

However, some approaches to active learning, e.g. based on classroom discussions and interactions, might be more suitable with small cohorts of students. Some studies have recommended replacing the lecture format with other resources such as videos or tasks. For instance, Price (1990) replaced the lecture format in an introductory psychology class with short assignments using readings of primary sources. The assignment required students to 'integrate, for other college students, central ideas in the text and/or lectures with ideas in the assigned primary sources, in a three-page paper' (p. 51). Based on feedback from students, the study suggested that students were enthusiastic about the module while their performance was comparable to that of students following a traditional lecture format. However, it is worth noting that it was a challenge to compare the grades of two groups of students following different teaching approaches when there is no random allocation into groups. It seems that some of these studies measured particular objectives which might focus on content that is not necessarily worth while for the students (e.g. procedures versus statistical thinking) or might lead to low student motivation. Teaching methods cannot be considered in an objective way, detached from institutional and societal culture.

Research outside statistics education further considered that lecturing is just as effective as other methods to teach information (Bligh, 1972) and recommended combining lectures with other methods (discussion methods, independent study, enquiry methods and so on. For example, on an introductory statistics module, Abbott and Falstrom (1977) found that adding frequent weekly testing, content unitisation, and learning objectives to a lecture was as effective (similar course achievement measured as cumulative points earned in the course) as personalised methods such as contact with a tutor, non-lecture unitised material for students to learn, and criterion-related mastery learning at individual rates. Using particular activities during lectures, such as 'interactive windows' during lectures may also have a positive effect on students' test results (Huxham, 2005).

The studies reviewed in this section used to examine the effectiveness of lectures, e.g. within the four areas proposed by Bligh (1972), include cognitive ability measures (students' marks on factual content, problem-soling) or non-cognitive measures (e.g. using questionnaires of student interest in the subject or students' use of textbooks). The research reviewed here suggests that findings from studies investigating the effectiveness of lecturing varies. Some of the studies investigating lecturing are small scale and might not take into account contextual factors (e.g. lecturer's background, discipline, students' intentions and activity in lectures). Further, student variables, such as preparedness for the class, can affect student performance on the tests used to measure the effectiveness of lectures.

Lecturing is generally combined with other teaching methods and approaches so the context of the module also matters when investigating teaching methods at university. In the end, research in higher education seems to suggest that the lecturer's skill is paramount for achieving Bligh's (1972) four objectives (p. 54). For example, a lecturer might use the lecturing time for questions and answers or for engaging students in active learning during or after the lecture.

This overview of studies debating the effectiveness or otherwise of lectures and of lecturing appears to propose conflicting views. In the field of statistics education, researchers and educators seem to believe that active learning methods better promote the practice of statistical skills. It is not clear however how lectures should be combined with practical work when teaching statistical skills. Since achieving the learning outcomes can depend on the interaction between the lecturer and the "context", such as the students, it is difficult to conclude which techniques are required. In my study, I do not seek to prescribe what should occur in order for students to achieve the learning outcomes but rather describe or characterise how teaching does occur based on observations and interviews with lecturers.

Statistics educators and researchers recognise the gap between scientific statistical practice, research into pedagogy of statistics and actual teaching of statistics (Ben-Zvi and Makar, 2012). The work of Ben-Zvi and colleagues is an example of a research programme that investigates how the learning environment can promote statistical reasoning about particular concepts, e.g. data analysis, data representations (Ben-Zvi and Arcavi, 2001; Ben-Zvi, 2004; Ben-Zvi et al., 2012; Abrahamson, 2009). By using classroom investigations, video recordings, in-depth interviews with teachers, tutorial clinical interviews with students, review of students' notebooks and research projects, the authors focus on how students develop an understanding of statistical concepts over time by solving statistical tasks and in dialogue with teachers. In this way, the authors provide a description of what it means to learn to reason about statistical concepts that have implications for curriculum design and teaching.

For instance, challenges for teaching and assessment of students' understanding would stem from the existence of multiple goals for students, the mismatch between real-world contexts and statistical contexts, the role of the technology in learning statistics and the need to work collaboratively (i.e. in groups) rather than individually. The authors suggest that group projects using extended tasks provide an authentic experience when learning statistics (Ben-Zvi, 2004). This work was however carried out with secondary school children and it appears that less is known about how learning is constructed at more advanced educational levels.

Beyond using teaching resources specifically designed for classroom use, statistics educators have also investigated the possibility of bringing statisticians' practice into the classroom in order to create a more 'authentic' learning experience for the students. Curricula based on statisticians' professional practice (Section 2.4.2.3) promotes learning with similar features to actual statistical practice, "organised around joint accomplishment of tasks, so that elements of the skill take on meaning in the context of the whole" (Resnick, 1987b, p. 18). Using technology and project-based learning, Blades et al. (2015) aimed to allow students to gain experience in the practice of statistics by experiencing the whole statistical investigative cycle.

Studying students' integration of statistical workplace-based knowledge, Bakker and Akkerman (2013) showed how students and teachers were able to make hidden processes overt through discussions and reflection. Some studies therefore promote teaching the accumulation of statistical skills similar to an apprenticeship, that encourage student observation between school and workplace/real-world applications of statistics.

In a series of studies, Derry and colleagues focused on statistical learning while students engaged in situated simulations of professional activities (e.g. conducting experiments, collaborating on tasks, presenting and debating). Derry et al. proposed a statistics course that focused on statistical authenticity, including simulations of real-world problem situations requiring statistical thinking (Derry et al., 1995, 2000; Bakker et al., 2008). The pedagogical strategies involved the production of 'a microcosm of a productive problem-solving community within larger society' in which students could negotiate problems emerging in their own adult communities (e.g. smoking ban in restaurants). The authors viewed statistical authenticity and the teaching approaches required to deliver this aim alongside two dimensions, cultural relevance and social activity. Teaching strategies with high social and cultural relevance included learning in situ (cognitive apprenticeships), situated simulations and collaborative analysis of relevant data. Lectures with relevant examples were lower in social activity than in cultural relevance, while decontextualised lectures and analysis of contrived data were low in both social and cultural activity.

In this model of statistical teaching, decontextualised teaching, involving for instance teacher-centred delivery of abstract statistical formulae, was assumed to be the least engaging form of statistical teaching since it was considered to be 'culturally irrelevant' and 'socially passive'. Further, students' statistical reasoning could be enhanced only if the use of examples, illustrations, analogies, discussions and demonstrations were also relevant to the students' cultures. Secondly, the design of the course, inspired by Dewey's (2011) view of school activity that extends experiences and practices of the adult world, assumed that group discussions, debates, role-play and guided discovery should be part of the social activities on the module.

Classrooms with high social impact become small communities that are able to simulate productive, motivated communities of practice. Such communities are not possible without the use of authentic statistics in which knowledge is constructed in everyday "contexts". Taking a socio-cultural view of learning (Section 4.2.1), Derry and colleagues considered the "context" of learning and the "context" of the statistical problems to be critical for students developing *usable* statistical reasoning. For example, the student activities involved experimentation and generation of hypotheses in small collaborative groups with scaffolding from mentors.

This series of studies with their focus on developing students' statistical reasoning skills do not identify and examine lecturers or mentors' teaching practices and processes in the module. However, this approach to developing statistical modules highlights that teaching is not free of context or lecturers' personal interests and knowledge of statistics, teaching or students. In this view, teaching is dependent on the cultural environment and social circumstances in which the teaching and learning take place. While the study offers valuable insights into intended and attained curricula, other aspects of the implemented curricula are missing, such as evaluating the teaching interventions or which mechanisms were more important to student learning.

2.6.2 Teaching resources

A number of studies have evaluated students' progress and attainment in statistics as a result of using different curricula and teaching resources. Statistics education research focused on students' use of teaching resources such as "context", statistical symbolism, statistical terms with lexical ambiguity, the role of statistical software in learning, and students' difficulties with statistical graphs such as boxplots or histograms (Ben-Zvi and Makar, 2012).

Some studies for example have investigated the role of *context* (situations) and structure in how students solve statistical problems. Beyth-Marom (1982) found that the way data were presented and the task instructions influenced participants' understanding of correlation (the tendency of two events to coexist). The instructions using lay or technical language influenced the interpretation of correlation given by participants. In this view, although statistical rules and ideas are difficult and counter-intuitive, a better understanding of students' perceptions before starting formal learning in statistics has implications for teaching and learning. For example, Beyth-Marom and Fidler (2008) concluded that when learning about correlation, students should be familiar with differences between everyday and technical statistical language. Further, the study recommended that students needed to understand the difference between variables and values and be exposed to symmetric variables first followed by asymmetric variables. Finally, data presented in a table format should precede other formats.

Quilici and Mayer (1996) were interested in the role of examples in helping students to learn to think by analogy, that is 'map a solution method from the source problem to the target problem, sometimes using a solution method abstracted from the source problem'

(p. 145). In particular, they investigated whether the design of examples can help students recognise which situations require which statistical test. In a series of three within-subjects experiments comparing three groups of students before and after experience with examples, the authors investigated whether students were able to sort problems on the basis of structure (i.e. *t*-test, chi-squared, correlation) rather than surface (i.e. context) characteristics. The study showed that when statistics-naive students are presented with grouped examples of a statistics *problem types*, they are more likely to focus on structural rather than surface features in judging problem similarities than equivalent students who are not exposed to examples.

Although this study was conducted with inexperienced students exposed to short-term interventions, it is possible that explicitly teaching students to pay attention to the types of independent and the dependent variables will improve their success with statistical problem-solving. However, statistical thinking can only materialise as part of a context and statistics may be seen as *as a web of interconnected reasons*, some statistical and some contextual (Bakker and Derry, 2011). Therefore, the importance of contextual factors should not be ignored. However, Quilici and Mayer (1996) show how features of examples influence student learning under experimental conditions rather than with (more) heterogeneous groups of students in the classroom context.

In a later study, Quilici and Mayer (2002) examined whether structural awareness can be taught to students on an introductory statistics module. The data showed that statistics students were better at identifying correlation problems than *t*-test or chi-square problems. However, in the case of inexperienced students, this research showed that teaching strategies (e.g. completing schema-training exercises that directed the students' attention to particular structural similarities) can help students pay attention to structural features rather than to surface features. The study concluded that one implication for teaching introductory statistics is to teach students to ignore the context (cover story) in problems and focus on the experimental design and types of independent and dependent variables (one or two groups, quantitative or categorical). Also, given students' success in identifying correlation problems rather than t-test and chi-squared, the study recommended that the teaching should focus on identifying the differences between chi-square and t-test problems. However, this research did not look at the effects of keeping similar surface features (story) and varying the structure, nor at the extent to which either structural or surface variability contribute to students' learning. Further, contextual knowledge is the knowledge about the context in which data are collected and analysed and is considered fundamental in developing statistical literacy, reasoning or thinking (Gal, 2002).

Students find it difficult to recognise an appropriate problem for solving statistical word-problems or recognising that a word problem learnt in one context is relevant to a structurally identical problem in a different context. Difficulties with structural awareness mean that students confuse the t test with χ^2. Quilici and Mayer's (2002) study demonstrates that it is possible to teach structural awareness for simple statistical problems, dependent on schematic knowledge or knowledge of problem-types. However, features of the problem's contextual features are likely to affect students' learning.

Ware and Chastain (1991) explored the effectiveness of a teaching strategy aiming to improve students' skills in using statistical tests (selection skills). The teaching intervention consisted of a handout containing a table which identified concepts of scale of measurement and function, i.e. various conditions of the dependent and independent variables and the particular statistical test suitable for selection in each case. The study concluded that the handout helped students develop selection skills and also that statistics students were better at selecting statistical tests than inexperienced students. The study shows the impact curricular choices (e.g. whether to focus on selection skills or not) have on students' learning. However, less evidence is available of the longer-term effects on students' learning of different teaching resources or teaching methods.

Research was also concerned with different approaches for improving longer term statistical reasoning in introductory courses (Chervany et al., 1977; Sahai, 1990; Garfield and Ahlgren, 1988). Some studies suggested that students seem to learn more effectively when teaching encourages statistical reasoning and interpretation beyond the use of statistical formulas (Bradstreet, 1996). Statistics courses that were tailored to engage students in processes of meaning construction and interpretation rather than focusing on symbols and formulae, detached from an understanding of statistical concepts, were assumed to be better for student learning (Batanero et al., 2000). For example, some studies recommend presenting simpler algebraic forms of equations which can better support students' understanding of statistical concepts. Sahai (1990) in the case of sample variance preferred the definitional formula $\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N}$ where $\overline{x} = \sum_{i=1}^{N} \frac{x_i}{N}$ instead of the computational formula $\sum_{i=1}^{N} \frac{x_i^2 - n\overline{x}^2}{N-1}$. Although Sahai's (1990) recommendation is that students should be able to calculate a finite population variance from a formula (e.g. $\sigma^2 = \Sigma(x_i - \overline{x})^2/N$ where $\overline{x} = \Sigma x_i/N$), teaching time should not be used on algebraic formulae. Instead, teaching resources should include the use of computers and software (see Section 2.6.3) and an emphasis on concepts, principles and applications.

Some studies also looked at whether including formulae in the teaching of statistics helped students understand statistical tests and overcome statistical anxiety (Bradstreet, 1996; Cobb, 1992). Seabrook (2006) conducted a quantitative study with undergraduate students investigating whether conducting calculations by hand such as working through mathematical formulae help students understand statistical ideas. The study used a pre-test/post-test design in which student performance was compared at the beginning and the end of the module and between a calculation task and five multiple choice questions aiming to test statistical thinking. For example, one of the five items testing statistical thinking tapped into students' selection skills (e.g. 'if your study has two variables measured at an interval level and you wanted to investigate the relationship between the two, what test would you use?'). The teaching of calculations is characterised as a calculation task included in the final examination which supplied students with a formula for Person's correlation r and probability tables asked students to calculate the correlation between two variables with twelve data points each. This study appears to analyse limited data from student existing assessments on an application of a formula to a dataset and only five multiple choice items as a measure of statistical thinking rather than looking at whether the teaching of statistics using formulae helped students' statistical thinking. Despite its limited design, the study concluded that competence at calculations accounted for approximately two percent of variance in statistical thinking.

Beins (1993) was interested in how to encourage students to learn 'interpretation' skills. To do this, Beins (1993) evaluated the use of assignments that required students to carry out data analysis to answer an empirical question and formulate an interpretative account without using statistical terminology. Similar to Ware and Chastain (1991), this quasi-experimental study compared four successive groups, each differing in the work performed. The first group was required to complete the same module plus the writing assignment for which they received extra credit. The second group followed the same module but instead of completing the writing assignment, the group practised generating conclusions during class meetings. The final group were required to produce conclusions independently on problems given in tests and homework exercises. Based on the final exam marks, the first group, which was exposed to writing and interpretation skills, showed highest scores on computational and interpretative skills but not conceptual scores. However, the study does not provide any information regarding possible differences between groups in prior-attainment and the students could not be randomly assigned to treatment groups. As well as the focus on interpretation skills in the first experimental group, it is possible that the effect in learning gain observed was also due to the amount of time spent completing the assignment (homework) rather than the value of the assignment.

2.6.3 Technological tools

Technological (computer) tools were also used in statistics education to address changes in content, pedagogy and teaching format (Chance et al., 2007; Bakker, 2004; Tintle et al., 2011; GAISE, 2016). For example, teaching interventions using technology have been developed to help students overcome some of their difficulties with understanding particular statistical concepts (Garfield et al., 2012b). Technological tools in the classroom are also believed to provide interactive teaching and learning opportunities through multiple representations and simulations (Garfield et al., 2015).

DelMas et al. (1999) conducted classroom research with students enrolled in an introductory statistics course. The aim of their action research was to evaluate an interactive, computer-based simulation software and hands-on activities used to help students visualise and understand sampling distributions, crucial in understanding statistical inference. The problem DelMas et al. (1999) identified was that their students, although able to pass the module, could not apply statistical concepts in their reasoning or in different contexts. For this reason, the study investigated the effectiveness (i.e. whether students achieve higher scores on post-test versus pre-test) using a new instructional approach based on the model of conceptual change that allows students to make predictions and directly test them. When compared to pre-test score, students' post-test scores using the new activity were significantly higher than those who used the initial activity. The study concluded that students' learning was enhanced when they became aware of their misconceptions, e.g. about chance and confronted them using empirical data.

Using the algebra-based introductory statistics course, in a series of studies, DelMas et al. (2007), Zieffler et al. (2010) and Tintle et al. (2011) used a 40-item multiple-choice test (the Comprehensive Assessment of Outcomes in Statistics, CAOS) to measure students' statistical literacy and reasoning of concepts such as distribution, centre and variability rather than applying a formal statistical calculation. Used on a comprehensive sample of American students, the limited gains on the CAOS test before and after undergraduate statistical modules lead some to suggest that the cause of this apparent lack of student progress in statistics is due to the curriculum (DelMas et al., 2007; Garfield et al., 2012b). As a result, Garfield et al. (2012b) called for a radically different curriculum for the introductory statistics course. The authors proposed a simulation-based approach to inference that requires students to create a model with respect to a specific context, repeatedly simulate data from the model and then draw a statistical inference. The focus of this approach is to start from informal statistical inference and then move students to formal simulation-based methods of statistical inference through the use of interesting contexts using real statistical enquiry and real-world data. The core principle is to enable students to *start to think statistically*, i.e. be able to recognise a suitable statistical model and use it when making an inference with context. This work has helped define what it is meant by statistical reasoning within the classroom context.

Ragasa (2008) compared a computer-assisted teaching approach where students often work cooperatively in groups in a computer laboratory with the 'traditional' teaching method consisting of lectures given by the lecturer, recitation and tutorial activities involving the topics discussed during the class. Although the experimental and control (traditional method) groups were not randomly assigned, thus limiting the generalisability of findings, the study was able to conclude that collaborative computer-assisted methods had a significant effect on learning but not on attitude. However, despite recommendations to include computer simulation methods to teach statistics, in a summary of the literature, Mills (2002) emphasised that there is little empirical evidence to support such recommendations. For instance, Mills (2002) questioned whether using computer simulations help students' understanding of t-distribution and related concepts (Gordon and Gordon, 1989). Computer simulations might help students for example 'see' the differences between sample means from sample to sample and illustrate effectively statistical concepts such as sample estimates, random variables, and estimated standard errors. However, robust empirical data is required to demonstrate the benefits of using computer simulations. Importantly, the study concludes that the students in simulation-based modules seemed to have developed a better understanding of basic statistical ideas than other students enrolled on algebra-based statistics modules. However, the students could not be randomly assigned to the two types of modules (non-simulation, algebra-based or simulation-based).

The reseach concerned with the impact of teaching methods (lecturing versus other methods), teaching resources (design features of problems and examples, computations versus conceptual understanding within a real-world context using technology, assignments or projects encouraging discussion and reflection versus problem exercises) and experiencing the whole investigative cycle on student learning highlight the critical importance of tools in learning statistics.

2.6.4 Students' difficulties with learning statistics

Some studies investigated how statistics can be taught at university to aid students overcome difficulties with learning statistical concepts and with non-cognitive factors such as motivation and disciplinary interests (Urquhart, 1971; Kalton, 1973; Jolliffe, 1976). Rather than investigating students' abilities in statistics, this research strand focuses on documenting students' difficulties with and misconceptions about key statistical concepts, such as understanding of the mean, independence, samples, sampling variability, sampling distributions, the Central Limit Theorem (clt), graphical representations, data collection and design, significance tests and statistical models such as correlation or regression (e.g. Swanson et al., 2014; Sahai, 1990; Mathews and Clark, 2007). In particular, research in the 1970s focused on informal (naive) statistical reasoning and after the 1980s on formal (following teaching) statistical reasoning and the links between naive and instructed reasoning.

In addition, research studies appear to agree that lecturers and their students believe that statistics is a difficult subject for students and consider that a number of changes are necessary in the learning environment in order to improve statistics education at tertiary level in a range of areas, including in content, teaching methods, teacher preparation and training, assessments, or the use of technology (Garfield and Ahlgren, 1988; Cobb, 1992; Gal and Garfield, 1997; Lajoie, 1998; Ben-Zvi, 2000; Watson, 2001). Some research in mathematics education sees the learning of statistics as a process through which students construct mathematical concepts and create knowledge through experience (Resnick, 1987a; von Glasersfeld, 1998). Statistics education research has investigated naive understanding of statistical concepts in the absence of instruction (Tversky and Kahneman, 1983) and on student informal reasoning as a precursor to formal inferential reasoning (Ben-Zvi et al., 2012). In general, the research into student learning of statistics at university has focused on the impact of specific teaching resources, students' misconceptions, difficulties with statistics and learning of particular statistical concepts and intuitive statistical reasoning.

2.6.4.1 Misconceptions about statistical concepts

Research into students' difficulties with statistical concepts is relevant for teachers of statistics to gauge students' starting point and the amount of conceptual learning that goes on in statistical modules. As Garfield noted

> "Ideas of probability and statistics are very difficult for students to learn, and

> conflict with many of their own beliefs and intuitions about chance and data. Students connect new ideas to what they already believe, and correct or abandon erroneous beliefs reluctantly, only when their old ideas don't work or are inefficient" (quoted in Cobb, 1992, p. 10).

Work on informal statistical reasoning was also considered useful in improving learning by having students "become aware of and confront their misconceptions".

A number of studies highlighted the problem with the students' heterogeneous backgrounds with regards to previous preparation and prior learning in mathematics and statistics. Jolliffe (1976) noted that students who did not take mathematics after the age of sixteen achieved lower results than the other students. To overcome issues with student preparation and anxiety with statistics, Jolliffe proposed a system of continuous assessment consisting of quantitative essays and assignments (work-books). However, an evaluation of this particular framework was not carried out.

Other studies showed that students continued to find it difficult to apply particular statistical concepts even after formal learning in statistics and investigated the study strategies of students in statistics at university. Derry et al. (1995, 2000) describe an innovative statistics course for undergraduate education focusing on simulated or realistic problems that aim to improve the statistical thinking skills of these students. The students who followed this innovative module appear to be more likely to recognise the limitations of correlation evidence on a study (an increase of about 1.26 standard deviation units from pre-test to post-test), understand the components of scientific experimentation ($t(13) = 2.17, p = 0.025$, *effect size* $= 0.85$) and random sampling ($t(13) = 6.35$, $p < 0.001$, *effect size* $= 1.32$). Further, these participants also improved their general reasoning capabilities.

Certain faulty reasoning still seemed to persist, specifically the distinction between random sampling and random assignment, estimation based on random sampling from a known population and students' critical evaluation. However, such findings from one-group pretest-intervention-posttest design, common in statistics education (e.g. Park et al., 2011), although found improvements from pretest to posttest in the quality of students' statistical reasoning (effect sizes of one pretest standard deviation), use opportunistic samples and therefore lack a randomised no-intervention control group.

Instead, Shaughnessy (1997) recommended that statistics education research should carry out longitudinal studies into the growth and change in students' thinking in variability, chance and data, starting with what they can do rather than what they are unable to do. For example, Shaughnessy recommended the use of datasets in contexts which would allow students to experience the concept of variability (spread) in several forms in the data, detect patterns in the data and make predictions. Rather than asking students to answer forced-answer questions, Shaughnessy recommended the use of more open ended questions that require students to 'to try and make sense out of data, to generate their own questions about data, to design their own graphical representations of data, to in fact even create their own *measures* on data' (p. 14). In essence, after a period from the 1980s/1990s when research used forced-answer responses to probe student misconceptions about statistical concepts (Shaughnessy, 1992), towards 2000s, Shaughnessy called for a change in methodology towards think aloud open ended tasks to probe how students think statistically.

Fong and Nisbett (1991) and Fong et al. (1986) looked at how students with formal training in probability and statistics compared to those without formal training. They showed that formal statistical training using examples enhanced students' use of statistical principles in reasoning. Further, statistics training (e.g. in the law of large numbers) using examples could be generalised across method, context, type of subject and domain. However, Fong et al. (1986) found that the statistics module failed to enhance students' statistical reasoning for problems set outside the classroom context (e.g. playing squash, effect of marriage on athletes' performance). In this study, the participants had problems transferring relevant knowledge and skills outside the statistics classroom. However, students' statistical reasoning was tested in a phone interview set outside the classroom context and with the declared aim of testing students' opinions about sporting activities. Also, Fong et al. (1986) tested the hypothesis that adults tend to use abstract rules in their application of statistical concepts rather than understand students' reasoning with statistics.

Batanero et al. (2004) further investigated introductory statistics students' understanding of the normal distribution as evidenced in three assessment tasks (e.g. find which variable fits a normal distribution well). The study found that many students correctly understood abstract properties, such as the usefulness of models, density curves, and areas under the normal curve and related concepts and properties. However, based on students' responses and interviews, the study showed that there was a discrepancy between the intended learning outcomes and the meaning of normal distribution acquired by the students. For example, some students confused the empirical data distribution with the theoretical distribution (e.g. for age) and also erroneously thought that a discrete variable with only three different values was normal. Such studies investigating students' ways of thinking and difficulties with particular concepts following teaching may help identify sequences of content and suitable resources when learning complex concepts such as the normal distribution.

Other research investigated why students reason incorrectly about particular statistical concepts (e.g. random sampling) and the impact erroneous understanding of statistical concepts has on students' statistical thinking. The focus in these studies is on how and why university students understand particular statistical concepts such as the mean, chance, uncertainty, independence, significance, correlation, random sampling and variability.

Pollatsek, Konold and colleagues carried out a series of studies into students' beliefs about the mean, random sampling, sample size, ideas about data and data analysis. For example, Pollatsek et al. (1981) interviewed university students as they solved problems involving the appropriate weighting and combining of means into an overall mean such as the grade point average problem in which grades have different weights in calculating an overall mean (the mean is considered one of the most basic concepts in statistics). The study found that although the participants could calculate the mean, a large proportion did not understand the concept of the weighted mean indicating that many students understand the concept of the mean computationally rather than conceptually.

In another study, Konold et al. (1993b) carried out think aloud interviews with twelve undergraduate psychology students to test the degree of understanding of the law of large numbers before instruction. This study suggested that students' difficulties with learning the law of large numbers were not based on incompatible intuitions but due to difficulties with basic mathematical concepts such as percentages and statistical concepts such as distributions of means and of samples.

In a study extending the work of Tversky and Kahneman (2009), Pollatsek et al. (1984) investigated whether participants hold a passive, descriptive view of random sampling or an active-balancing model (the idea that things 'will even out', e.g. when flipping a coin) in which earlier trials influence later trials. Based on think-aloud interviews, Pollatsek et al. (1984) suggested that the participants held the belief that the population mean was the best guess for both sample means. This was in spite of presenting alternative solutions or participants showing good comprehension of the rationales underlying those solutions. The authors conclude that students' *heuristic representativeness* is different from the commonly held belief that students' misconception is of active balancing (e.g. expecting that a tail will come up after a series of heads thus revoking the notion of independence). With this research, Pollatsek et al. (1984) showed that students might hold different conceptions of statistics. Although the research did not focus specifically in improving students' learning and teaching and studied students' statistical reasoning outside the statistics classroom, the study participants were university students who were exposed to more realistic problems.

2.6.4.2 Non-cognitive factors

Beyond cognitive statistical skills, an important body of research has investigated the effects on student learning of non-cognitive skills, such as the students' attitudes towards and motivations for studying the subject, a well-documented challenge in undergraduate statistics. For example, Gal and Ginsburg (1994) documented that students may find statistics difficult due to non-cognitive factors such as 'feelings, attitudes, beliefs, interests, expectations, and motivations'. The authors stated that in addition to a statistics module facilitating statistical thinking, students should also 'emerge from statistics classes without apprehension or negative feelings about learning more statistics'.

Motivational research is relevant in statistics education which emphasise the use of authentic contexts. The assumptions are that authentic data/problems might increase students' motivation and lead to improved student competence. Research has also investigated the effect non-cognitive factors such as negative attitudes or beliefs towards statistics have on students' achievement in statistics (Gal and Ginsburg, 1994; Budé et al., 2007).

Budé et al. (2007) examined the motivational constructs and their effect on students' achievement on a statistics course. Taking a quantitative approach, the authors investigated relationships between the module grades and students' activities within this statistics module, affect and study behaviours. The study concluded that the resources used in a module, such as tasks or exercises, were important in stimulating students to study the material. Second, the study showed that students who liked statistics and were interested in the module, saw the value and relevance of statistics and were resilient obtained a higher exam grade. Based on these results, the authors conclude that statistics modules should be interesting, challenging and enjoyable to students.

However, the study lacks specific recommendations regarding ways to achieve these curricular aims. Further research would therefore be required to identify teaching and learning strategies which ensure positive outcomes for the students. Such studies could for instance be carried out from either or both the lecturer and students' perspectives. It might be important to understand how lecturers ensure a positive environment in the classroom. On the other hand, further research might investigate how students perceive the resources provided in lectures and tutorials.

This research has focused mainly on student-related characteristics such as attitudes towards statistics, enjoyment, interest and nature of tasks presented. Students' perspectives are relevant to classroom teaching. Another area of research in my view needs to focus on the teacher and investigate the attitudes with which teachers (lecturers) approach statistics (integrative motivation), view the student (instrumental need for achievement and self-confidence) or the "context" or situation within the module (group-specific motivational components). For example, studies could look at the teaching situation, how to ameliorate motivation and diminish de-motivation through teacher teaching, materials and teaching methodology. There is very little research in statistics education looking at student anxiety and pre-knowledge or at how to design motivational classroom interventions. In my main study (Chapters 4 and 5), I became particularly attentive to non-cognitive factors from the lecturers' perspective.

2.7 Implications for teaching statistics and research

The topics of these research studies reflect the emphasis in statistics education on developing conceptual understanding and statistical literacy, reasoning, and thinking rather than procedural understanding, computations, formulas, techniques and procedures (Garfield and Ben-Zvi, 2004). One implication for teaching is to specify the learning outcomes in statistics courses in terms of module specifications but also the collection of teaching, learning and assessment resources. For instance, if the module aims to develop students' statistical thinking, the obvious solution is to include assessments on higher levels of statistical thinking and reasoning skills.

This review has also highlighted students' challenges with learning statistics due to both cognitive and non-cognitive skills. In designing curricula, it seems that the module content and the sequencing of this content are important, although lecturers' own learning goals for the students and beliefs about the students and about statistics are likely to impact on what students achieve by the end of their learning.

While these conclusions regarding student learning are fairly obvious, the research is less clear regarding the effectiveness of teaching methods (e.g. active learning versus traditional lectures). The evidence supporting particular curricula and teaching methods is inconclusive since in some cases the evidence consisted of quantitative measures of unknown reliability and validity.

Another challenge for teaching statistics is to integrate everyday contextual knowledge with statistical investigations and connect these experiences to teaching practices. One such issue is the balancing of theoretical statistical principles with situation-specific forms of competence and bridge the gap between university and workplace. Studies investigating the connection between workplace and school statistics or between novice and expert practices have offered some understanding regarding the nature of statistics and of learning statistics in different environments.

However, this research has predominantly focused on students' activity and less is known about the activity of statistics lecturers and the role of different types of tasks and other resources across contexts. There appears to be an understanding and consensus of what to teach but less is known about how lecturers teach in real situations. The large majority of the research surveyed helped with defining concepts used in the analysis (Chapter 5) and provided background to the lecturing activity I observed in my study. In addition, there was limited literature directly relevant to the methodological approach and the aims of my study, i.e. to characterise the teaching of statistics at university from a sociocultural perspective. My study aims to focus specifically on lecturers' activity before, during and after lecturing.

Although some research directions are promising, further research in characterising the teaching of statistics at university could help to further explain what it means to teach statistics at university.

2.8 Summary of chapter

Since the data revolution of the 1960s with the advent of new graphical representations (stem and leaf, boxplots) and computers (for data analysis and/or for learning) statistics education research has witnessed an explosion of research into new pedagogies. In general, the statistics education literature relied on both professional studies, i.e. educators' reflections on teaching practices and research studies, i.e. systematic enquiries based on empirical data. The literature is concerned with several aspects of teaching and learning statistics across educational levels, including the development of statistical thinking, reasoning and literacy in adults and learning at primary, secondary and university levels.

At university level, research focuses on curriculum content, including specifying the nature of the statistics students need to learn, the use of teaching resources and the impact of different teaching methods on student learning. Statistics education has focused on practical problems such as teaching interventions using innovative computerised tasks, and active or collaborative learning in large classes.

A number of studies investigate students' difficulties with learning and using statistics. In this category, studies identify several challenges with teaching and learning statistics at university, including high student numbers, writing and communicating statistical findings, students' affective and motivational issues with learning statistics, students' prior background in mathematics, students' understanding of the non-statistical *context* in statistical problems and data interpretation.

In general, there appeared to be some agreement regarding what constitutes statistics and how it should be taught. Further, the agreement about what constitutes statistics has resulted in agreement in terms of the content and goals for teaching (e.g. statistical reasoning). There was however limited evidence of 'what works' in practice and there seems to be a disconnect between curricular aims and content and the specification of learning outcomes for students. Further, there was limited evidence in the literature surveyed that particular teaching methods are better than others in improving students' learning outcomes. However, the content and teaching methods differ across contexts due to differences in culture, pedagogy, curriculum (learning outcomes) and availability of resources. Importantly, I found that there was limited literature that focused on the practice of teaching statistics at university with qualitative analyses or from a sociocultural perspective.

3 | Pilot Study

Chapter contents

In this section I describe the beginning stages of my research journey and the first phase of my research project which I called the pilot study. This involved interviews with twenty statistical methods lecturers and observations of teaching on one statistics module, which I present below.

In the first stage of my study (Table 1.1 in Chapter 1), I aimed to survey the the kind of skills, knowledge and understanding statistical methods lecturers perceived as important for their students. Based on the literature review (Chapter 2), I considered it was important to gain an understanding of relationships between lecturers' beliefs and curricula (Section 2.4) within the context of teaching and learning statistics in universities in the UK. This stage in my research process was useful as I could identify a specific challenge in the teaching of statistics at university which defined the purpose of my pilot studies, i.e. to investigate statistics lecturers' beliefs about intended curricula.

Literature review 1 *(Chapter 2)*
↕
Pilot study 1 *(interviews, 9 lecturers)*
↕
Pilot study 2 *(interviews, 11 lecturers)*
↕
Pilot study 3 *(observations of teaching)*
↕
Main study methodology *(Chapter 4)*
↕
Main study data analysis *(Chapters 5)*

Figure 3.1: The structure of the Pilot study

For this purpose, I carried out three pilot studies (Figure 3.1) in the initial stages of my study, as follows:

- **Pilot 1** involved repertory grid interviews with nine lectures (Section 3.2.1).
- **Pilot 2** used the data analysis in Pilot 1 to design repertory grid interviews with eleven lecturers to gain a deeper understanding of lecturers' beliefs about statistics curricula and validate some of the finding.
- **Pilot 3** I observed teaching of two lectures and two tutorials for one module (Section 3.3).

The pilot studies were emergent as, rather than starting from a hypothesis, they aimed to build an understanding of the teachers' beliefs regarding the knowledge, skills and content intended on statistical modules, grounded in the data collected from interviews with statistical methods lecturers (Glaser and Strauss, 1967). Hence, in this chapter, I discuss what influenced my methodological choices, the research design and findings. The pilot studies influenced the development of my research questions and further refinement of data analysis in my main study (Chapters 4 and 5).

3.1 Identifying a specific challenge

As shown in Table 2.1, p. 13, over the past three decades, reforms in statistics teaching and learning have dominated the statistics education research, professional and pedagogical literature. At tertiary level, research since the early 1990s has made recommendations for changes in the intended statistics curricula, more data analysis and less probability and pedagogy, fewer lectures, more active learning and the use of technology (Cobb, 1992; Moore, 1997). There is broad agreement that the key learning goals in introductory statistics are statistical literacy, reasoning and thinking (Bradstreet, 1996; Cobb and Moore, 1997; Gal, 2002; Broers, 2006, Chapter 2). However, the literature also shows that statistics has remained a challenging subject for lecturers to teach and for students to learn (Kalton, 1973; Bryce, 2002; Garfield et al., 2012b; Tishkovskaya and Lancaster, 2012).

Further, at the tertiary level, statistics courses for non-statisticians, i.e. students who take a compulsory introductory statistics module, face particular challenges given the profile of some of the lecturers who may not be statisticians and students' cognitive and non-cognitive skills and knowledge (Gal and Ginsburg, 1994; Gordon, 2004; Allen et al., 2010). Although findings from mathematics education suggest that teachers' pedagogical content knowledge impacts on the quality of teaching, in statistics education teachers' pedagogical content knowledge is less researched and hence well understood, particularly at university level (Callingham and Watson, 2011).

As discussed in Section 2.5.1, teachers' beliefs and conceptions about important aspects of the curriculum that include intended (or planned) curriculum comprising aims, goals and objectives, content, teaching strategies and assessments have a key role in how well students attain the curriculum. For example, Pajares (1992) suggests that it is important to understand the relationship between teacher beliefs, teacher knowledge, teaching practice and student outcomes in order to characterise teaching (see p. 37 above). In this study, I focus on lecturers' beliefs about intended curricula and pedagogy, which are assumed to be *tacit knowledge* (p. 38), in order to draw attention to the key role of the lecturer in learning statistics at university (Handal and Herrington, 2003) and to make lecturers' tacit knowledge explicit (Greatorex, 2002). In my interpretation, lecturers tacit knowledge of curricula might be inferred as personal constructs or schemas. A suitable methodology for my pilot study would need to be able to make such tacit knowledge explicit.

3.1.1 Pilot study research questions

This chapter therefore presents an empirical study into lecturers' beliefs and conceptions about the undergraduate statistics curriculum. In this pilot study, I set out to explore lecturers' planning of undergraduate statistics education. **The general research question** which I investigate in this thesis is:

> *What insights can I gain about the teaching of statistics at university by studying the teaching activity at two levels of context, the macro and micro?*

As explained in Chapter 1, this question is important because the practices and processes of teaching a subject matter such as statistics in an integrated account of the macro and the micro are relevant to the quality of education experienced by students. With my focus on lecturers' intended and implemented curricula in my thesis, my **pilot study research question** is as follows:

> *What characterises the participants' beliefs regarding the teaching of statistics at university?*

This question is divided into five specific research questions shown below.

1. What characterises the participants' beliefs about the learning outcomes important for students to attain on statistical modules?
2. What characterises the participants' beliefs about the content of statistical modules?
3. How do lecturers' beliefs about statistics curricula compare to actual module specifications?
4. What characterises the participants' beliefs about the teaching and learning of the statistical investigative process components?
5. What characterises the participants' beliefs about the key statistical skills students should learn on statistical modules?

3.1.2 Methods for researching teacher beliefs about intended curricula

In tertiary statistical education, the methods used for researching beliefs about intended curricula have included interviews, open questions, focus groups (Gal and Ginsburg, 1994; Davies et al., 2012; Price and Rust, 1999) and questionnaires (Gardner and Hudson, 1999) with lecturers, employers and students, standards and guidelines produced by standard setting bodies (Franklin et al., 2007) or opinions from subject experts (Gould, 2010; Nicholl, 2001; Hawkins, 1997; Higgins, 1999; Raymond and Neustel, 2006). For example, in order to investigate the use of different assessment methods, Garfield (1996) used a 5-point scale to measure students' beliefs about statistics and its value. However, the study did not survey the teachers' beliefs about statistics, its value or curricula.

As a first step in the development of a test on how well students presented with a research problem are able to select a suitable statistical procedure for analysis, Gardner and Hudson (1999) conducted a survey with lecturers in order to check the validity of the topics included in the test. To do this, the authors compile a list of 34 topics based on an analysis of introductory statistics texts, journal articles and their own subject matter expertise. They then ask statistics lectures to rate the importance of the statistical topics in a statistics course on a 5-point scale. This ranking of statistics topics was then considered in test construction; the least important topics were disregarded from the test, while more items were included against the most important topics. This method of validating the test content however focused exclusively on domain-based content and not on the context of teaching and learning or learning outcomes.

Another approach is to classify learning outcomes into categories, such as the process components of the *statistical investigative process* - formulate research questions, collect data, analyse data, and interpret results (Franklin et al., 2007) or whether there is a requirement for statistical reasoning (see Newton et al., 2011, in Section 2.4.3, p. 32). Studies that focused on developing content for education in a particular domain have also been based on frameworks or models that describe the skills, knowledge and understanding in that domain. For example, in an analysis of video recorded lessons and interviews with primary teachers, Burgess (2008) developed a model of knowledge needed by statistics teachers that included statistical reasoning based on a theory and model of statistical thinking by Wild and Pfannkuch (1999, discussed in Section 2.4.1.3).

Classroom teaching experiments and artefacts of student work (assignments, classroom tests and observations) in conjunction with cognitive models of children's mathematical learning were used within a constructivist view of learning (Inhelder and Piaget, 1958) to validate a particular curriculum, i.e. determine what statistics ought to be worth knowing and doing (Cobb et al., 1991; Steffe and Thompson, 2000). For my pilot study however, I was interested in capturing lecturers' tacit knowledge of intended curricula, what curricular learning outcomes and content lecturers believed to be important for students to learn and how to teach it, rather than instances of attained curricula (e.g. student assignments).

In conjunction with other sources of information, the repertory grid technique (RGT, explained in Section 3.2), initially developed in clinical psychology for providing therapy to individuals (Kelly, 1955), has been adapted to capture tacit knowledge in a domain such as beliefs about self in relation to others for a range of applications (Kelly, 1955; Fransella et al., 2004). In educational research, RGT has been used to articulate qualitative information such as assessment criteria, grade descriptors, marking criteria, curriculum statements or competencies (Bjorklund, 2008; Greatorex, 2002; Honey, 1992; Rayment, 2000; Senior and Swailes, 2004; Tofan et al., 2011).

Bjorklund (2008) applied RGT to elicit the (tacit) criteria two secondary level teachers use in assessing student technology projects. The study proposed that assessment criteria represent knowledge assumed to be subjective and outside teachers' awareness. Bjorklund considered that ordinary interview techniques are not useful since such knowledge of assessment criteria teachers use is not verbal and the interviewees might not be aware of their decisions and actions. The teachers were asked to select 7-8 artefacts or different projects from their own class. I explain terminology used in repertory grid research below in Section 3.2, including grid, elements (columns) and constructs (rows). This study used the triadic method which compared three elements (artefacts) for the generation of nine different constructs (differences between elements) on a scale of 1 to 9. In this way, the study identified constructs that refer to both student dispositions and behaviours in the process of creating the artefact as well as with the product itself. Given the focus in statistics education on cognitive (Section 2.4.1) and non-cognitive skills (Section 2.6.4.2) when defining curricula, I consider that RGT could be valuable in eliciting lecturers' beliefs about statistics curricula.

Kandola and Pearn (1992) proposed that the repertory grid technique can be used to develop competency statements for human resources purposes. By comparing and contrasting triads of types of managers (i.e. good, poor or average at handling people), participants were able to produce statements of how two managers were more alike and how one was different from the third. These contrasting statements, e.g. 'prepare statements with care and adequate research' versus 'no attempt to assemble relevant facts', were then used as scales on which each of the types of workers could be rated on a 7-point scale. The authors showed how this approach could be used to 'surface' participants' perceptions and arrive at a list of behavioural and competence indicators.

Hoogveld et al. (2002) also used RGT to investigate the way trainee teachers approach the design of a new instructional unit. One of the measures used in this study required participants to provide instances of instructional design that they compared and described. These descriptions were then analysed by comparing them to a prototypical model of instructional design in order to explore teachers' instructional design practices. It might be possible therefore to compare lecturers' intended curricula with learning outcomes from actual module specifications used in higher education institutions.

In another study, Greatorex (2002) used an adapted RGT to formulate grade descriptors at different grade levels for Accounting examinations. Senior examiners in this study compared student answers to questions who achieved different grades (the elements). The next step was for examiners to identify *qualitative differences* between Accounting A-level scripts at different grade levels which would form the basis of the grade descriptors. In this study, accounting experts compared triads of borderline scripts, e.g. two scripts near the bottom of grade B and one near the top of grade C, and described how two scripts were the same yet different from the third and the context in which these statements were exhibited. In this way, subject matter experts were able to evidence their views about which qualities are associated with each grade based on their own personal and perhaps tacit knowledge. In this way, the examiners' tacit knowledge which they use to make judgements about student performance was made explicit as grade descriptors. By asking the participants to compare their own objects of interest such as skills and knowledge using RGT, I expect to gain a valuable insight into the participants' tacit knowledge and beliefs of the curricular, didactic and pedagogic processes on their modules, which they may otherwise not verbalise using other methods, e.g. a questionnaire, structured or in-depth interview (Gentner and Medina, 1998).

In studies comparing qualifications, e.g. from different Awarding Bodies in the UK, bipolar constructs resulting from RGT were also used to create a questionnaire of the relative demands placed on candidates (Elliott and Greatorex, 2002). The questionnaire is then used by a larger sample of examiners to rate different assessments (defined as elements) against the same constructs that form the questionnaire. Using a different implementation of RGT, Rayment (2000) investigated art teachers' perception and understandings of the national curriculum framework by providing eight bipolar constructs and six National Curriculum statements (the elements). When selecting RGT as the methodology for my study, I considered that it was possible to formulate a list of statistical skills and knowledge or learning objectives based on the data generated by several participants. The analysis of this list would then give an indication of the nature of the learning objectives several lecturers believed to be considered relevant for introductory statistics.

Since my focus in this pilot study was on beliefs and conceptions about lecturers' intended statistics curricula, I considered the repertory grid technique (RGT) to be a suitable methodology for identifying lecturers' subjective beliefs within the time available for interviews (one hour). Further, I considered that qualitative analyses using classification methods (e.g. Franklin et al., 2007; Newton et al., 2011; Hussey and Smith, 2002, or comparisons with intended curricula from existing module specifications) would shed further light into lecturers' planning of their teaching. In this pilot study, I was interested in lecturers' beliefs about what and how introductory statistics should be taught and what it means to do, understand and teach statistics rather than knowledge of the content of introductory statistics or knowledge for teaching.

3.1.3 Ethical issues

Throughout the data collection and analysis process, I was mindful of the ethical implications of researching and recording lecurers' activity and views about teaching statistics. The study followed the British Educational Research Association guidelines for educational research (The British Educational Research Association [BERA], 2011) as well as the University ethical guidelines. At each step, e.g. observations of teaching, interviews, collection of documentation, I sought and received the explicit consent from participants to extract the data (Wagner, 1997). I also informed participants of the purposes and implications of this study and provided a broad description of the data analysis procedures. I also considered how my own reflective research might affect others, and the impact the research might have on lecturers and the statistics education community. I used pseudonyms and generic descriptions of the modules to maintain the anonymity of the participants (Miles et al., 2013).

Corbin and Strauss (2008) define research sensitivity as 'having insight, being tuned in to, being able to pick up on relevant issues, events, and happenings in data' (p.32). Sensitivity is the interplay between researcher and data. In my Pilot study using interviews with lecturers, I aimed to pick up on subtle verbal responses and non-verbal cues during the interview. The interviews I conducted in a sensitive way and guided by ethics to ensure that the data collection process was beneficial to both participating lecturers and the researcher (Corbin and Morse, 2016). RGT can provide rich insights into participants' belief systems. I was mindful of my participants' comfort during interviews to ensure the procedure was not tiring and that the pace of the interview was sustained.

Protecting my participants' anonymity, confidentiality and well-being was particularly important in order to be able to convey the meanings of what the participants were saying about their experiences in statistics education. Further, I aimed to present the views of participants with respect to what they wanted their students to achive by the end of their learning by trying to put myself in their role through immersion in the data. During data analysis, I looked out for subtle nuances of meanings in the data provided by the research participants by applying multiple analytic strategies.

3.2 Pilots 1 and 2: the repertory grid technique (RGT)

In order to start answering the research question and identify the statistical skills, knowledge and understanding students ought to learn, I set up an exploratory study with statistical methods lectures using the repertory grid technique. This pilot study was an emergent study which, rather than starting from a hypothesis, it aimed to build an understanding of the intended curricula (learning outcomes and content) of statistical modules, grounded in the data collected from interviews with statistical methods lecturers (Glaser and Strauss, 1967). My intention was therefore to conduct an exploratory study to investigate its usability for the purpose of identifying curricular intentions.

RGT, initially developed in clinical psychology for providing therapy to individuals, had been adapted to capture *tacit* knowledge in a domain for a range of applications (Kelly, 1955; Fransella et al., 2004). As a constructivist theory, RGT asserts that people experience and describe their environment in terms of bipolar personal constructs (tacit knowledge) that they use to construe a meaningful world (Grice et al., 2004).

RGT is considered to be highly personal and subjective since the way individuals make sense of their experiences often differs (Fransella et al., 2004). The technique involves interviews, face-to-face, online or computer-based. The interviewee is presented with a grid comprised of **elements** as column headings, **bipolar constructs** as rows and ratings linking an element to a construct (explained in further detail below). Elements are the object of study, cases, examples, ideas within a specific domain, e.g. statistical skills and knowledge. Elements are judged according to constructs. Kelly (1955) proposed that every person experiences or perceives the world in an individual way. These perceptions are described as constructs. The interviewee's systematic comparisons or interpretations of the **elements** in terms of the research question result in bipolar constructs (e.g. easy-difficult), the ways the interviewee makes sense of or construes the elements. A grid can capture a person's perceptions, associated feelings and intuitions held by the interviewee about the research question and so I considered RGT to be suitable in making lecturers' tacit knowledge *explicit* (as in Section 3.1.2).

RGT has been applied in various ways since Kelly's first conception of personal constructs theory. For example, one way to elicit personal constructs from elements is the triadic version which asks the interviewee to compare three elements. The interviewee then decides which two elements are *similar* yet *different* from the third in terms of the research question. A comparison between elements generates a **construct** which becomes a row of the grid with two poles: the positive construct (how two elements are similar to each other) and the negative construct (yet different from the third). A construct is therefore not the same as a concept since it has two poles, a similarity (positive pole) and a difference (negative pole). A construct with one pole such as ‘good with numbers’ (positive) must be described with its contrasting pole since the opposite of ‘good’ could be ‘bad’ but also ‘weak’, so eliciting a bipolarity is essential in a RG interview (Kelly, 1955). The next step is to link (all) the elements to the construct on a **scale**, e.g. 1 to 7, at one end the positive, similarity construct (=1) and at the other end the difference, negative construct (=7).

3.2.1 Interview designs

The interviews were designed to last one hour. I considered that it was important to ensure lecturers’ participation in the study and comfort during interviews. In my study, I used two interview designs, summarised below as follows:

- **Pilot 1, Design 1**

  1. Interviewee and researcher agree on the topic and research question.
  2. Interviewee generates 9 elements representing *the statistical skills and knowledge the lecturer would like his/her students to learn by the end of the module* and writes them on cards.
  3. Researcher selects **three** elements written on cards at random.
  4. Interviewee compares the three elements to elicit a bipolar construct.
  5. Interviewee links elicited constructs to elements by rating all elements on a scale 1 to 7, at one end the positive construct and at the other end the negative construct.
  6. Researcher fills in the grid (elements, constructs and ratings).
  7. Repeat steps 3 to 6 until the interviewee cannot produce any more constructs or the combinations of elements are exhausted.

- **Pilot 2, Design 2**

  1. Interviewee and researcher agree on the topic and research question of the study.
  2. The interviewee receives 24 elements written on cards and discusses them with the researcher.
  3. The interviewee has the option to add or take away elements.
  4. Researcher selects **three** cards at random.
  5. Interviewee compares the three cards against a supplied construct.
  6. Interviewee links all elements to the construct provided on a scale 1 to 7.
  7. Researcher fills in the grid (ratings).
  8. Repeat steps 4 to 7 for the provided constructs.
  9. Interviewee has the option to add constructs (following Steps 4-7 in Design 1).

The first interview design involved nine participants who elicited both elements and constructs (discussed below). RGT was used in Pilot 1 to elicit lecturers' teaching intentions (the elements written on cards) and secondly to use these elements to elicit beliefs and conceptions about statistics (the bipolar constructs). The repertory grid data from these nine interviews in Pilot 1 was then used to supply both elements and constructs with eleven participants in Pilot 2. In this study, I considered that an investigation of both individual and supplied construct relationships in the grid matrix might provide information about any distinctive meanings that had been attached to the supplied constructs and elements (Rayment, 2000; Yorke, 1978).

3.2.1.1 Participants

The study involved twenty statistical methods lecturers from a range of disciplines. Based on a list of statistical methods lecturers I created for this study using institutional databases and personal contacts, I deliberately invited statistics lecturers to participate if they had been teaching a quantitative or statistics methods module for at least three years at the time of the study. I considered that participants' expertise in teaching statistics was important since RGT required them to *construe* in their own words what teaching statistics meant for them in relation to their own *contexts*.

I contacted participants directly via email, explaining the purpose of the study and the time required for the interview. Participants on this study self-selected to participate and therefore comprised a sample of convenience. Convenience or opportunity sampling in this study could affect the general validity of any inferences drawn from the results of the RG interviews (Martínez-Mesa et al., 2016). The sampling process ended when a time limit of two academic terms was reached. RGT could be used with one participant (as a case study) and to a large number of participants, depending on the requirements of the study, the methods of data analysis or logistical considerations. In my study, I ensured that the lecturers were from a range of disciplines. Table 3.1 summarises the number of participants by discipline. My sampling for this study was more strategic and purposive because I was focusing on lecturers' unique contexts and I recruited participants prior to the data collection. I then examined each case and used within-case sampling (statistical content, statistical thinking, statistical process components and learning outcomes).

The first design consisted of eight statistical methods and one quantitative methods (with emphasis on mathematical methods) lecturers from two higher education institutions in the UK. Three lectures were from Business and Economics, three from Psychology and Social Science, one from Sciences (e.g. Biology), one from Geography and one from Mathematics. The remaining eleven lecturers in the second study design were from two higher education institutions, four teaching in the Psychology and Social Sciences, three Science, two in Geography and two in Engineering. One participant in Design 2 found it difficult to elicit repertory grid data and was therefore excluded from the analysis. In total, lecturers taught in three different universities at the time of the study, with a majority of 18 lecturers working in the same university.

Table 3.1: Number of participants by discipline

Discipline	**Number of participants**		**Total**
	Design 1	**Design 2**	
Business and Economics	3	–	3
Psychology and Social sciences	3	4	7
Science	1	3	4
Geography	1	2	3
Engineering	–	2	2
Mathematics	1	–	1
Total	**9**	**11**	**20**

3.2.1.2 Agreeing on a topic

In both designs, I started the interview with agreeing on the topic and ensuring that the interviewees understood the research question. I therefore started the interviews with a brief introduction into the purpose of the interview and the definition of statistical literacy, reasoning and thinking proposed by Ben-Zvi and Garfield (Ben-Zvi and Garfield, 2004, p. 7 and Appendix A). This was an opportunity to discuss my research more generally, obtain consent from participants and explain the purpose of my study in further detail. This stage was important in defining the *context* of my research to ensure homogeneity, i.e. a similar understanding of the interview research question (in Section 3.2.1.3) to ensure the elements are representative of the area under study (Kelly, 1955).

3.2.1.3 The Pilot interview research question

The module specifications in higher education in the UK emphasise aims (e.g. *learn basic statistical theory and procedures* or *acquire skills in statistical thinking*), intended learning outcomes (e.g. *on completion of the module students should be familiar with the use of statistics in the analysis of data, and should be able to select appropriate statistics for research tasks and compute them accurately*), content (e.g. *descriptive versus inferential statistics*) and methods of teaching, learning, assessment and feedback. I therefore explained my research question to participants, as follows:

> *"What kind of statistical knowledge and skills do you consider important for your students to learn on introductory statistical modules?"*

3.2.1.4 Choosing elements

In standard RGT procedures, elements are determined first and constructs are elicited based on distinctions made amongst these elements (Bell, 2004). Eliciting elements is an established process of defining the sample of elements (e.g. Bjorklund, 2008, asked teachers to select student artefacts as elements for the study). In **Pilot 1**, to elicit elements, I asked participants to articulate their own elements or objects of interest, that is the statistical skills and knowledge they considered important for their students to learn in terms of a specific module and student cohort.

I then gave each participant nine blank paper cards and asked the participant to write down on each card one skill or knowledge and understanding they considered essential for their students to learn on statistical modules. In repertory grid terminology, this is called an element. Although the participants could identify more or less than nine of these elements or cards, this number was considered sufficient for capturing the range of knowledge and skills on a statistical module and for using them in the time available (Curtis et al., 2008; Kandola and Pearn, 1992). This resulted in a list of statistical skills and knowledge (9 elements $\times$ 9 participants = 81 elements).

I considered reasonable to ask the participants, as subject matter experts, to articulate their own objects of interest (such as statistical skills and knowledge, learning outcomes or content) in terms of a specific module and student cohort. A list of statistical skills, knowledge and understanding identified by lecturers on these modules was likely to help me gather some empirical evidence for what lecturers considered important when defining curriculum and teaching statistics.

For **Pilot 2**, I created a summary of elements and constructs emerging from the first interview design. The elements from Pilot 1 (reduced to 24 as described in Section 3.4) were then supplied as elements in Pilot 2. It is common practice in RGT to supply the list of elements to interviewees (Fransella et al., 2004). With these participants, I used a summary of elements and constructs emerging from the first interview design as I considered the eleven participants to be equivalent to the first nine interviewees and therefore able to provide a view on the emerging themes.

Further, I assumed that the list of elements produced in Pilot 1 were representative and understood by the participants in Pilot 2. The elements were also assumed to be homogeneous, meaning that they were at the same level of construing from the participants' point of view. The elements elicited by other participats were "within the range of convenience of the constructs used", i.e. reflective of the context presented by the statistics students should learn (Fransella et al., 2004, p .18). To ensure the elements were representative and homogeneous, participants could add or remove elements based on their own context.

3.2.1.5 Construct elicitation

Constructs are discriminations that participants made between the elements, the statistical skills and knowledge. The basic assumptions applicable to construct elicitation are that constructs should be (1) permeable in the sense that they apply to a range of situations, e.g. statistical modules; (2) pre-existing or there is some degree of reliability in the constructs elicited; (3) understood by the researcher; (4) representative of the way the interviewee perceives statistics skills; (5) personal to the interviewee, i.e. reflective of their own practice in statistics education and importantly (6) bipolar, stating what an element is and is not. In my study, I aimed to design grid formats that maximise the usefulness of constructs (Fransella et al., 2004). Based on previous research using this technique (Shadbolt and Milton, 1999), comparisons of elements were likely to produce statements of these elements, which further specified them and went beyond the use of value or generic verbs (e.g. ‘understand inferential statistics’, ‘have knowledge of statistical techniques’, ‘apply statistical techniques’) often found in benchmarking statements or professional body standards (as in Section 2.4.3).

The participants compared the three cards and considered the following question:

> *“In terms of the skills, knowledge and understanding you would want the students on your module to learn by the end of their studies, which two cards are alike yet different from the third?”*

The twelve **constructs** used in Pilot 2 were also arrived at from an analysis of the constructs elicited by participants in Pilot 1. For example, the participants considered whether particular elements had high or low value on the job market in students’ field of studies, were easy or challenging for students to learn, required understanding on intuitive level or deep knowledge of the skills and how should such skills be learnt by students (independently through own activities or in lectures).

The participants were also asked to elicit their own constructs by comparing three elements (cards) selected at random from the pack. The alternation between ratings of provided and solicited constructs ensured the interviewees were kept engaged during the procedure. The procedure also allowed me to discuss the elements and the constructs emerging from Pilot 1 even when the participants did not rate them, because the participant might not have considered the construct relevant in his/her context or due to time constraints.

For example, one participant rated four provided constructs (out of twelve) and elicited two new constructs. We also discussed the remaining eight constructs without rating them. In this case, the resulting repertory grid contained 139 ratings and five missing ratings where the participant chose not rate a card on a particular scale.

3.2.1.6 Rating: relating elements to constructs

Lecturers were also required to relate the elements and constructs, either elicited in Pilot 1 or supplied in Pilot 2, by rating elements on a scale of 1 to 7 against the constructs. I anticipated that a 7-point rating scale was sufficiently wide to allow participants to identify greater levels of discrimination (Yorke, 1978).

Each time the participant specified a scale of contrasting statements (a row) in this way, the elements (columns) were linked to the construct (rows) by rating each element from 1 to 7 against the scale. A rating of one indicated the element that was most like the left-hand positive construct (e.g. ‘method of analysis’) and seven most like the right-hand negative statement (e.g. ‘sampling’). Ratings of 2, 3, 4, 5 and 6 went to elements in between these extremes. I discuss an example of a repertory grid from one participant in Section 3.4.1.

This process was repeated in each interview until the participant could not come up with any more statements (rows). The resulting data for each participant was therefore in the form of measurements on a number of variables (i.e. constructs) on a scale of 1 to 7.

3.2.2 Resulting repertory grids

Table 3.2 on page 82 shows an example of a completed repertory grid form from one participant in **Pilot 1** where the nine elements are written as column headings. The participant provided skills and knowledge such as element E_1 'be able to use simple inferential statistics', E_2 'appreciate different sampling methods', E_3 'have a basic appreciation of multivariate methods' and so on. The participant then provided a similarity statement (e.g. construct C_{1+}) that described what the pair of elements had in common on the left-hand column and a separate difference statement (e.g. C_{1-}) that described what made the singleton different from the other two on the right-hand column. For example, when comparing the elements E_2, E_8 and E_9 shown in the column headings in Table 3.2, the participant indicated that E_8 and E_9 were similar and the similarity statement to be C_{4+} 'methods of analysis', but felt that E_2 was about C_{4-} 'sampling'. In this way, the participant built a 'scale' with two contrasting or opposing ends, i.e. C_{4+} 'methods of analysis' at one end and C_{4-} 'sampling' at the opposite end. In this example, the interview resulted in fourteen constructs (variables), nine elements and 126 ratings.

Table 3.2: Example of a repertory grid from one participant

		Elements*										
	Construct (positive pole, C_+)	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	Construct (negative pole, C_-)	
C_{1+}	theoretical underpinnings of statistics	7	4	4	6	1	4	6	5	2	application of technique	C_{1-}
C_{2+}	easier to grasp	1	4	3	3	1	5	2	7	6	misunderstood by students	C_{2-}
C_{3+}	purely statistical	7	4	3	3	1	1	1	2	1	non-statistical skill	C_{3-}
C_{4+}	communication and interpretation skills, social skills, organisational skills	1	1	3	7	7	7	7	7	7	logic of research process that students need to follow	C_{4-}
C_{5+}	not essay style [skill]	4	7	7	1	2	1	1	1	1	essay skills	C_{5-}
C_{6+}	report structuring	5	1	1	6	3	5	7	6	4	not related to report writing/structure	C_{6-}
C_{7+}	less requirement of English language skills	1	7	7	1	4	1	1	1	1	language skills required	C_{7-}
C_{8+}	crucial to the curriculum	6	1	1	1	1	1	1	7	4	not critical to the curriculum	C_{8-}
C_{9+}	presence within lectures - high (repeated)	4	2	4	5	3	1	6	6	3	one off topic or two off	C_{9-}
C_{10+}	introductory	2	7	1	4	1	3	4	4	7	more advanced	C_{10-}
C_{11+}	value in job market high (research skills)	3	1	1	3	1	1	2	7	7	low value in the job marked	C_{11-}
C_{12+}	hand computations required (numbers)	7	7	7	1	7	3	1	1	4	no hand computations	C_{12-}
C_{13+}	essential for good statisticians	7	4	3	1	1	1	1	2	1	not essential in statisticians	C_{13-}
C_{14+}	low level	1	4	5	4	4	5	4	6	7	high level	C_{14-}
C_{15+}	least demanding year 1	NA	NA	5	4	4	4	NA	4	NA	most demanding year 1	C_{15-}
C_{16+}	least demanding year 2	NA	NA	NA	3	4	3	3	4	5	most demanding year 2	C_{16-}
C_{17+}	least demanding year 3	NA	NA	3	2	2	2	2	3	4	most demanding year 3	C_{17-}

*The participant elicited 17 bipolar **constructs** (positive C_+ and negative C_-), nine **elements** (E_1 to E_9) and 153 **ratings**.
E_1: team work
E_2: ability to write a report using statistics
E_3: ability to understand quantitative research in journals
E_4: t-tests
E_5: difference between descriptive statistics and inferential
E_6: statistical inference
E_7: chi-square
E_8: one sample tests
E_9: relationship between effect size, power and significance

3.3 Pilot 3: observations of teaching

Towards the end of my repertory grid study, I contacted a lecturer teaching statistical methods to observe his teaching. The module was an introductory module taught in the programme's second year of study. I considered this module to be typical of most students' experience of higher education. The lectures I observed (four hours in total) were delivered to a large group of students (nearly 150) and the teaching strategy involved a two-hour weekly lecture and two hours in the laboratory every other week. I was able to observe two lectures covering regression and correlation, classic topics in introductory inferential statistics.

The lecturer taught the lectures using slides and gave students the opportunity to carry out simple calculations during the lecture using handouts with 'gappy notes' and ask questions. I made the decision to sit in the middle of the lecture theatre to be able to hear the lecturer and also have a view of most of the students. In that early phase, I observed that the lecturer explained the 'harder' statistical concepts while students solved some of the exercises using 'basic' calculations (plugging numbers into formulae).

In the two laboratory sessions (four hours of observations), I followed the lecturer who predominantly answered questions about what I considered to be 'basic' concepts (significance testing, the differences between regression and correlation, the regression equation) when compared to what he taught in the lecture earlier in the week. Towards the end of each session, the lecturer summarised the objectives of the session and the answer to the questions on the sheet.

Although I planned a detailed observation sheet for each of the sessions following a quantitative kind of observation, I realised that such a structured, systematic approach in a natural setting (the lecture theatre or computer laboratory) where I did not know in advance what to look for were not possible. It was clear to me at this point that I could not capture the lecturers' everyday, social setting and behaviour over many weeks using pre-defined categories and codes. As my background was in cognitive psychology and measurement, my initial instinct was to count and measure what my participants were doing. Soon after I entered the lecture theatre I realised the richness of the data which could be brought to light through more intrusive strategies using participant observations within an interpretative paradigm (Section 4.1).

Despite the tentative nature of the work I carried out during the pilot phase, I feel that this research study with lecturers has influenced to a great extent the decisions I took regarding the methods of data collection and analysis in the main phase of the research study. This pilot work helped me to clarify my research questions and methodology in the main study, which I explore in Chapter 4.

3.4 Data analysis of repertory grid interviews

The data resulting from these interviews was therefore in the form of qualitative statements (words), quantitative data (ratings) and audio-recordings of conversations with the participants about these issues.

In designs 1 and 2, each repertory grid had the same number of different elements or observations (9 in design 1 and 24 in design 2) and a varying numbers of different constructs or variables. Therefore, the grids provided by each participant had in common only the number of constructs. To answer the research questions, for both repertory grid designs, I analysed each repertory grid separately as a case study using both statistical analyses of ratings and qualitative analyses using classifications and open coding.

The quantitative data analysis involves *statistical analyses* using statistical software and also, importantly, a *qualitative interpretation* of the statistical output. The data analysis procedures for all grids involved three stages. In the **first stage**, quantitative data were analysed for each interviewee using statistical analyses. In addition, the Pilot 2 repertory grids, since the aim was to obtain ratings against a common set of elements (observations) and constructs (variables), I intended to analyse using statistical methods in order to aggregate the outcomes from the study in a systematic way. In the **second stage**, the reduced data from phase 1 was analysed using qualitative data analysis methods. Finally, in the **third stage**, the interview audio data I analysed using open coding. These procedures are detailed in this section.

3.4.1 Statistical analyses of ratings

The data for each repertory grid I analysed with Idiogrid 2.4 software to undertake singular-value decomposition (SVD) as a data summary or data reduction tool (Grice, 2002). SVD enables both elements and constructs to be represented together. Idiogrid is a quantitative data analysis software package that allows the analysis of repertory grids, such as the ones collected in this study. Customarily, in a dataset, the persons (observations) are organised as rows and the variables such as test scores as columns. In a repertory grid such as the one presented in Table 3.2, this is reversed as the observations are the elements (i.e. E_1, E_2) in each column and the variables are the constructs with two contrasting ends (i.e. C_{1+} versus C_{1-} to C_{14+} versus C_{14-}) on each row.

Taking the data in Table 3.2 as an example, for the nine ratings for each of the fourteen statements C_1, C_2,..., C_{14} against each element, I have 153 ratings (17 constructs $\times$ 9 elements). To make sense of this data, one option would be to compare actual ratings for each element and construct or calculate the correlation coefficients among constructs. This however would result in too many comparisons or correlations to interpret. Since the constructs are all about statistical skills and knowledge, factor analysis could be used to model the correlations among the statements using a smaller number of underlying or hypothetical factors (Manly, 2005). Statements that correlate highly are considered to be similar in meaning and therefore represent a factor (Afifi and Clark, 1996). Since the Factors are uncorrelated with each other, each Factor is considered to measure a different 'dimension' of the data, in this case a participant's view of the statistical skills their students ought to learn. An example of the steps involved in factor analysis for one repertory grid taken from one participant is provided in Appendix B.1, p. 267.

The Factor analysis reports distances between constructs showing the likelihood that constructs appear near each other by chance. Further, for each grid, cognitive maps group constructs with similar meanings together into clusters and elements relating to those clusters. To ensure the validity of the procedures, we also checked the plots against the grid data to ensure patterns were fairly consistent across the two. Figure 3.2, p. 90 shows such a SVD plot with the elements (red dots and black font) and the constructs (blue dots and font). Here, red dots represent element Factor scores, blue dots construct Factor loadings and grey lines the eigenvectors.

This participant construed the element 'difference between descriptive and inferential statistics' as being 'crucial to the curriculum', 'easier to grasp', 'introductory concept' about 'theoretical underpinnings of statistics' and with 'value in the job market'. Similarly, 'team work' was a non-statistical skill, closely related to 'communication and interpretation skills'. In this participant's view, the 'ability to understand and read quantitative research papers' does 'not require hand calculations' as it is about 'reporting' skills.

The ten grids from Pilot 2 interviews, I intended to analyse using a Generalized Procrustes Analysis (GPA) with elements matched and a varying number of constructs (Grice and Assad, 2009). GPA is a multivariate statistical technique for analysing three-dimensional data matrices. This technique has been suggested for analysing both aggregate grids (multiple grids in one analysis) and individual grids. The goal of this analysis was to identify common patterns in how the participants organised the skills and knowledge they considered important for their students to study on statistical. In essence, the analysis promised to construct rating scales from the constructs rated against a common set of elements (cards). This process allowed for the use of different constructs by participants and of a different number of constructs in each interview. So the repertory grids may be matched according to the elements but not the constructs. In this way, the analysis was meant to identify agreement amongst participating lecturers regarding the statistical skills and knowledge most important in undergraduate education. To achieve this, GPA involves a combination of Procrustes rotation and Analysis of Variance (ANOVA) that computes a consensus matrix which is an average of the rotated matrices of ratings.

However, following extensive data analysis of the repertory grid data collected in this study, I noticed that the order in which the repertory grids were entered into the analysis mattered to the output of the analysis, i.e. if the grids were entered in different order in a table, the results were very different. For this reason, I did not include this often used analysis in the analysis presented in this Chapter. Instead, due to the issued with aggregating the information from Pilot 2 using GPA, the analyses of these eleven repertory grids used Factor analysis and qualitative analysis of the interview data.

3.4.2 Qualitative analysis of elements and constructs

To manage possible overlap between the elements obtained in this study (Raymond and Neustel, 2006), I analysed the data using **qualitative techniques** by classifying the elements and associated constructs in various ways based on sequential and conceptual frameworks presented in previous research (Allan, 1996; Newton et al., 2011; Wild and Pfannkuch, 1999). I anticipated that such an analysis would allow me to examine the statistical skills and knowledge the participating lecturers believed to be important for their students to learn, the participants' beliefs about their students' learning and to identify which aspects of such a schema or framework may or may not be required of students in a range of domains. If such classification schemes were present in the data, then I could use them in further research in my Main study (Chapter 4).

In order to identify **characteristics in the elements** provided by participants, I classified the repertory grid elements into the following categories:

1. The conceptual types of learning outcomes proposed by Allan (1996): statistical-based, personal transferable, generic (research question 1).

2. Whether the repertory grid elements represented content or a learning outcome as an indication of how the participants construed the statistical curriculum (Hussey and Smith, 2002) (research question 2).

3. The same classifications were used with a **list of actual learning outcomes** from actual statistical modules (research question 3).

4. The sequential process components of the statistical investigative process: formulate research questions, collect data, analyse data and interpret results (Franklin et al., 2007) (research question 4).

5. Whether there was evidence of expectations for statistical reasoning in the elements and/or constructs provided by participants, i.e. student will reflect on or evaluate statistical procedures. One element could be classified against one or several process components as well as statistical reasoning (research question 5).

The modules used in this study were all delivered at university in the UK. The modules were initially identified for possible consideration through a keyword search using the terms 'quantitative' (14 results) or 'statistic*' (15 results) within the module content of all module specifications available on a university virtual learning environment. The criteria for inclusion in this survey were the presence of a significant proportion of statistical knowledge, skills and techniques in the module content and/or the application of these to solving domain-relevant problems. Modules that mainly covered mathematics, probability or domain-specific methods were excluded from the analysis because their learning outcomes and content were not specifically relevant to my study.

All the parts of the module specifications were summarised in a database for each module, containing the following fields: primary subject domain, module title, code, level, credit rating, aims, nature and weighting of assessments, time required for the completion of each assessment, timing of the assessment during the module, administration conditions (invigilated or take away), assessment methods (constructed-response or selected-response), feedback arrangements, total number of measures used in the module, the nature and type of teaching activities, learning outcomes and content.

In total, eighteen module specifications were included in this analysis. The modules covered six domains, Business and Economics, Engineering, Psychology, Sociology, Geography and Mathematics. The module specifications resulted in 28 module aims and 197 intended learning outcomes. The modules included in this part of the analysis are outlined in Appendix B.2, p. 273.

The content analysis using pre-defined classifications I carried out for the actual module learning outcomes, the repertory grid elements and the Factors (constructs related to elements) as explained above. The analysis of the data from the first nine interviews involved open coding taking a grounded approach in order to identify common themes amongst participants. This analysis of the first nine interviews informed the design of the repertory grid interviews with the final eleven participants.

3.4.3 Qualitative analysis of interview audio recordings

In addition to the content analysis of repertory grid elements and constructs, I produced qualitative summaries of the audio recordings which I analysed using open coding with the view of identifying themes and sub-themes to substantiate the qualitative analysis using classifications and the statistical analysis using factor analysis. Transana Professional version 2.50 was used to analyse the interview summaries. Transana is a computer programme that is used in the qualitative analysis of large collections of text, still image, video, and audio data. The analytic process then proceeded by comparing different pieces of data (elements, constructs, open interview data). In my study, I intended to characterise the context of these lecturers' teaching by identifying a set of issues and conditions that gave rise to problems or circumstances ranging from the most macro to the micro. In this way, I was hoping to be able to identify tensions or contradictions in the teaching of statistics by exploring the participant's context in which this teaching occurred.

3.5 Findings from repertory grid interviews

The data analysis involved a complex set of procedures, which revealed that our participants focused simultaneously on curricular content, outcomes, pedagogic and contextual issues when construing their own intended curriculum. This section presents the data analysis and findings from repertory grid interviews that aimed to identify a suitable procedure for characterising module outcomes.

The interviews in **Pilot 1** lasted from 50 to 110 minutes. During the interview, each of the nine participants provided nine elements, i.e. skills, knowledge or understanding, giving a total of 81 (=9 elements x 9 participants) ratings (Section 3.2.1.4). Each participant produced a varied number of constructs (or contrasts), ranging between 8 to 17, giving a total of 143 constructs (Section 3.2.1.5). In order to analyse this interview data both qualitatively and quantitatively using Factor analysis of ratings, I asked the interviewees to rate each of the nine cards on a seven-point scale formed by the bipolar contrasts. In total I obtained 1,287 ratings (9 elements × 143 constructs). This process of generating elements and constructs I explain in more detail above in Section 3.2.1.

Similarly, the interviews in **Pilot 2** lasted from 1.04 to 1.42 hours and 796 minutes (approximately 13.5 hours) in total. The participants rated twenty-one new constructs and fifty eight provided constructs. Ten participants agreed that the twenty-four elements provided to them were representative of the range of statistical knowledge and skills they considered important. One participant was excluded from the analysis as the constructs were not the participant's 'personal descriptions of the domain' (Section 3.2.1.1). One participant added one element ('measurement error') and one other participant excluded three elements ('dependent and independent data', 'assumptions of statistical tests' and 'randomisation').

3.5.1 Conceptual types of learning outcomes and content

To characterise the participants' beliefs about the learning outcomes important for students to attain on statistical modules (pilot research question 1) and about content (pilot research question 2), I first looked to see whether the repertory grid data represented a system of *learning outcomes* for the students to learn by the end of the module, i.e. well-defined rules and representations, or *content* statements.

The classification of the 81 repertory grid elements provided by participants into the three types of learning outcomes (statistics-based, personal-transferable and generic Allan, 1996) revealed that 52 (68%) were statistical-based learning outcomes (e.g. knowledge, understanding and application of statistical methods and techniques), 20 (26%) personal-transferable (e.g. information technology, team work and communication skills) and 4 (5%) generic (e.g. dangers of 'blind statistics', ability to read critically or critically interpret data).

Second, the 197 actual module learning outcomes were classified into the three types: statistical, personal-transferable and generic. This classification resulted into 202 categories of which 103 (51%) were domain/statistical-based, 82 (40.5%) were personal-transferable and 17 (8.5%) were generic learning outcomes (Section 2.4.3).

Finally, the categorisation of 81 elements and the constructs that explained them using the Factor analysis plots revealed that 52 (73%) were statistics-based, 19 (26%) were personal-transferable and 1 (1%) generic. In thirteen cases, the constructs changed the way I interpreted the elements on their own. For example, 'statistical considerations need to be included at all stages of a study' I found difficult to interpret on its own. The constructs relating to this element were 'easy concepts for students to grasp; knowledge of concepts, right/wrong answers; concept used to communicate statistical procedures, findings and recommendations'. This might mean that knowledge of statistical concepts is required for communicating the analysis, and may therefore be classified as a statistics-based type of learning outcome. It could also mean that it is about communication skills, a personal-transferable type of learning outcome.

Regarding the participants' beliefs about content (pilot research question 2), of the 81 repertory grid elements provided by participants, 41 (50%) represented statistics-based content statements (e.g. 'regression', 'correlation', 't-test', 'report writing') rather than a learning outcome. It may be that the participants construed their intended curriculum considering the content of the syllabus, textbooks, time allocated to different activities and assessment tasks rather than intended learning outcomes (Hussey and Smith, 2002).

To investigate how lecturers' beliefs about statistics curricula compares to actual module specifications (pilot research question 3), the analysis of repertory grid elements and Factors (element-construct pairs) indicated a different emphasis than in the case of the actual module learning outcomes. In the case of repertory grid elements, 68% were statistics-based when compared, 71% for Factors but only 51% for the actual module learning outcomes. This might stem from the experimental design which asked participants to produce statistical knowledge, skills and understanding that students should learn. It is also possible that regulatory requirements in higher education impose a particular format on the way module specifications are designed. It might also mean that participants focused more on statistical-based learning outcomes, although further information may be required before drawing such a conclusion.

3.5.2 Classification into statistical process components

To characterise the participants' beliefs about the teaching and learning of the statistical process components (pilot research question 4), the classifications of the 197 actual learning outcomes into the four statistical process components (formulate research questions, collect data, analyse data, interpret data) revealed that analysing data involved 48% of learning outcomes and interpreting results 24.8%. The RG elements resulted in 93 classifications of the four statistical process components. The majority of elements were about analysing data (49%) and interpreting results (43%). This suggests a similar finding as in the analysis of actual module learning outcomes that also seemed to emphasise the two later stages of the statistical process.

In Table 3.3, I show the number of learning outcomes by the four statistical process components and by statistical reasoning. The classifications of the 197 *actual* module learning outcomes into these categories revealed that analysing data involved 48% of learning outcomes and interpreting results 24.8%.

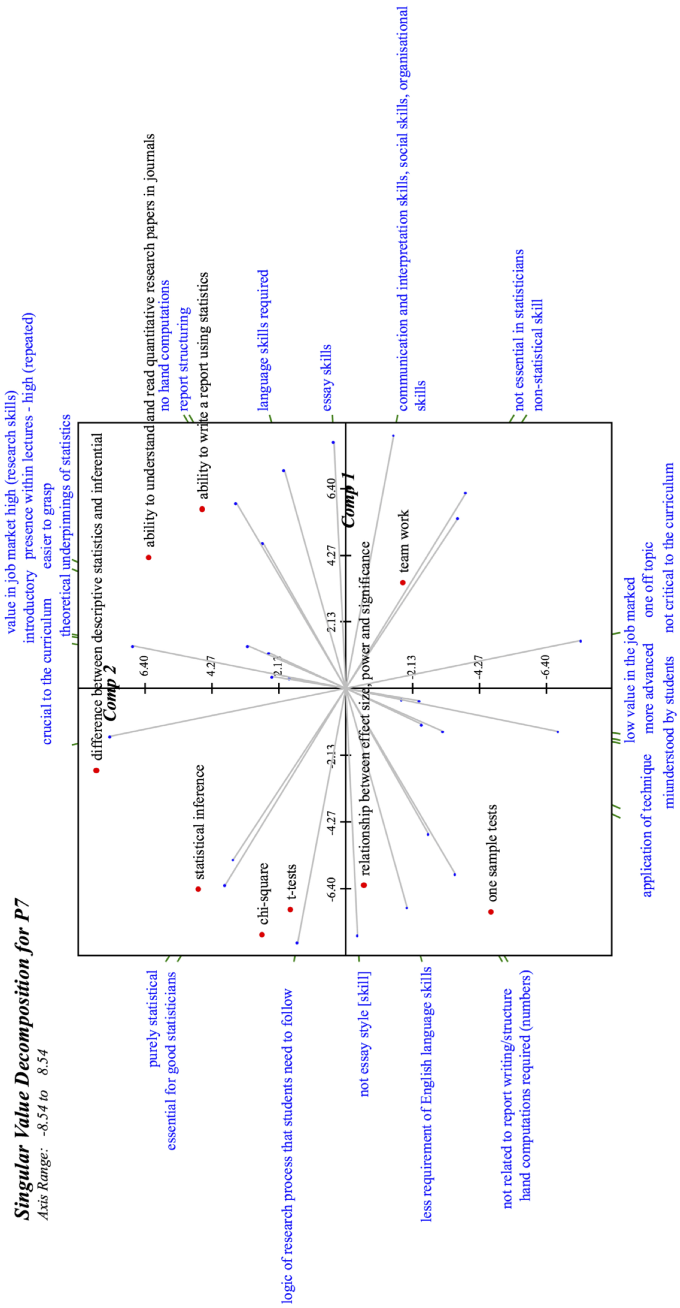


Figure 3.2: Example of a Singular-value decomposition (SVD) plot (Idiogrid 2.4.)

Table 3.3: Classifications by process components and by statistical reasoning

	Process component				
	Formulate Questions	*Collect Data*	*Analyse Data*	*Interpret Results*	*Total*
Actual module learning outcomes (=197)					
N by process component*	42	24	116	59	241*
% in overall number by process component	17.43%	9.96%	48.13%	24.48%	100%
N by statistical reasoning	23	1	13	15	52*
% in overall number by statistical reasoning	54.76%	4.17%	11.21%	25.42%	21.58%
Repertory grid elements (=81)					
N by process component*	3	4	46	40	93*
% in overall number by process component	3.23%	4.30%	49.46%	43.01	100%
N by statistical reasoning	0	0	4	14	16*
% in overall number by statistical reasoning	-	-	4.30%	15.05%	19.35%
Repertory grid Factors (=32)					
N by process component*	0	1	16	21	38*
% in overall number by process component	-	2%	42%	55%	100%
N by statistical reasoning	0	1	3	9	11*
% in overall number by statistical reasoning	-	2%	7%	23%	28.9%

Note: * one learning outcome, element or Factor could be classified against one or several process components

3.5.3 Expectation of statistical reasoning/thinking

Table 3.3 shows that all statistics modules included at least one learning outcome that promoted statistical reasoning (pilot research question 5). Out of the 241 classifications, 21% (52) involved statistical reasoning. This might mean that, based on my interpretation of these learning outcomes, statistical procedures rather than statistical reasoning, conceptual understanding or conducting statistical studies were emphasised in the module learning outcomes included here. Of the 81 elements, 18 (19%) were classified as requiring statistical reasoning.

Using the SVD plots (Figure 3.2, p. 90), each pair of element-construct was also coded into the four statistical process components. The results of this analysis suggest that the proportion of statistical reasoning was 29% (11 Factors out of 38) of classifications. This may mean that although there was less expectation of statistical reasoning in the module learning outcomes and the elements and constructs when considered in isolation, the participants in fact described the elements as involving a higher proportion of statistical reasoning when considered in relation with their constructs.

The emphasis on content rather than intended learning outcomes may also explain the higher proportion of statistical-based elements. Further, the participants' reliance on content rather than statistical reasoning was despite the fact that the participants were presented with a definition of statistical literacy, reasoning and thinking (see Section 3.2.1) at the beginning of the interview, which was then discussed in the context of their module. On the other hand, RGT required participants to provide a list of elements within a short time and not overtly defined as learning outcomes. Also, the factor analysis and the classifications in Table 3.3 I carried out inductively, based on mathematical modelling and subjective judgements that made the link between observable data and my hypothetical Factors.

3.5.4 Participants' beliefs about the intended curricula

The next step in my analysis was to investigate the grid data (constructs and elements) from all the participants in order to identify common themes amongst participants and characterise the collection of constructs that attracted most interest using open categories. The main themes were further categorised into sub-themes which I linked to the pilot study research questions (Section 3.1.1, p. 70). This process was iterative in the sense that for each new theme produced, I re-inspected the data to check whether an element or construct could be classified against this newly defined theme. Once all the themes and sub-themes were finalised (see Appendix B.4, p. 285), I carried out the analysis multiple times to ensure it was accurate. I also discussed these themes with two researchers in a meeting in order to confirm these interpretations of the data.

In this way, I identified three main ***themes***:

1. **Teaching with "context":** statistical knowledge and understanding versus application of statistical methods and techniques within a "context" (*23 classifications*).

   1.1. Knowledge of theory, conceptual knowledge and understanding **versus** applying techniques in a real world context (**pilot research question 1**).

   1.2. Knowledge and understanding of statistical methods, techniques and procedures (**pilot research question 1**) **versus** statistical thinking and reasoning (evaluate or reflect on statistical procedures, integrate and synthesise the context knowledge with statistical knowledge, work with models to draw inferences from data and communicate a statistical argument about the real situation) (**pilot research question 5**).

   1.3. Statistical (exploratory techniques, inferential statistics, graphical representations, probability, estimation, significance tests, ANOVA, regression, correlation, bi-or multivariate methods, time series analysis) **versus** non-statistical outcomes (mathematics, confidence with statistics, team work, logical thinking, communication skills – report writing, critical thinking, creativity, using resources – information technology) (**pilot research question 2**).

2. **Teaching of statistical process components:** formulating questions and data collection versus analysing and interpreting data (*8 classifications*).

   2.1. Collecting data **versus** analysing and interpreting data (**pilot research question 4**).

   2.2. Analysing data (applying techniques) **versus** interpreting data/models (communicating with data) (**pilot research question 4**).

3. **Student learning:** curricula that students ***can*** learn versus ***should*** learn (*22 classifications*).

   3.1. Easy (basic, straightforward, intuitive, understood, learnt from textbooks) **versus** difficult (demanding, challenging, complex, advanced, hard, misunderstood, mathematically-based, learnt by doing) ideas/ concepts/ methods (**pilot research question 1**).

   3.2. Knowledge or skill important (relevant, essential, core) **versus** not important (not essential) to learning particular aspects of statistics (**pilot research question 1**).

   3.3. Need to cover in the module (presence within lectures, high value in employment) **versus** do not need to cover (low value in employment or for carrying out particular statistical analysis) (**pilot research question 1**).

This analysis indicates that issues with student learning and curriculum design (theme 1 and 3) attracted most interest with 23 classifications and 22 respectively. This included aspects such as ‘is easy versus difficult for students to learn’, ‘is important versus not important in learning statistics’ or ‘need versus don’t need to cover in the module’. These constructs included clear opposites (e.g. easy versus difficult) and were therefore more accessible for participants to produce using the repertory grid technique.

The participants also seemed to characterise the statistics students should learn in terms of knowledge and understanding versus application (skill) (theme 1) and statistical versus non-statistical (theme 1). Seven factors were classified as statistical knowledge and understanding of methods and techniques versus statistical reasoning and thinking (theme 1, sub-theme 1.2). Nicholl (2001) emphasised that achieving an appropriate balance between the theory and application of statistics is a major challenge in teaching and learning statistics at tertiary level. This dataset seemed to associate communication skills (theme 2, sub-theme 2.2) with statistical reasoning and thinking (theme 1, sub-theme 1.2).

Constructs contrasted interpreting data, which was about the meaning of data, with analysing the data, which was about applying techniques and procedures. Further, the Factor analysis for participants 2 and 4 suggested that these participants associated statistical thinking (sub-theme 1.2) with higher-level statistical abilities while techniques/procedures were associated with basic abilities (sub-theme 3.1). On one occasion, statistical thinking seemed to be required in employment (sub-theme 3.3, participant 8). Five participants (1, 2, 7, 8, 9) provided elements which included report writing and six participants (1, 2, 6, 7, 8, 9) the ability to read quantitative research. Two factors associated communicating with data (theme 2.2) with statistical thinking (sub-theme 1.2) and contrasted 'communicating with data' with techniques/data analysis. Aspects of the statistical process I classified against eight Factors. There were no grids which referred to the entire statistical process, but the emphasis across participants appeared to be on the analysis and interpretation of data, the later stages of the statistical process.

In summary, statistical thinking (sub-theme 1.2) seemed relevant for interpreting and communicating with data (sub-theme 2.2), which were also more demanding skills. The participants in this study believed that statistical thinking, interpreting and communicating data were contrasted with statistical techniques and procedures were easier/less demanding skills (sub-theme 3.1). The finding that the participants on this study referred only to certain aspects of the statistical investigative process was similar to the classification of actual module learning outcomes (as in Section 3.5.1). The actual module learning outcomes surveyed here also seemed to emphasise only certain stages of the statistical investigative process, i.e. the analysis and interpretation of data with little opportunity for students to experience the first two stages of formulating research questions and collecting data.

The analysis of *all* interviews (Pilots 1 and 2) resulted in a coding scheme produced using the emerging topics, the constructs and elements provided to participants and the repertory grid data. In total there were ten parent codes and sixty five codes in the code system. The code system is summarised in Appendix B.5. Statements within the interview summaries were qualitatively reviewed and coded to as many parent codes and codes as applicable. The ten themes summarised in Appendix B.5 expanded my interpretation of the emerging themes and sub-themes identified in the analysis of the Pilot 1 interviews.

Based on my coding of this data, the notion of *relevance* of students and lecturers' field of study and of the statistics in the teaching of statistics made the link between teaching and statistical processes, summarised in the diagram in Figure 3.3. I expressed these links as contradictions within the teaching of statistics at university.

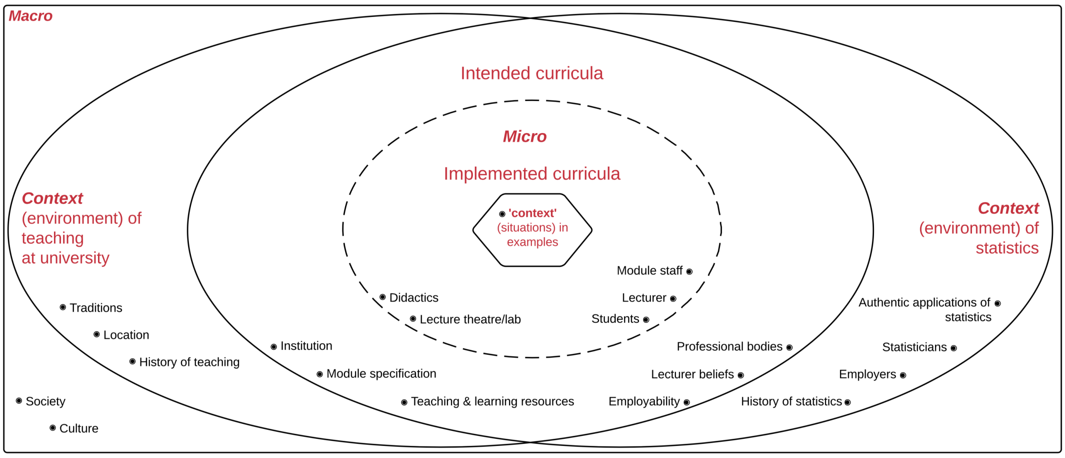


Figure 3.3: Pilot study conceptual framework

3.5.4.1 Teaching processes: contrasts among curricular, didactic, pedagogic and institutional

A first contrast, between curricula (what students should learn) versus pedagogy (beliefs about what students can achieve), highlighted a tension between what lecturers considered important for these students to achieve on their modules and the lecturers' experience with the student cohorts. Two lecturers believed that, although students on their modules were not introduced to 'complicated statistical techniques', the mathematical underpinnings were important. For Participant#1, it was important to know that 'there are some mathematical underpinnings', but that the focus on his module was on applications of techniques in context: 'in psychology you will need to understand what is the point of using that technique in relation to what it is the researcher is trying to answer, the research question'. Participant#12 however, although agreed that the extent to which students were able to understand mathematical underpinnings was a function between student and topic, especially considering some of the students' low mathematical attainment, without mathematics, the learning could only focus on routine applications of techniques:

> **Quote 1 (Pilot 2, Participant#12)**
> Most applications of statistics involve making a decision, which recipe to use. What matters is the process of making that decision.

Students on his modules were able to follow 'recipes' but 'did not know what the recipes were made of'. However, statistics was about 'translating the real world into a mathematical framework', about seeing 'the connection between a linear equation and a description of some experiment that they want to do'. While in his view, most students were able to grasp some of the mathematical detail, it was difficult for them to decide 'what [statistical test] to apply and getting that real world situation into the list of two numbers'.

Here, a second contrast became apparent, between curricula, what students should learn versus statistics that, in some of the participants' view, relates context to mathematics. While Participant #2 considered challenging translating the real-world context into mathematical meaning, for Participant#1 the challenges were about formulating the real world problem into a research question and then relating the statistical analysis back to the original problem. These issues show the inter-dependence between statistics, curriculum design, teaching strategies, student engagement and student learning (didactic), lecturers' beliefs about statistics and about the students (pedagogy) and institutional constraints.

The third contrast was about the didactic principles (how to teach) of these two lecturers. Participant#12 intended to start from contextualised real-world examples, while Participant#11 suggested an approach starting from general principles towards contextualised knowledge. For Participant#1, a Psychologist, the approach to planning his teaching was to start with applications of statistical techniques in a context relevant to students' field of study and emphasise personal-transferable outcomes (critical thinking, team work). This lecturer believed that students can only learn statistics in the workplace, rather than at school, although all his students were able to achieve the module learning outcomes. Participant#5, an Economist, believed that the teaching should start from abstract statistical concepts or objects, followed by a large number of practical examples in a context. This tension between teaching general statistical principles and concepts and providing contextualised examples also reflected the importance these lecturers placed on theoretical, scientific knowledge versus practical, contextualised statistical applications. This resonates with Eichler's (2007) finding that there are no 'pure' *traditionalists*, i.e. teachers that establish a theoretical basis for stochastics, algorithmic skills and abstract structures but no applications. My data suggested that statistics lecturers are closer to *application-preparers* who want students to grasp the 'interplay between theory and applications' and to *everyday-life-preparers* who develop statistical methods while examining applications.

To some participants, students' independent study, after the lecture, was key to learning statistics. In addition, Participant #2 believed that the crowded curriculum on his module (one topic every week) meant that the module did not go into enough detail and that students 'did not have the opportunity to gain a feel of what was going on'.

3.5.4.2 Statistical processes: contrasts among statistical process components, applications, analysis and theory

The coding for statistical processes centred around contrasts among statistical process components, statistical applications, statistical analysis and statistical theory. The lectures' views differed with regards to which processes were important for their students. Participant #2 believed that without the knowledge and understanding underlying the practical applications of statistics, students would not be able to progress to the later stages of interpretation and communication with statistics, considered to be higher level. Applications of statistical techniques involve a decision about which technique to use, so students are required to understand the process of making that decision through the integration of examples and theoretical underpinnings. Participant #1 considered however that the initial stages of formulating the research question and design of studies then finally building a narrative about the statistical analysis were more important since the analysis and reporting stages can follow conventions within the context of a particular domain. So, on the one hand, statistical modules could focus on learning non-routine techniques that are not well specified and on the other hand on becoming familiar with routine statistical techniques applied to particular situations.

3.5.4.3 Contrasts between teaching and statistical processes: the relevance of context

The participants considered that statistics relates context to mathematics. Participant #1 viewed statistics as a continuum between applying, practical issues such as statistical software and techniques and theoretical knowledge. He conceptualised statistical analyses involving three phases, a pre-decision phase which is about defining the research question, a practical phase and a post-decision phase which the analysis means. Theoretical knowledge underpins the first and third phase, which he considered challenging for students, while the practical phase was about using software and following pre-set procedures:

> **Quote 2 (Pilot 2, Participant#9)**
> That's what I try with my students. Press these buttons, see what comes out. But what does that mean? The decision is (.). There is a pre-decision phase, and a post decision phase. Why are we doing this in the first place, what is the question And a post (decision phase). OK, I have done this, what does it mean? And the practical being in between, which actually does not matter too much. Just doing. They have no problem doing it. But when they look at the output, they do not know what it means. Statistical packages can get something out of it.

The above quote suggests tensions between the statistical investigative processes that involve applications in context (non-statistical), curricular decisions (what to teach) and didactic (how to teach), whether for instance students should learn the theoretical bases of statistical methods versus learning to apply statistics in context, focusing on routine procedures rather than a deeper, theoretical understanding of techniques.

Further, the participants discussed the relevance of the statistics taught on their modules to students' field of study, degree programme and future potential employment, suggesting a contrast between institutional (control over what lecturers teach) versus societal (what is valued in the job market), pedagogy (lecturers' beliefs about students) and statistics. Participant#17 recognised that some statistical concepts and topics are only important to some disciplines, such as *dependent and independent data* which is not very important for engineers. Participant#1 considered that statistics is a useful tool only to those who apply particular experimental methods, while similarly Participant#2 recognised that it is not necessarily the case that the students on his module will require statistics in future employment. Further, Participant#1 considered that the teaching of statistics to Psychologists should start from a context, 'an everyday description', a psychological construct such as differences in memory or emotions between two groups and then 'tell students that this is a concept which statisticians call variance'.

3.6 Conclusions emerging from the Pilot studies

The interviews using RGT allowed lecturers to construct in their own words the beliefs about the intended curricula in their own context. The classifications into the statistical process components, whether there was a requirement for statistical reasoning/thinking and the types of learning outcomes revealed some similarities between actual official curricula and the grid data. It may be however that the broad nature of some of these elements produced by participants might not reflect the implemented teaching (Pierce and Chick, 2011).

For example, 'statistical inference' may involve a number of activities and reasoning processes, which the participants did not express in these brief elements. The analysis indicated that the participants focused more on content than on learning outcomes. This finding may however be due to the interview design, which predominantly focused on statistics. Despite some possible limitations with the research design, we consider RGT to be useful in capturing lecturers' tacit knowledge about their intended curriculum. Future studies could for example allow participants more time to prepare for the RG interview and provide richer elements for construct elicitation.

3.6.1 Implications of using the repertory grid technique

The participants in the Pilot 1 interviews, elicited the elements, constructs and ratings, thus making them relevant and well known to them and reducing the amount of researcher bias into the grid structure (Curtis et al., 2008; Fransella et al., 2004). This repertory grid interviewing design, that allowed the participant to provide the grid data, appeared to be useful in such an exploratory study which sought to identify a diverse range of beliefs about statistical skills and knowledge. The grids could then be analysed efficiently, in contrast to other qualitative techniques such as in depth or structured interviews, which necessitate significant time committed to data analysis (Fransella et al., 2004).

The interviews which supplied the elements and constructs (although the participants could elicit their own constructs as well) also proved useful in validating and discussing the grid data with the participants and for carrying out the overall data analysis. The qualitative analysis of content that involved grouping, categorising and frequency counting depended on the way I interpreted the data. While this is a challenge for any qualitative study, in the case of repertory grids this is augmented because the data provided by participants has very specific characteristics, e.g. brief constructs. In order to obtain more information in the time available, in Pilot 2 I supplied the elements and constructs to participants. This approach was useful for confirming the findings from data analysis of Pilot 1 interviews and for being able to clarify the emerging themes.

In order to identify whether the repertory grid elements obtained were clear and specific as discussed in Section 3.2.1, I inspected the 81 elements and the statements that describe them. In my view, the elements provided by the participants in the study seemed less clear or specific than the actual module learning outcomes. For example, I compared the grid data for participant 3 to the actual module learning outcomes the participant was using at the time of the interview. One module learning outcome, ‘be able to store, analyse and present geographical data in a common spreadsheet programme such as EXCEL’ I related to the grid elements ‘be able to describe data using descriptive stats and appreciate graphical methods’ and ‘be able to interpret graphs and tables and draw reasonable conclusions’. These examples appear to show that the module learning outcome was more specific since it included the means of analysis (information technology) and the type of data (geographical). On the other hand, the grid elements included information normally classified as module content (e.g. regression, correlation, confidence intervals). The grid elements seemed to be missing the “context” in which the student was expected to demonstrate learning or how well the student was expected to do that (Otter, 1994).

The elements provided by participants led them to elicit constructs which were repeated across participants, meaning that it was possible to identify common themes. However, the elements did not appear to vary enough, i.e. represented a narrow range of statistical abilities or represented predominantly domain-based content (e.g. t-test, time series techniques, knowledge of statistical inference, decision making) rather than learning outcomes.

The constructs produced by participants do not appear to include a range of situations, circumstances or activities that students should carry out in practice, although they reflect the complexity of work on undergraduate statistical modules. The proportion of elements produced by participants that I described as content (50%) rather than a learning outcome suggests that this study may be a first to provide some empirical evidence that some lecturers construe the skills and knowledge students ought to have in terms of content rather than outcomes (Hussey and Smith, 2002).

3.6.2 Implications of pilot study

Intended curricula can only be framed in general terms rather than be pre-specified, before any meaningful teaching and student learning have taken place (Maher, 2004). These findings and previous research into intended curricula in higher education (Hussey and Smith, 2002, 2003; O’Donovan et al., 2004; Maher, 2004) suggest that module specifications cannot communicate how to teach or how well students should perform (Raymond and Neustel, 2006). However, the challenge with specifying curricula does not mean that it should be left tacit since it is an essential part of effective learning, teaching and assessment (Gal et al., 1997).

In my main study therefore, I aim to carry out observations of teaching activities and semi-structured interviews with lectures regarding the module intended and implemented curricula. Amongst other aspects, observations of teaching could be used to confirm whether it is the case that lecturers consistently focus on content rather than learning outcomes.

With the interviews I carried out with lecturers, I identified examples of the underlying knowledge, skills and understanding students need on statistical modules such as the importance of statistical techniques (e.g. descriptive and inferential statistics), numeracy, using resources and communication skills (e.g. report writing) to the statistical curriculum. In terms of the general principles involved in tackling statistical problems set in a real-life context, the participants appeared to focus on statistical analysis and interpretation of results more than on formulating research questions and collecting data.

In general, the pilot interviews with lecturers highlighted the concern with combining (often established) knowledge and understanding of statistical theory with applying statistical methods and techniques to solving realistic problems (theme 1). Especially relevant in the age when statistical software is widely available is the question of how much and what theoretical understanding does one need to be able to run a statistical analysis to solve a problem. Some of these lecturers were part of a module team that was involved in designing the module specification within institutional and regulatory frameworks, guided by their own beliefs about what statistics is and what students should learn during their studies. In other cases, lecturers taught existing materials, guided by their own beliefs about statistics and about the students.

In the pilot interviews, the lecturers I talked to were also concerned with how much statistical knowledge (e.g. conceptual knowledge of statistical tests) and non-statistical knowledge (e.g. communication skills, team work, critical thinking, creativity) students needed to be successful in statistics (theme 1.2). Another concern for these lecturers was which statistical processes students needed to experience (e.g. formulating a research question, collecting data, analysing and interpreting it) versus what was possible to experience in a few weeks during a module (e.g. no opportunity for data collection at the introductory levels due to time constraints) (theme 2). Such issues contribute to the so-called *official, intended student learning outcomes (curricula).* So, at the macro level, it is important to consider the links between *the statistical practice processes* and *the teaching practice processes* as realised in the *intended student learning outcomes* for a module.

The in-depth analysis of the qualitative data of the interviews revealed a number of tensions in the teaching of statistics stemming from conventions of how statistics can be applied in the context of a particular field of study, the varied importance of statistics to different disciplines, achieving a balance between theoretical knowledge (statistical) and practical applications (non-statistical) and the challenge of representing a real-world context in a mathematical framework. Further tensions derived from whether students can do more than learn basic principles, follow recipes versus gain the ability to think about the world in a mathematical way, and know what the recipes are made out of. For the interviewees, the challenge was to integrate theoretical knowledge and practical applications.

However, the interviewees did not agree on how to integrate theory (abstract concepts without context) and practice (applications of statistics in a real-world context). Do students require the language of mathematics or only concepts taught within a context, do they need to be able to decide what statistical test to apply or be told what test to apply and finally should they be able to get the real world into mathematics or provide a narrative account of the analysis using specific conventions for communicating statistics without mathematics?

There were some indications that these tensions or contrasts in the statistics classroom were also reflected in lecturers' approaches to teaching. It seemed that Participant#1 viewed the didactic role of the lecturer in providing concrete examples of psychological constructs in a context then statistical objects, while Participant#12 suggests that his teaching involves presentation of abstract statistical concepts, discussion of general ideas followed by a (large) number of concrete, practical examples in context.

However, from this interview data, I could not gauge the nature of the contexts, examples or exercises these lecturers were using in their teaching or how they realised or implemented statistical applications and theory considering the institutional, didactic, pedagogic and curricular constraints that surrounded their modules. This pilot study helped me identify aspects of characterisation of context that merit further investigation: (1) the teaching processes and practices (curricular, didactic, pedagogic and institutional) and (2) the statistical process and practices (applications, analysis and theory). Further data gathered from observations of actual teaching on statistical modules and open ended interviews with the lecturers teaching on those modules could help me determine what characterises the teaching statistics at university in different contexts.

3.7 Summary of chapter

In my pilot study, twenty lecturers were able to described the statistics curricula they believed were important for students to learn on introductory modules. I was also able to gain insights into teaching processes and practices through close observations of two lectures and two laboratory sessions. The findings from the pilot interviews data and the conversations I had with each participant allowed me to gain insights into the possible contextual influences from outside the classroom, at the macro-level of analysis. Since the repertory grid technique allows participants from different contexts to describe statistics education in terms of contrasts, I also regarded this dataset useful for characterising tensions and contradictions in the teaching of statistics at university. Importantly however, the themes emerging from the qualitative/quantitative analysis formed the basis for producing a 'rough' first draft of a conceptual framework in Figure 3.3. This conceptual framework was important in guiding my research design (in the next chapter in Section 4.3) and in particular refining my research questions, data analysis and interpretations of my main study.

4 | Main Study Methodology

Chapter contents

In the previous chapter, 3, I describe details of my Pilot study methodology, analysis and findings. The pilot study was useful in providing rich insights into lecturers' beliefs about intended curricula on statistical models. The output of the pilot analysis was an initial conceptual framework containing a set of components which became the building blocks of my main study. In this chapter, I discuss the research paradigm, theoretical perspectives and methodology of my main study, a qualitative enquiry into lecturers' teaching practice. Figure 4.1 represents in brief the relationships among the main factors influencing the main study methodology.

Methodology is a description of *methods*, "how" the research was done and also about "what influenced the researcher to choose to do the research in the manner described" (Burton, 2002, p. 1). In this chapter, I discuss "why" I chose to carry out my research in the way that I did and what influenced my assumptions regarding my *methods* of data collection and analysis, my *research paradigm*. Further, I also look at what theories inform the choice

Pilot study *(Chapter 3)*

$\updownarrow$

Research paradigm $\longleftrightarrow$ ***Theoretical perspectives*** $\longleftrightarrow$ ***Research design***

$\updownarrow$

Main study data analysis and findings *(Chapter 5)*

$\updownarrow$

Summary and Conclusions *(Chapter 6)*

Figure 4.1: Three factors influencing the methodological process

or research topic, research questions or data collection methods, the *theoretical perspectives*. Finally, I outline "how" I carried out the research, the *research design* of my main study.

During the data collection and data analysis processes, "why" I made particular choices was intertwined with "how" I carried out the research (as in Figure 4.1). In other words, the research paradigm and theoretical perspectives are intrinsically linked to my research design.

4.1 Research paradigm

In this thesis, I sought to *characterise* lecturing as teaching practice on statistical modules with the broad aim to explore current ideas and practices in statistical education at undergraduate university level. My study aimed to explain the teaching of undergraduate 'in-service' introductory statistical modules from the perspective of the lecturers. I was in particular interested in the insights that I could gain about the teaching of statistics at university level by studying the way lecturers taught in the lecture theatre.

With this topic to study in mind, I turned to considering how to go about investigating it, crystallise my beliefs, assumptions about the context I was about to investigate, values about the aim of my inquiry, methods and so forth, what Schwandt (2001) called a **methodological or research paradigm**. Guba and Lincoln (1994) define a paradigm as

> *"the basic belief system or world-view that guides the investigator, not only in choices of method but in ontologically and epistemologically fundamental ways."* (p. 105)

Chilisa and Kawulich (2012) propose that the methodological process is informed and influenced by three factors (1) the research paradigm, (2) the theoretical perspectives, literature and research practice and (3) the value systems and ethical principles (axiology). Thus, questions of paradigm need to precede questions of method or methodology (Figure 4.1). In this section therefore, I discuss my research paradigm followed by the theoretical perspectives in Section 4.2.

Particular paradigms have been described in several ways, depending on the nature of reality, what counts as knowledge, values and researcher's role in the research process (Chilisa and Kawulich, 2012). Two common examples of paradigms are *positivist* and *interpretative* (Charmaz, 2006).

Since the main reason for doing the research study was to *understand*, *describe* and *interpret* human behaviour, I placed myself within the **interpretative paradigm** which is assumed to be compatible with qualitative, ethnographic or naturalistic methodology (Creswell, 2003).

In my thesis, I sought to actively form empirical insights about teaching practice on statistical modules as a participant-as-observer (Gold, 1958). My aim was to *observe* and *interpret* lecturers' behaviour before, during and after teaching in order to *understand* and *explain* the **context of teaching** and link it to the **teaching process** (concepts discussed in more detail in Section 4.2). Thus the interpretative paradigm was appropriate since the focus was on the interpretations or meanings associated with human behaviour.

The interpretative paradigm sees the nature of **reality** (ontology) as socially constructed, depending on a personal or social construct (Guba and Lincoln, 1994). There are multiple, intangible, social realities which depend on an individual or group's experiences of the world. As more is known, the assumption is that the investigation will diverge rather than converge. Thus, although aspects of reality are inter-related and can be shared among individuals, groups and across cultures, reality may be local and specific in nature (Guba, 1981).

Corbin and Strauss (2008) propose a paradigm for identifying contextual factors and linking them to process. **Contextual factors** are conditions that shape the nature of situations/ interactions or problems to which individuals or groups respond with actions/interactions or emotions. *Process* is about the flow of such actions/ interactions or emotions that occurs in response to events, situations or problems. A change in context may lead to a change in process. Hence interpretative researchers often address the *context* and *process* of 'interaction' among individuals or groups (Creswell, 2003).

In this view, the research design (developed in Section 4.3) would involve making interpretations of *conditions* - why, where, when, how, what happens, *actions/ interactions/ emotions* - the responses made by individuals or groups in response to situations, problems or events and *consequences* - what happens as a result of the inter/actions or events.

The **knowledge** (epistemology) that develops through an interpretative lens is based on the assumption that humans develop *subjective* meanings of their experiences. This raises the issue of "objectivity" in the data collection and analysis (Guba and Lincoln, 1994). As an interpretative researcher, I bring to the research situation my own paradigms, perspectives, training, knowledge and biases. The choice of paradigm, topic for study, methods of data collection and analysis and how I interpret and report the findings of my study are value-laden in nature. My role as a researcher is to look for multiple meanings and views which are varied and complex with many categories, rather than narrow, with few categories. To create meaning, I can start from a broad and general question which I can refine through discussions, interactions or negotiations with others in the social context.

In my study, I negotiated meanings with the participants in interviews and other social interactions, e.g. during observations of teaching (before and after a lecture, during breaks or in workshops, with my supervisors in meetings or with other researchers when presenting, discussing and sharing my research. The way I understood and made sense of the participants' actions and beliefs was part of constructing an interpretative account of the teaching of statistics I focused on in this study. Yet, could my research study have validity and/or replicability (reliability)?

As discussed earlier, for Corbin and Strauss (2008) a change in the context leads to a change in the process which means that a contextual change can affect my interpretations and the meanings I am making about my observations. Therefore, within an interpretative paradigm, a study would not be **repeatable** under the same contextual conditions twice and so the interpretations (results) would not be replicated in the same way on multiple occasions.

The notion of *truth* or **validity** would depend on context, space, time, individuals or groups since knowledge is subjective, relating to unique facts or events that cannot be generalised into common reality. Instead, "working hypotheses" relate to a particular context (Guba, 1981). For example, teaching experiences, narratives, beliefs and claims about the context of teaching can be seen as legitimate knowledge. Although researchers within the interpretative paradigm such as grounded theorists may not embrace quality criteria of validity and reliability (Corbin and Strauss, 2008; Charmaz, 2014), it is important to consider whether the research design leads to interpretations which can be *trusted*. It is essential to be aware when/whether the quality of the research activity is threatened in order to avoid any pitfalls (Corbin and Strauss, 1990).

Instead, qualitative researchers offer alternative criteria for judging the quality or credibility of an inquiry (Guba and Lincoln, 1994). For example, trustworthiness, credibility, transferability, confirmability or authenticity are such criteria. Guba (1981) proposes that the **trustworthiness** of an inquiry can be affected by four factors:

- *Factor patternings* which produce effects of non-interpretability instead of credible actions and plausible findings.
- *Situational uniqueness* which lead to non-comparability instead of transferability and context-relevant findings.
- *Instrumental changes* which produce instability instead of dependability and stable findings.
- *Investigator predilections* which lead to bias instead of confirmability and investigator-free findings.

In an interpretative inquiry, researchers adopt certain procedures to preserve the quality of a study. For example, to preserve the **credibility** of a study, researchers use prolonged and persistent engagement and observation. Guba (1981) suggests that the researcher does not make sense of the reality he/she is observing in isolation. During and after the data collection, in order to preserve **credibility** of the research process which lead to finding plausible patterns in the data, the researcher could use peers (e.g. lecturers, supervisors or other researchers) as a sounding board of the outcomes of a study and triangulate the data (using more than one method to collect data on the same topic). **Transferable**, context-relevant findings require the collection of thick data that focuses on a small number of cases but on a large number of variables, theoretical sampling and the development of thick descriptions which capture several, alternative points of view (Guba and Lincoln, 1994) arrived at by my submersion in the setting of the study, in conversations with lecturers, interviews and formal observations (Miles et al., 2013).

Further, instead of reliability, the interpretative researcher is interested in **dependability** or stability of findings through the use of overlap methods and a systematic approach which is transparent and communicable. Such a research design involves research questions that are clearly linked to methods of data collection and analysis and to findings (Miles et al., 2013).

Confirmability, objectivity or investigator-free findings are possible through triangulation and procedures that are made apparent. Further, the findings and conclusions are explicitly connected to the data, while other competing conclusions have been considered or the methods and procedures are detailed and can be audited.

Finally, instead of validity, truth value in qualitative research is represented by **authenticity**. Issues of authenticity are represented in descriptions which are context-rich, meaningful and "thick". The key to demonstrating authenticity is for the researcher to provide an account which is coherent and systematic. In other words, the researcher can present a trail from research question, research design and findings. For Corbin and Strauss (2008), good quality research is interesting, logical and makes sense to the reader in stating that "quality qualitative research resonates with the readers and participants' life experiences" (p. 301).

In the next section, I present my theoretical perspectives, the theories that inform the choices I made later in my research design and the literature review. The ways in which I view reality, how I come to construct knowledge, along with the theoretical perspectives I hold about the teaching of statistics, together with the literature reviewed in Chapter 2 and my own value system discussed in this section are intended to strengthen the quality of my findings.

4.2 Theoretical perspectives

Theoretical perspectives, frameworks or approaches can provide the conceptual guide for choosing the concepts to be investigated, what theories and beliefs guide or inform the research, suggesting research questions and for framing the findings (Maxwell, 2012). A theoretical perspective is the philosophical world-view informing the methodology and is different from methodology, which is about the process, strategy or design of a study. In this section, I aim to reflect on the design and implementation of my research in order to argue "why" I made particular choices. My purpose is to further frame and structure my research study.

Glaser and Strauss (1967) propose that theory is a "strategy for handling data in research, providing modes of conceptualisation for describing and explaining" (p. 3). Thus, theory can be *descriptive* and *prescriptive.* A theory of teaching can *describe* how teaching occurs, a process as it *does* go on and also *prescribe* how it *should* occur to optimise student achievement (Gage, 2009). However, the distinctions between prescriptive and descriptive theories of teaching are rather blurred, since the process of teaching relates to student understanding. Further, a theory of teaching is assumed to apply to several aspects of teaching, including subject matter, school level, types of students, educational objectives, schools, cultures and geographic areas.

Thus, in this section, I summarise the topics I focused on to arrive at an understanding of a model of teaching, of the content of teaching and of teaching process. In my research I differentiated between a "model" and a "theory". Using analogies from science, Schoenfeld (1998) makes the point that a model and a theory it embodies "do not represent absolute claims to **truth**; they are working descriptions that help us grapple with complex phenomena" (p. 14). The models built should be constantly tested against *reality* and judged by "ways they [the models] support predictions, by their explanatory power, and by their scope", the number of cases to which a model applies.

For my theoretical framework, I draw on two rather different descriptive perspectives: Vygotskian sociocultural theory (SCT) of developmental teaching and a theory of classroom teaching developed by Nathaniel Gage (Gage, 2009). In my research, SCT perspectives on teaching and learning emerged towards the end of my Pilot study when I attended a number of seminars lead by Stephen Lerman (e.g. Lerman, 2001). The complexities and opportunities offered by SCT approaches appeared to fit in with my focus on teaching processes and context. Further, I describe Gage's theory of *classroom* teaching. Considering my interest in direct observations of teaching, I found that this model, although grounded in previous educational research studies not based on SCT perspectives on teaching, to be a useful *tool* for conceptualising some of the aspects I was concerned with in my research.

I discuss these two perspectives in turn in Section 4.2.1 and Section 4.2.2 respectively. In my research design, I adopt a grounded analytical approach with a different theoretical frame, namely SCT principles on developmental teaching and learning. I justify the use of a grounded approach within a SCT frame in Section 4.2.3.

4.2.1 Vygotskian sociocultural theory of teaching and learning

Sociocultural or sociohistorical theories of teaching and learning have as basis the writings of Vygotsky (1987, 1978b), Leontiev (1972, 1978) and Luria (1928) and more recently researchers such as Wertsch, Rogoff, Cole, Engeström, Lave and Lemke (Lemke, 1990; Lave and Wenger, 1991; Cole, 1996; Wertsch, 1998). SCT is a *theoretical* and *methodological* tool which can support the modelling of social, cultural and historical processes that influence human functioning rather than offering a practical approach to teaching.

My theoretical approach needs to be sensitive to the experience of teaching statistics at university from the point of view of the lecturers. SCT encourages a focus on teaching, learning and its context. Consistent with my theoretical framework based on the ideas of Vygotsky, I explore the network of relationships between lecturers, statistics as the subject matter and context. The grounded approach to analyses of data such as observations of teaching, interviews and documentation provides the means for investigating the specific context and circumstances in which my research took place. My application of the theoretical framework based on Vygotsky's SCT provides insights that go beyond the immediate context and circumstances in which I carried out this research study and invigorate my analysis.

Research in mathematics and statistics education emphasise how mediating cultural tools such as numeracy, mathematical challenge, questioning, technology, knowledge or expertise impact collective activity (Greeno, 1991; Engeström, 1995; Jaworski and Potari, 2009). In my focus on characterising teaching of statistics, the contributions of tools, context and social interactions are relevant to questions as to what statistics is important to learn and how it is taught at university. Following a discussion on **the relevance of SCT to adult education** in the next section (Section 4.2.1.1), in order to present a model of how teaching statistics at university is derived from pedagogic and statistical activity, relevant to my characterisation of teaching in context are therefore Vygotskian concepts of **teaching activity** (Section 4.2.1.2), including the concept of *relevance*, **mediation** (Section 4.2.1.3), **the zone of proximal development (ZPD)** in which learning occurs (Section 4.2.1.4), including the concepts of *development*, *contradictions* and *scaffolding* and **context** (Section 4.2.1.5), including the concepts of *context as physical environment*, *context as intellectual environment*, *scientific and everyday concepts*, *vertical and horizontal knowledge*, *representations*, and *macro and micro context*.

4.2.1.1 Relevance of sociocultural theory to adult education

In this section, I discuss how Vygotskian developmental psychology can be applied to the study of teaching and learning of statistics to adults at university. SCT studies have been concerned with examining children's development and so the relevance of SCT to adult education might not be straightforward. Vygotsky was concerned with studying children's enculturation into the values of their community over long periods of time, arguably much longer than learning statistics at university. For example, Vygotsky's (1978a) work on language development focused on two functions, the long and gradual differentiation between social contact or communication and representation which grows out of the social-communicative function. With this approach, research aims to understand how language and symbols mediate meaning, and thus **social interactions**, historical and cultural context have a fundamental role in teaching and learning activities of all learners, including adults (Vygotsky, 1978b).

In Vygotsky's approach, relevant in the context of teaching adult learners is the notion that external activity precedes individual internal activity. Vygotsky (1981) recognised the differences in cognitive responses to stimuli between child and adult subjects and made the assumption that adults are fully developed organisms. In a biological sense, that adults are fully developed beings is a perfectly reasonable assumption. However, in SCT, culture plays a central role in the formation of human thinking. It is reasonable to conclude then that **development** (Section 4.2.1.4) could and does occur throughout the lifespan as adults encounter new forms of mediation (Lantolf and Thorne, 2007).

Two projects carried out by Luria (1976) I believe epitomise the relationship amongst mind, culture and society: the cross-cultural work in Central Asia and relating psychological processes and brain function. Both projects were based on the underlying idea that brain function depends on and develops through interaction with a culturally organised environment (Wells, 2007). In the 1930s, in collaboration with Vygotsky, Luria planned and carried out two cross-cultural experiments with groups of adults in Uzbekistan and Khirgizia, ranging from illiterate peasants to trainee teachers. The project investigated whether development in the new socialist economy with access to schooling has an effect on the content and form of adults' thinking (Cole, 1996). Luria (1976) wrote that

> "*when [the subjects] acquire some education and participate in collective discussions of vital social issues, they can readily make the transition to abstract thinking. The acquisition of new experiences and new ideas impart added meaning to their use of language so that words become the principal agent of abstraction and generalisation. At this point, people dispense with graphic thinking and codify.*" (p. 99)

In spite of criticisms of the research carried out by Luria, his findings supported the initial hypothesis that since higher mental functions have their origins in the activities of particular cultures, they also differ depending on the culture to which individuals belong to (Cole, 1996). Luria and Vygotsky's experiments with adults looked to demonstrate how "man adapts himself more perfectly to the surrounding world" and that "the tools used by man not only radically change his conditions of existence, they even react on him in that they effect a change in him and in his psychic condition" (Luria, 1928, p. 493). Engeström (1996) for example explored work of general practitioners in a health centre as collective, institutionally organised **activity** (Section 4.2.1.2), which is not usually conceptualised as learning.

After Vygotsky's death, Luria studied medicine and turned his attention to the study of the relationships between brain function and cognitive activity. This work went back to Luria's earlier interest in children's development and aphasics he had done with Vygotsky (2012). Luria gave an account of *neuroplasticity* which refers to changes in brain function and structure that can arise in different contexts. By introducing concepts such as *neuroplasticity*, Luria promoted the idea that the higher mental functions of the mature brain mediated by cultural systems of signs are capable of change and transformation. Luria's study of brain function originates in and supports Vygotsky's (2012) work on the relation between thought and speech.

Teacher development is another example of a research strand that has been conceptualised through SCT. For instance, a number of studies have conceptualised development from SCT approaches, including teacher development mediated by cultural tools, the ZPD in which teacher development occurs and scientific and spontaneous concepts used to analyse adults' teaching or other work practices.

For example, Gordon and Fittler (2004) conceptualised **teacher development** through the concept of teaching activity and a focus on tensions or **contradictions** (Section 4.2.1.4) that stimulate development, such as the tension between adopting teacher-centred and student-centred teaching approaches (Section 2.5). Gordon and Fittler (2004) thus conceptualise the teaching activity as a 'complex and dynamic phenomenon'. The teaching activity took various forms, which lead to tensions between these different forms of activity and subsequently to development and improved capacity to support student learning. In SCT approaches, the development (e.g. of a teacher) is **mediated** by cultural tools (Section 4.2.1.3).

In another study, Goos (2005) and Blanton et al. (2005) use Vygotsky's **ZPD** (Section 4.2.1.4) and Valsiner's (1987) extension to the zone of free movement and the zone of promoted action to characterise the development of teaching and teachers. With these concepts, Goos (2005) sought to identify possible relationships between the context, actions and student teachers' beliefs, and how these relationships might change over time or across different classroom or school contexts.

In a case study focusing on the transformation of a teacher from pre-service, novice teaching into expert teaching, Au (1990) uses observations of teaching and interviews to capture the relationships between the development of the teacher's knowledge of practice and changes in her inner thinking. Au's (1990) study shows how Vygotsky's **scientific and spontaneous concepts** (p. 122), the role of speech in thinking and the development of consciousness, although usually used to analyse teaching aimed at children, can also provide useful insights into the complexities of teaching aimed at adults while still providing some general principles of development. Engeström (2000) also applied Vygotsky's (2012) developmental learning as a creation of scientific and everyday concept formation in the case of junior physicians' work practice.

To clarify the relationship between teaching and development, Vygotsky (1978b) compared first and second language learning. In first language learning, children acquire their native language spontaneously as they engage in everyday social activity. Unlike first language acquisition, second language acquisition happens consciously. In this way, Vygotsky explained the difference between unconscious, spontaneous knowledge and conscious, scientific knowledge in different **contexts** (Section 4.2.1.5).

Based on Vygotsky's approaches to second language learning, Washburn (1994) noted that adult teachers work in the learners' ZPD since the data showed that the subject could not produce a sentence before the teaching intervention (consisting of a variety of teaching strategies) but could do so after the intervention. Lantolf and Thorne (2006) also applied spontaneous and scientific concepts to adult learners to show that first language mediates second language learning and that Vygotskian concepts can be fruitfully applied in investigations with adult participants. However, the range of SCT applications reviewed in this section have not been applied to investigating the teaching of statistics at university from the lecturers' perspective. In the rest of this section, I discuss relevant concepts to my theoretical framework, such as teaching activity, mediation, ZPD and context.

4.2.1.2 The concept of teaching activity

In my approach to SCT, the emphases are on the teaching actions and activities as mediated by tools within the various contexts in which they occur (Wertsch, 1995). In this section, I discuss definitions of teaching (pedagogy and didactics) from the SCT lens, followed by a discussion on mediation in the following section. I discuss mediated activity as the unit of analysis in Section 4.7.2.2.

Pedagogy in SCT is a fundamental human activity and a central concept in Vygotsky's approach (Moll, 1990). Pedagogy (androgogy for adult learning), as a term, within SCT, is interpreted as referring to "forms of social practice which shape and form the cognitive, affective and moral development of individuals" (Daniels, 2001, p. 1). Further, teachers' pedagogy is about making decisions, negotiating between different sources of experiences, judgements and understandings of the activity within a *context* (in 4.2.1.5). A teacher's craft is the repertoire of skills, strategies, methods, approaches and practices. My interpretation from a SCT lens is that the craft of teaching statistics is not static, it changes continuously in response to classroom experiences. *Teaching and learning processes* are more than *interactions* between individuals or *transmission* of knowledge and skills. Daniels (2001, p. 4) makes the point that during the course of their own development, individuals "actively shape the forces that are active in shaping them".

In Bernstein's (2000) social theory, the challange was to connect practices within and between macro and micro contexts in the shaping of consciousness when he defines pedagogy as

> "a sustained process whereby somebody acquires new forms or develops existing forms of conduct, knowledge, practice and criteria, from somebody(s) or something deemed to be an appropriate provider and evaluator – appropriate from the point of view of the acquirer or by some other body(s) or both" (p. 78)

In this definition, institutional rules, discursive practices and even policymakers are involved in pedagogic practices (Daniels, 2015). However, Bernstein (2000) distinguishes between *institutional* pedagogy carried out in over a long period of time in formal settings and *segmented* or informal pedagogy carried out in everyday experiences and practices by informal providers. SCT provides a compatible account that emphasises individual agency through the concept of mediation (Section 4.2.1.3). Using both views of pedagogy can facilitate addressing relationships between macro (wider) and micro (local) contexts of activity. I discuss further Vygotskian and Bernsteinian theories in Section 4.2.1.5).

Didactics on the other hand studies the relationship between students, teachers and the subject-matter content (Holmqvist Olander, 2016). The aim of didactics is to plan and organise learning experiences, with the teacher being a necessary component of the learning process. Thus pedagogy is about the philosophy, sociology or psychology of education, while didactics is a form of specialised, domain-specific pedagogy. The focus in pedagogy is on different types of teachers, contexts of teaching and teaching and learning processes (Watkins and Mortimore, 1999). SCT emphasise the role of teaching in extending knowledge through pedagogic practice, understood as "a fundamental social context through which cultural reproduction-production takes place" (Daniels, 2001, p. 20). The teaching and learning environment is conceptualised to be a "shared problem space" in which students co-construct knowledge under the teacher's guidance (Haenen et al., 2003).

For instance, Vygotsky (1978b) criticises the teaching of writing as artificial, self-contained training when not founded on the needs of the child as they naturally develop and on children's own activity: "given from without, from the teacher's hand". Vygotsky further states that learning music should not be about learning to strike keys while reading music but about the essence of music itself. Rather than being taught as a motor skill, the teaching of writing or music should be taught as systems of symbols and signs whose mastery affect a critical turning point in the cultural development of the child. The teacher's role is to help the child discover that one can draw speech, or to shift the child's activity from drawing things (e.g. letters) to drawing speech/words (the meaning of letters).

This view of teaching and learning as affecting a child's cultural development could be extended to adult learners. Gordon (1993) used SCT as a conceptual model to investigate the complexity of learning statistics as mature students at university and the role of teaching in learning. In Gordon's (1993) study, students' approaches to learning statistics were determined by their history of learning experiences, by their beliefs about mathematics and statistics and by the teaching and learning context or setting. In this study, SCT is useful in uncovering the complex processes of learning statistics as an in-service subject. SCT promotes a unified historical line from a child's make-believe play, drawing and writing. In Gordon's (1993) analysis, students' learning approaches were influenced by their prior experiences with learning school mathematics.

Relevance

From a Vygotskian point of view, reading and writing need to be **relevant** or 'necessary for something', otherwise they are mechanical and boring (Vygotsky, 1978b). The mature students in Gordon's (1993) study perceived that the main purpose of studying statistics at university is to recall information for an assessment and as irrelevant to their future careers. One implication for teaching statistics from a SCT perspective is that statistics needs to be taught "naturally" in order to be relevant and meaningful to students. Natural methods of teaching reading and writing involve appropriate operations on the child's environment. For Vygotsky (1978b), teaching needs to bring the child to an 'inner understanding' of writing, to arrange that writing as 'organised development' rather than learning. Drawing and playing are considered to be preparatory stages in a child's development of written language. A teacher's role is "to organise all the preparatory stages or actions and the entire complex process of transition from one mode of written language to another" (p. 118).

The teaching of statistics is a complex cultural activity which needs to be taught 'naturally' and 'relevant' to students' lives. From this perspective, Gordon (1993) defines the teaching of statistics as

> "*The guiding of students to view their mathematical learning as relevant and meaningful, rather than as the process of transferring a body of mathematical knowledge from teacher to student.*" (p. 45)

One of the challenges experienced by students in learning statistics is to know how to make formal or generalised knowledge of statistics relevant to current and future contexts (see Gordon, 1993, Section 2.6.4). Further, statisticians use formal knowledge to solve everyday problems. These actions and interactions across contexts (e.g. from school to work) affect not only the individual but also the different social practices in general (Akkerman and Bakker, 2011). Bakker et al. (2008) for example discuss how statistical inferential reasoning can be different when applied at school or in the workplace. In the workplace, contextual reasons can overpower statistical reasons. At school however, statistical tests are used to test the accuracy of statements. Different uses and applications of statistical tests, what counts as evidence, inference, or conclusion are motivated by cultural norms and values. Such workplace knowledge might contrast with textbook knowledge since it depends on the context in which it is represented (Brown et al., 1989). Teachers might need to guide the students in gaining awareness of the relevance of and connect formal knowledge gained in lectures and other forms of social interactions at university to future informal, practical knowledge. Developmental teaching holds that formal teaching and development are interrelated and that when development follows formal teaching, it creates ZPDs. Following the concept of mediation in the next section, I discuss the Vygotskian concepts of development, learning and ZPD in Section 4.2.1.4.

4.2.1.3 The concept of mediation

The focal point in SCT is to provide and account for learning and development as *mediated* by indirect processes with an emphasis on *speech*[1], which is a form of behaviour. In a process of formation, the individual acts upon and is acted upon by *social, cultural and historical factors* by means of *mediators* (Cole, 1996). A critical issue then is "what is the pedagogy of such a process of formation?". The starting point in Vygotsky's theory is the idea of mediation of elementary mental functions by so-called tools (Vygotsky, 2012; Davydov and Radzikhovskii, 1985).

The basic mediational triangle in Figure 4.2 represents a Vygotskian view of an activity system. Based on Vygotsky's concept of mediation, the basic triangle of mediation represents the subject, object and artefact (or tool) relationships, bringing together the artefacts

[1]In activity theory, a tradition also linked to Vygotsky's work, the emphasis of analysis is on activity itself. In other approaches such as situated action models, the unit of analysis is everyday actions of persons in a setting.

with human actions in order to dispense with the individual/social dualism (Cole, 1996; Engeström, 1999). The triangle has the subject and object both *directly* connected to each other through a natural path, "unmediated" (at the base of the triangle). Subject and object are also *indirectly* connected through a medium of artefacts. The "cultural" or "mediated" functions are those where subject and object are linked through artefacts. Both mediated, cultural and unmediated, natural, elementary routes can operate 'synergistically'. In my study, lecturers represent the subject of activity and other people in the sociocultural settings such as the students and their statistical meanings as the object of activity (Engeström, 2000; Rezat and Sträßer, 2012) or the statistics modules and the students on it (Trowler and Knight, 2000).

In my work, I considered that even this basic notion of human thought as interweaving of direct and indirect, objective and subjective aspects of experience was sufficient since my focus was on individual lecturers. Thus, the complexities and opportunities offered by the subject-object-tool relationships captured the aims of my study in which the subject was actively engaged to establish the link between subject and object. In Chapter 5, I present in detail how I applied the concepts captured in the mediational triangle to my data.

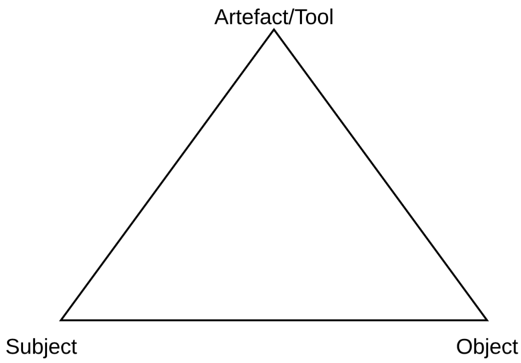


Figure 4.2: Mediational triangle

Language and signs were the main mediational tools in Vygotsky's work. Vygotsky described **psychological tools** as devices for mastering thought processes (Vygotsky, 2012). Psychological tools such as language, counting systems, mathematical symbols, works of art are artificial and a means of social communication with a definite meaning that has evolved in history and culture (Davydov and Radzikhovskii, 1985). **Material tools** serve as the mediator between the human hand and the object upon which the tools act (Kozulin, 2012). Tools are externally directed towards the object and subsequently become means of controlling one's own mental processes.

Cole adopts Wartofsky's (1973) definition of artefacts/tools for a more differentiated description on three levels. On the first level, *primary* artefacts are tools used in production, e.g. words, writing instruments, characters, and are transformed by human activity. *Secondary* artefacts are representations of primary artefacts and of modes of action using primary artefacts, such as beliefs, traditions, norms, recipes. *Tertiary* artefacts are imaginative artefacts, such as works of art. Tertiary artefacts are about the meaning of representation, about mediation, and secondary artefacts are reflexive and primary artefacts are qualitative (Seeger, 1998). Artefacts then can be material and also conceptual, affective or idealised representations. In my research, I conceptualise primary tools/artefacts to represent textbooks, examples used in lectures or university infrastructure such as software. Statistical models and formulae, teacher beliefs, curricula were part of secondary tools while lecturer narratives about statistics represented "imagined worlds" and idealised representations of statistics.

Vygotsky (1978a) was interested in the relationship between **development** and **learning**, with special emphasis on learning once children reach school age. For Vygotsky, teaching was not separated from learning; mediation by tools, peers and teacher comprise learning (Lerman, 1996). The pedagogic tools or strategies used by lecturers to help students' learning constitute a ZPD, which I discuss next (Moll, 1990). The use of units of analysis conceptualised in terms of the use of psychological tools in contexts raises questions of differences between contexts, such as differences in pedagogic practices which represent differences in contexts (Daniels, 2001). I consider the concept of context in Section 4.2.1.5.

4.2.1.4 The concept of zone of proximal development

A SCT perspective on *learning* is reflected in the general law of cultural development which considers that psychological functions (attention, memory, cognition) appear on two planes: first, on the social plane and then on the psychological plane. Vygotskian developmental psychology proposes two **developmental levels** or "zones": (1) actual development, what the child can do on their own or 'imitate' in collective activity and (2) potential development, what a child can do with adult guidance or in collaboration with (more capable) peers. In *Mind and Society*, Vygotsky (1978b) valued dynamic assessment of children's intellectual abilities instead of measures that are interpreted as static such as IQ scores. Vygotsky proposed the concept of ZPD as a metaphor to explain how social and participatory learning takes place. Vygotsky stated that

> "*the distance between the actual developmental level (of the child) as determined through problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers.*" (Vygotsky, 1978b, p. 85).

This definition highlights the difference between *supported* and *unsupported* performance. The implication of ZPD is that a novice learner requires some help in order to master a task or make progress. With ZPD, Vygotsky explains the key difference between learning before starting formal schooling (pre-school) and formal school learning. ZPD defines mental development prospectively as functions that are in the process of formation, maturation and development. The difference between "pre-school", everyday knowledge and "schooled", scientific knowledge (Section 4.2.1.1) can also be extended to learning across the lifespan (Lantolf and Thorne, 2006).

Development

SCT claims that development takes place under conditions of teaching, which in turn organises the concrete learning activity and its formation. Linear theories of development explain development in a pre-deterministic way made up of developmental sequences, leading to a fixed mature state. Children's development is an intramental (psychological) process as they move from unorganised heaps to complexes to concepts and this applies to development across lifespan. Intermental (social) ability occurs in the relationship between people. In this non-dualist conception of mind, mediated processes in which culturally produced tools, including forms of talk, representations (ideas, beliefs), signs and symbols shape and are shaped by human engagement with the world (Daniels, 2001). Vygotsky used the term "zone" rather than level to emphasise that development is a dynamic and continuous process. Instead of following a linear course or pre-determined course of events, the "zone" represents a phase in development where a person is only able to perform a task and internalise it with the help and supervision of someone more experienced (Cole, 1985; Tabach, 1999). In this respect, "zone" is the social context of learning in which tools are made available and meanings are shared between teachers (adults) and students (children). At the level of individual adult learner, development can be understood as "cycles" of cultural transformations or changes as a result of the interaction between active organism and active context (Tolman, 1999). Through these cycles, transformations or change, people acquire established culture and also formulate desirable culture. The focus is on processes of learning and development as well as its product.

From Vygotsky's writings on the teaching of writing, music or Latin grammar (Vygotsky, 1978b, 2012), it seems that many activities undertaken by children are akin to those undertaken by adults when learning a new skill. In Vygotsky's (1978b) view, school should make every effort to push children to develop what is "intrinsically lacking in their own development" (p. 88). Learning oriented towards developmental levels that have already been reached is seen as ineffective from the child's overall development and thus teaching needs to aim for a new stage of the development process, to be "ahead of development".

This means that all adults, in all cultures, have higher sign-mediated systems (e.g. memory, counting, writing) and that in order to understand the cultural development of the child, it is necessary to know the structure and function of higher systems, their origin and development to "full maturity and death" (Vygotsky, 1978b). In this stance, development is not stage-like, natural and a universal path towards maturity or growth. Development is characterised by a process in which new stages arise out of "an actual confrontation between the organism and the environment" (Stetskenko, 2011, p. 32), an adaptation of the organism to the environment. So Vygotsky advocated researching child development by understanding the origins of higher mental processes and uncovering their development through to their most advanced adult forms. From here arises the idea that people, including adults, develop knowledge based on life experiences, mediated by cultural tools and within a social and cultural context.

In SCT, the role of culture and the social context of the learner or teacher are important in constructing knowledge rather than the individual developmental change (Vygotsky, 1978b). The mechanisms of individual developmental change is rooted in society and culture. In this view, mind is a complex network of general capabilities, independent of the other and developed independently. As a teacher himself, Vygotsky (2012) proposed that teachers' beliefs and activity are influenced by their powers of observation, attention, memory, thinking. Further, a student's improvement in a specific ability will result in an improvement in all his/her general abilities. In his view, the study of a subject (e.g. Latin grammar) can improve a students' attention in all other subjects. The assumption is that mental capabilities function independently of the material within which they operate, and that the development of one ability entails the development of others. The child learns structural principles whose sphere of operation is other than just the operation since learning and development do not coincide. Davydov (2010) argued that

> "*Cycles of development always precede formal teaching cycles. Teaching must be at the tail-end of development and development always pushes formal teaching and learning ahead.*" (p. 12)

Teaching therefore can be structured in such as way to provide information and concrete skill and also promote cognitive development in the student[2]. However, Vygostky's view that there is such a thing as transfer (e.g. from scientific to everyday knolwedge, p. 122) has proven difficult to substantiate with empirical findings (van der Veer, 1998).

In the socio-cultural perspective, language and forms of understanding that are embedded in social practices and contexts are seen as important cultural resources available to learners. More experienced participants in the teaching activity such as lecturers have a key role in **mediating**, **scaffolding** (a form of teacher support or intervention for the initial performance of tasks discussed on p. 119) and extending learners' knowledge about the world. The teacher's role is seen as connecting learner's existing cultural and social understandings before further "deep" learning can take place. Learning, seen as conceptual change, implies that the student's mind possesses some *prior knowledge*.

The teaching and learning processes are considered to be "effective" if learners are aware from the very beginning of what some of the different aspects of the learning task are. Under the guidance of the teacher and mediated through a learning task (tool), learners are supposed to develop independent learning processes through their own activities. Vygotskian ideas focus therefore on the content of teaching where the task of conceptual development is intertwined with the task of developing students' conceptual reasoning in a given field

[2]Vygotsky's (2012) view that teaching needs to be ahead of development contrasts with practices which theorise that development is ahead of learning as in the case of Piaget or which theorise learning as development as in the case of Skinner.

(Kozulin, 1998). So, the teacher's role is to diagnose what the student already knows in order to improve student learning. Teacher's *guidance* is fundamental since the teacher presents the learning task and the related knowledge and skills to be learnt in a meaningful way ('as a meaningful whole').

The teacher is also responsible for stimulating and maintaining learners' attention and focus through the teaching and learning process. Since teaching needs to be ahead of development, knowledge is also expected to have affective, motivational and cognitive value (Vygotsky, 1987). From this perspective, teaching is developmental since it affects the *maturation* of a learner's (child) higher mental functions in a cooperative process. The learner develops or matures through adult intervention/assistance and participation. However, teachers cannot transmit knowledge directly. Concepts and meanings develop as a result of people interacting with the environment. In *Thinking and Speech*, Vygotsky (1987) gave an example of unsuccessful teaching (Tolstoy teaching literary language to children) to illustrate the difficulty of transmitting knowledge in a straight line. Vygotsky proposed that formal learning directly influences development. The process of concept formation through interaction with others and development leads to the formation of higher **developmental levels** when the method of teaching is assumed to be complex, subtle and indirect.

Contradictions

In teaching statistics, development can be understood by analysing data for **contradictions**, disruptions, tensions or innovations within concrete modes of the activity - before, during and after teaching. In Gordon and Fittler's (2004) study on teacher and teaching development, one tension was between the roles of novice and experience teacher. In the Pilot study data analysis Chapter 3, the RG interviews were designed to identify contradictions, contrasts or opposing ideas and descriptions of the obejcts that were compared.

However, Engeström (2001) identified contradictions as sources of change or development. The *contrasts* in the Pilot study RG interviews do not represent 'problems' or 'conflicts' and therefore they may be used in a further analysis using a SCT lens since they represent contradictions as perceived by the subject of the teaching activity. For example, when new students start a module, i.e. begin to enter a teaching activity system, the lecturers' tools for teaching may become inadequate. A contradiction can arise between the new object and the subject's previous teaching activity. These contradictions of the activity are the catalysts for change and development of the activity system (Engeström, 1996).

In the Vygotskian perspective, learning, as a higher form of human mental activity, is a socially mediated process influenced by tools, the most important of which is language. Further, learning is a process of internalisation in which skills and knowledge are transformed from the social to the cognitive plane. Children actively appropriate tools to their own uses rather than passively accepting them (Cole, 1996). The main assumptions are therefore that learning precedes development and that mediation, through language as the main tool, is central to learning. Adult students learning statistics require both concrete and abstract thinking and lecturers need to push students to develop what is intrinsically lacking in their own development. In my analysis, I was interested in capturing lecturers' teaching methods for structuring the social interactions, leading the students through the steps of a task and how culture influences the selection of learning contexts, the tools they use to mediate learning and development.

Scaffolding

With ZPD, Vygotsky explained how people make progress through formal teaching. Lave and Wenger (1991) distinguish between a 'scaffolding', 'cultural', 'collectivist' or 'societal' formulations of the ZPD. In my study, I considered that the lecturer needs to diagnose what is lacking and provide guidance to advance development. This help offered by a teacher has been called **scaffolding**, which can involve a variety of teaching methods used to move students progressively towards improved understanding by "making connections to what they already know". In her study on adult students learning to weave, Greenfield (1984) noted that scaffolding was an important theme in analysing learning to weave under the supervision of an expert weaver. The novice weavers were able to complete a piece of woven material indistinguishable by that produced by more experienced weavers only with the help of a teacher. In this context, Greenfield conceptualised that teachers scaffolded the learning by operating above the actual developmental level, in the learner's ZPD through joint activity. A 'scaffolding' interpretation of ZPD is therefore relevant in the context of lecturing since the term scaffolding could be taken to infer a 'one-way' process whereby the lecturer constructs the scaffold alone and presents it for use to the novice students.

In Greenfield's (1984) study, the teacher's involvement in the weaving process (e.g. taking over difficult technical parts of the process) functioned as a scaffold in helping the learner to complete the task. Teacher's activity during the weaving process observed by the learner is a further learning opportunity and internalisation of the teacher's skill by the learner. Similarly, I considered that lecturers' discourse in micro-teaching in the lecture theatre, could function as a scaffold for the students. In their discourse, lecturers may be able to demonstrate sensitivity to the students' level of understanding on statistical or context knowledge and support what the students can already do. However, lecturers' scaffolding might not be individualised and differentiated in the same way with small versus large student groups. Vygotskian theories stress the fundamental role of *social interaction* in learning and the development of cognition, with ZPD as the primary activity space in which learning occurs. Lerman (2001) suggests that ZPD is a symbolic space involving individuals, practices and their activity. In my understanding, ZPD is the *difference* between a learner's present developmental age and knowledge and what the learner can achieve under a teacher's guidance. ZPD defines those functions which are in an embryonic state, that will mature tomorrow or are in the process of maturation. ZPD therefore can account for cycles and maturational processes that have been completed but also those currently in a state of formation.

Interpretations of Vygotskian theories of teaching and learning emphasise a contextualised and social approach to education which considers that psychological function is *socially distributed* in a community of learners (Kozulin, 1986; Wertsch, 1988; Cole and Engeström, 1993). Within this learning environment, the teacher's role is to organise teaching resources (tasks). Tasks and teacher explanations function as 'scaffolds' for learner sense making and construction. The teacher and the learner negotiate, exchange and **co-construct** a concept's essence. In Vygotsky's (1987) words,

> "*It is not things or reality that push the child's mind along the path of development. Reality is itself processed and transformed by the mind. Left to itself, the child would achieve the development of nothing but gibberish. Reality would never teach him logic*" (p. 82)."

In the context of teaching of adults, as is the case at university level, teaching represents the means through which learning and development is advanced. Vygotsky criticises teaching that lags behind developmental levels or processes instead of focusing on emerging functions and capabilities. This implies that teaching methods need to adapt to the historical and cultural setting in which students live.

My interpretation of teaching from a SCT lens is about coaching or guiding a learner's acquisition of knowledge in context. Learning is a process of enculturation, adopting the language, behaviour and norms of a social group, becoming members of that culture (Brown et al., 1989). The view is that for students to learn domain-specific concepts, students need more than scientific, abstract concepts and self-contained learning tasks (Brown et al., 1989, also Section 4.2.1.5). Teachers need to show learners how to use a domain's conceptual tools in *authentic* activity.

A teacher's role is to act as a practitioner does in using concepts to solve real-world problems. In my understanding, a statistics teacher or lecturer needs to use learning activities to solve important problems. SCT perspectives would not necessarily exclude textbook examples and teacher-lead explanations. The issue for teaching is to distinguish between school learning activities and authentic activities. Within SCT, tools are formed and shaped through cultural-historical processes and in turn tools shape the individuals who use them to act on the world. SCT therefore can help understand the ways in which human action shapes and is shaped by the contexts in which it takes place (Daniels, 2015). The final part of my treatment of Vygotsian SCT considers context as a multifaceted concept.

4.2.1.5 The concept of context

Social, cultural and historical factors, conceptualised as **context**, are critical for the development of individuals since in the Vygotskian view, all knowledge is gained through learning. According to Cole (1996), individuals are active agents in their own development but do not act in a context (setting) entirely by choice. Human activity and action are mediated in context. Following Vygotsky's distinction between sense and meaning, Cole and Gajdamaschko (2010) conceptualise two notions of context as a situation or environment, referring to a physical environment or set of circumstances and a broader relational construct between the individual and the social situation of development which "weave" together. Also important in my analysis of 'context' are Vygotskian scientific and everyday concepts, Bernteinian vertical and horizontal knowledge/discourse, representations and macro and micro contexts.

Context as physical environment

Educational research studies have conceptualised context as the physical and socio-economic environment, understood as a situation, set of circumstances within which students interact with the tools and people in different ways. Such research might focus on influences in the environment, such as community or classroom contexts, on the teaching or on a student's cognitive and non-cognitive skills and development. In Chapter 2, I discuss some of the literature on factors that influence teaching practice, including the nature of statistics, lecturers' beliefs about a range of factors including statistics, teaching and students. Corbin and Strauss (1998) created a coding device (the consequential/conditional matrix in Figure 1.1, p. 3) to show intersections of macro/micro conditions and consequences on actions. The device was intended to help theoretical sampling (Section 4.2.3) decisions by locating the contexts in which the conditions occur and the connections and paths between them. In Section 4.2.2.2, I further show how Gage offers a 'nested' representation of context variables that influence the teaching process. I found these two representations of "context" useful in moving the data analysis forward in a pre-established direction by highlighting the need to map conditions, contexts, consequences and actions (the basic components of separate human activities) to the data.

Cole and Gajdamaschko (2010) also describe context as a nested set of socio-ecological arrangements with various inter-relationships in 'proximal contexts', which are further embedded in successive layers of sociocultural contexts. Successive layers of context can also be conceptualised as constituting different planes of analysis containing specific aspects of developmental processes, such as micro- and macro-levels of context (Engeström, 1999; Daniels, 2001). In my study, the micro-level analysis privileges the study of the detail of interactions in the lecture theatre. On this basis, the macro sociocultural context analysis would be related to predictions from the micro pedagogy. By combining a grounded analytical approach within a SCT perspective (Section 4.2.3), the focus of the analysis on "context" encouraged a broader analysis of structure and processes of the teaching activity in the specific setting of the lecture theatre.

Context as intellectual environment

Vygotsky differentiated between general, scientific or abstract and everyday, specific or concrete forms of cognition. He proposed that tools have more influence if they are less 'context-specific' (Bakhtin, 2007). As an illustration of the difference between sense and meaning, Vygotsky (2012) gives an example from a fable (the dragonfly and the ant) which ends with the words 'go to dance!'. In Vygotsky's interpretation, the words have a definite, constant meaning, but in the context of the fable they acquire a much broader 'intellectual and affective sense'. The sense of words is enriched by the sense they can acquire in different contexts. (Vygotsky, 2012) coined the fundamental law of dynamics as

> "*A word in a context means both more and less than the same word in isolation: more, because it acquires new context; less, because its meaning is limited and narrowed by the context.*" (p. 340)

From a linguistic viewpoint, the sense of a word changes in different minds and different situations in almost infinite ways. Vygotsky further suggests "zones of sense", a 'nesting' of different types of context in language, from sentence to paragraph, from paragraph to the book, the book from all the other works of an author and so on. In my study, I was interested in context as culture or intellectual environment and its relationships to thinking processes evidenced through lecturers' oral speech in lectures. The nesting of different meanings became of particular interest, as I discuss in Chapter 5. In statistics, context knowledge plays a critical role in thinking about data and is a motivating factor in students' learning (Langrall et al., 2006).

Scientific and everyday concepts

Vygtostky proposed that any learning that a child experiences at school has a previous history. The interpretation of ZPD is based on Vygotsky's distinction between scientific and everyday concepts and on his argument that a mature concept is achieved when the scientific and everyday versions have merged. For example, learning arithmetic starts with everyday knowledge and experiences with quantity outside school. Learning in school is concerned with "the assimilation of the fundamentals of scientific knowledge". Thus learning that occurs in different contexts (inside or outside school) is associated with the assimilation of different concepts: **everyday concepts** which precede **scientific concepts**. Referring to first language acquisition, Vygotsky further observes that development and learning are interrelated since the children learn language under adult guidance in a social setting.

Vygotsky emphasised that formal education can help students learn in the ZPD decontextualised, scientific concepts within a discipline. Connections to students' everyday concepts happen afterwards. Scientific concepts are fundamentally different from developing everyday concepts, outside of school since everyday concepts are not capable of abstraction. Abstract scientific concepts have the capacity to develop in a learner's consciousness and increasingly become more concrete through interaction with the teacher and learning tasks. Vygotsky (2012) used a geographic analogy to describe relationships of concepts:

> "*If we imagine the totality of concepts as distributed over the surface of a globe, the location of every concept may be defined by means of a system of coordinates, corresponding to longitude and latitude in geography. One of these coordinates will indicate the location of a concept between the extremes of maximally generalized abstract conceptualization and the immediate sensory grasp of an object i.e.,*

> *its degree of concreteness and abstraction. The second coordinate will represent the objective reference of the concept, the locus within reality to which it applies. Two concepts applying to different areas of reality but comparable in degree of abstractness e.g., plants and animals could be conceived of as varying in latitude but having the same longitude.*
>
> *The "longitude" of concepts will, thus, be the characteristic of thought processes, while the "latitude" will be the characteristic of their objective reference. [...] This position of a concept within the total system of concepts may be called its measure of generality.*" (p. 291)

In Vygotsky's law of equivalence described above, two concepts may apply to different areas of reality but may be comparable in degree of abstraction, i.e. their respective measure of generality. Vygotsky goes on to describe relations between concepts as relations between numbers. Any number, in his representation, can be expressed in an infinitive number of ways because of the infinity of number and because the concept of number controls all its relationships to other numbers.

Vygotsky argues that the ways in which tools are used vary depending on *context* and the student's own development (Vygotsky, 1978b). Since an everyday or a scientific concept can only be interpreted within the form or context they are used, Vygotsky (1998) also emphasised the changes that allow a child to move from an 'everyday' concept, then to a 'pseudo' concept and finally to a scientific concept. Vygotsky (1978b) goes on to say that

> "*The child does not suddenly and irrevocably deduce the relation between the sign and the method for using it. Nor does she intuitively develop an abstract attitude derived, so to speak, from "the depths of the child's own mind." This metaphysical view, according to which inherent psychological schemata exist prior to any experience, leads inevitably to an a priori conception of higher psychological functions.*" (p. 45)

Gradually being guided by an adult, the student becomes more able to abstract and generalise from concrete, everyday perceptions of objects. He warns that a pseudo concept can appear to be the same as a scientific (true) concept and that the movement from pseudo concept to scientific concept requires the guidance of an adult. To develop scientific concepts, the students needs to 'abstract out' the concept (Confrey, 1995). Teaching as *guidance in transfer* involves an intelligeble, shared context by the participants in the activity (Rogoff, 1982). Teaching constructs a context in which new information is compatible with student's current knowledge and skills and supports generalisation to occur.

Vertical and horizontal knowledge

SCT considers discourse to be at the centre of human development of higher intellectual functions. Bernstein (1999, 2000) is another author who contrasted and further refined specialist, abstract knowledge and concrete, everyday and local knowledge. As discussed in

Section 4.2.1.2, Bernstein provides an account of pedagogic practice in which macro (large scale, institutional) levels are integrated with micro (small scale, classroom) levels of analysis. Thus an analysis of specific pedagogic discourses as a set of rules that mediate the learning of statistical knowledge can enrich a characterisation of teaching discourse and practices in undergraduate statistics.

Pedagogic discourse refers to *what* is taught, the statistical, scientific content and skills and *how* they are taught by the teacher and learnt by the student. In his theory, Bernstein distinguishes between vertical and horizontal discourses (Bernstein, 1999, 2000). **Vertical discourse** takes the form of specialised, systematic and hierarchical discourse, as in the sciences. Knowledge in the vertical discourse is achieved through 'explicit rules of recontextualisation'. **Horizontal discourse** is usually everyday or "common sense" knowledge. Knowledge is achieved through 'oral, local, context dependent and specific, tacit, muti-layered and contradictory across but not within context' (Bernstein, 1999, p. 159). To make specialised knowledges more accessible to students, in education, Bernstein claims that "segments of horizontal discourse are recontextualised and inserted in the contents of school subjects" (Bernstein, 1999, p. 169). Thus, Bernstein's theory is a powerful way of explaining the interdependence between the development of scientific and everyday concepts. For instance, as a theory which transcend educational stage or context, it helps to account for those aspects of schooling that are meant to add to everyday knowledge without the development of scientific concepts. Further, the development of scientific knowledge must take place inside, rather than outside, school. In this interpretation, ZPD is therefore the distance between cultural, historical forms of activity or vertical knowledge, as provided by the sociocultural context and the everyday activity or horizontal knowledge. My analysis thus focused on processes of socio-cultural-historical transformation.

In my study, the lecturing situation is an attempt to open up ZPD involving a lecturer and the students on the module in which lecturer's discourse is made up of "scientific" concepts. Lecturers' discourse may allow for contradiction to arise when coming into contact with students' discourse which constitutes "everyday" concepts. Lerman (1996) considers that valuing of decontextualised, intellectual thought, divorced from personal and social elements, becomes formal discourse. However, Vygotsky pointed out that both scientific and everyday concepts contributed to each other, which result in "webs or patterns of conceptual connection" (Daniels, 2001, p. 53). In the case of teaching, Ericsson and Simon (1998) contrast verbal descriptions with **explanations** of thinking:

> "*Although verbal descriptions and explanations may not reflect spontaneous thinking with complete accuracy, such verbalizations present a genuine educational opportunity to make students' reasoning more coherent and reflective.*" (p. 183)

Consistent with Vygotsky's (1987) view of unity between thought and speech, Ericsson and Simon propose that participants in the teaching and learning activity have to 'unpack' complex silent thoughts into understandable ideas through verbalising, describing and explaining their thinking.

Representations

Any concept may have countless equivalent representations. Equivalence between representations depends on the relationships of generality between concepts. Representations, as a concept, are a mental state with specific content, a reproduction of a former mental state. Structurally equivalent presentations in teaching through pictures, symbols and signs where something can be in place of something else meaning that a symbolic similarity is sufficient (Seeger, 1998). Leontiev (1978) for example was critical of using particular representations in school since there may be a long way from concrete to abstract and the student already needs to know what an image is meant to illustrate. If what the image is meant to represent is not obvious to the student, the representation can become an additional obstacle in the learning process.

In the statistics classroom, there is a cultural practice of using representations (graphs, images, narratives) as means to meaningful statistics which require transformation of everyday knowledge (non-statistical, outside of statistics) into statistical concepts. So significant in my study was Vygotsky's system of two types, everyday concepts and scientific concepts that a learner develops (Vygotsky, 1987). The learning of these two types of concepts is context dependent. Everyday or spontaneous concepts emerge from personal life experiences. Scientific or "true" concepts represent scientific knowledge and are systematically learnt or developed during teaching and learning experiences.

Mathematical and statistical concepts can show similar generality as the words 'go to dance!'. The numbers in a statistical analysis represent discrete objects and statistical models can apply to an almost infinite set of circumstances. The general rules of parsimony in statistics require the simplest model or statistical theory with the least assumptions and variables but with greatest explanatory power. General principles imply that they can be transferred to a great number of applications. From a SCT lens, I considered that Vygotsky's system of concepts and lecturers' representations in lectures were relevant in characterising lecturers' use of context knowledge in lecturing.

Macro and micro contexts

In my view, a focus on scientific and everyday concepts in statistics lectures has important implications for teaching and learning. Although everyday concepts are expected to be 'outside' of school, in statistics the boundaries are less clear given the importance of "context" in developing statistical thinking (Makar and Ben-Zvi, 2011). Concepts are not 'absorbed ready-made', and the lecturer, through the medium of speech, is likely to play a (central) role in guiding students to reach awareness and mastery of their own thoughts. Although some of the knowledge statistics lecturers might present in lectures is *situated* in the context of those lectures, unlike some of the statistics students might use in the workplace, such examples are not in contradiction with ideas of generality and abstractness (Bakhtin, 2007).

Cole and Gajdamaschko (2007, 2010) suggest that people are influenced and shaped by forces of history and culture, i.e. "context". The same tool, for example a textbook, could mean different things in different modules and to different students. The environment does not possess an objective, absolute meaning irrespective of the people who live in it (van der Veer, 2007). Students develop new and different understanding through learning and so the environment can also acquire different interpretations as the student develops. Thus, in my interpretation of context as a relational construct, it is challenging to define the environment since individuals attach different interpretations or meanings to aspects of the environment. Further, the environment is social and it changes in response to mediated actions. However, in my study, the role of *context* may support different and competing interpretations of teaching process between statistical modules.

In my analysis, I was interested in extending this notion of "context" to include the context of the situation but also the context of culture and history (Lemke, 1997). In my conceptualisation, context can be physical or intellectual, micro or macro (Figure 1.1). A university as institution or a module can be described as a stable frameworks, with lecturers and students acting in relation to the university or module. Associated with each level of analysis, my unit of analysis is the individual lecturer engaging in goal-orientated teaching activity under conventionalised constraints presented at different levels of context (Section 4.7.2.2).

Educational research interested in researching teaching practice has recognised the importance of integrating findings across multiple contexts (Summers and Davis, 2006). Despite such interest, there has also been some confusion about what the terms *micro* and *macro* actually mean. For Hammersley (1993a), macro and micro refer to 'differences in scale of the phenomena which are held to explain the events under study' (p. xv). For example, an explanation of the influence of a particular teaching style on student engagement in a lesson (Section 4.2.2.4) is a micro explanation. The lecturer's beliefs about his choice of teaching style to be dependent on institutional constraints or expectations might be macro explanations. Hammersley (1993a) suggests that macro and micro are types of *explanations* and conceptualises micro-macro context as dimension rather than a dichotomy. The focus of the analysis was on both the micro context of the pedagogical structures and the macro context of the socio-cultural-historical teaching practice and taking into account tensions or contradictions in the context of changing shared teaching and learning practices.

4.2.2 Gage's theory of classroom teaching

As I showed in the previous section, within a SCT of developmental teaching and learning, the students' conceptual change is at the centre of the learning process, yet emphasising the prominent role for the teacher, peers and objects in *guiding* the students' thinking processes. Vygotsky's writings do not provide detailed descriptions of models of developmental teaching. Outside SCT of developmental teaching, Gage offers a concrete-objective theory of school-based teaching which views the teacher variables as fundamental in influencing student thought processes and achievement. This presentation of Gage's theory of classroom teaching and related studies is based of my ongoing literature review for the main study and is intended to define some of the terminology used in the Main study.

Gage (2009, p. 2) describes the term **teaching** as "all of the events which might have a direct effect on the learning of a human being". Research on teaching has *teaching* at its centre since it is the teacher who oversees all aspects of teaching, except for aspects which are predetermined, such as content of prescribed curricula. Everything else, the teaching style or the use of teaching resources are considered to be within a teacher's control.

At the centre of all these actions is the link between the teacher's subject-matter knowledge of the curriculum with knowledge about students. Subject-matter knowledge includes knowledge about the concepts which form the basis of teachers' planning. Knowledge about the students is concerned with knowledge about student learning and needs.

Thus, in his theory of teaching, Gage assumes that teaching is a natural social phenomena which is fundamentally the same across time or cultures, i.e. universally valid. Smith (1963 in Gage, 2009) defines teaching as

> "*a system of action involving an agent, a situation and an end-in-view, and two sets of factors in the situation: one set over which the agent has no control (for example size of classroom) and one set which the agent can modify with respect to the end-in-view (for example assignments and ways of asking questions).*" (p. 4)

Thus teaching actions are highly dependent on the teacher but also on the context in which the teacher operates. Gage proposes a model of teaching with six basic categories and 15 pairs of categories based on a historical overview of research into teaching, represented in Figure 4.3. In the second part of my main study, I used relevant concepts in Gage's theory of teaching to guide my observations in lectures and structure my literature review.

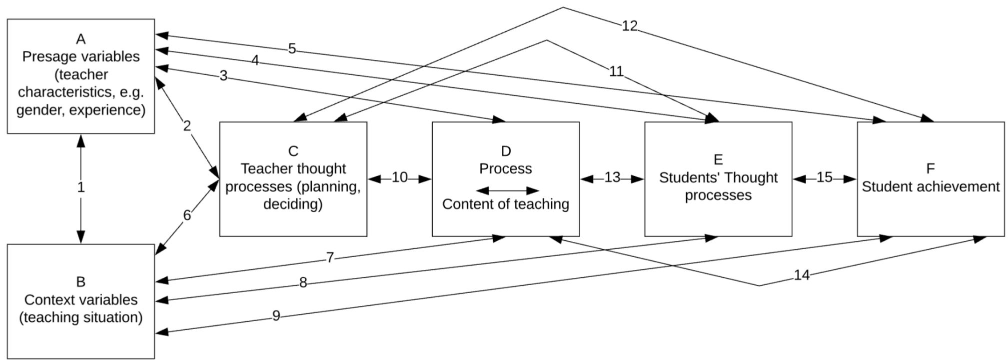


Figure 4.3: A paradigm for the study of teaching (from Gage, 2009, p. 47)

4.2.2.1 Presage variables (teacher characteristics)

This category consists of teacher characteristics such as gender, age, teaching experience and also intentions, beliefs, attitudes, values, etc. It also includes knowledge about teaching and knowledge about the subject matter or content. A number of conceptualisations of teaching knowledge have been proposed in the research literature (e.g. see Mali, 2016).

4.2.2.2 Context variables (teaching situation)

In this category, Gage includes characteristics of the country nested within a region, a community, a school and a class. Important in Gage's model is not the definition of particular aspects of learning, but the interactions between them. For example, researching teaching can build some understanding of the interactions between presage variables (A in Figure 4.3), context variables (B) and teacher thought processes (C).

4.2.2.3 Teachers' thought processes

A teacher's thought processes are studied before, during and after the teaching takes place. Research thus can focus on teachers' thought processes before teaching is interested on planning, how teachers organise facts, concepts or principles being taught. Research can also focus on teachers' activities during teaching, momentary thought processes when dealing with pedagogical content knowledge and *after* teaching when reflecting on teaching. Thought processes can also include attitudes, motivations, values, emotions. Presage variables (teacher characteristics) can interact with thought processes to impact on the teaching. Clark and Penelope (1984) proposed that teacher's behaviour is substantially influenced and determined by teachers' thought processes. The teacher's thought processes are about their thinking, planning and decision-making, which are a major part of the psychological context for teaching. Clark and Penelope's (1984) review highlight that the teacher is a highly reflective professional and that the teacher's thought processes (e.g. planning) have real consequences in the classroom (as discussed in Section 2.5).

4.2.2.4 Process and content of teaching

Gage (2009) proposed that the content of 'what' teachers teach deserve as much attention in research on teaching as 'how' they teach. The model separates teaching processes from cognitive processes. *Teaching processes* refer to the teacher's verbal behaviour and cognitive or social-emotional interactions with students. *Cognitive processes* are about students' mental activities while learning.

4.2.2.5 Students' thought processes

Research in mathematics education on students' experiences and thought processes, whether verbal or written, has looked at student views, expectations, attentional processes, motivation, memories, understanding, beliefs, attitudes, learning strategies, and metacognitive (monitoring-own-thoughts) processes (Mapolelo, 2009). In statistics education, processes of statistical thinking have also been the focus of a number of studies (see Chapter 2). In the context of my study, although I was aware of students' activity and thought processes, it did not form a focus of study.

4.2.2.6 Student achievement

Student achievement is a critical goal in teaching since it deals with the question whether, at the end of the teaching and learning phase, students *know* more than they did before engaging in the learning activity. In the hierarchical model of teaching process developed by Thomas (2011), mathematical competence is the highest goal in the hierarchy and involve the lecturer in the study verbalising his intentions, while being the culmination of all the other activities that lead up to *competent* knowledge. In addition to knowledge objectives, the lecturer might also have cognitive objects, of the type comprised in intended learning outcomes. In the pilot study (in Chapter 3), I was particularly interested in what lecturers desired for the students to achieve in their lectures, thus linking the *process of teaching* category with the *content* category in Gage's model. As in Treffert-Thomas' model of teaching process, the content or *what* the lecturer is teaching is part of the of process of *how* the lecturer is teaching, his actions, beliefs and behaviours.

4.2.3 A grounded approach within sociocultural perspectives

Adopting a grounded analytical approach (GAA) within an existing theoretical perspective as the one discussed in Section 4.2 is not without challenges (Seaman, 2008). According to Glaser and Strauss, theory is a "strategy for handling data in research, providing modes of conceptualisation for describing and explaining" (Glaser and Strauss, 1967, p. 3). In *The Discovery of Grounded Theory* published in 1967, Glaser and Strauss defined Grounded Theory (GT) as 'the discovery of theory from data' (p. 1). In this sense, this initial conception of GT saw *discovery* as emergence of categories of theory from data. Glaser and Strauss further stated that

> "*In the beginning, one's hypotheses may seem unrelated, but as categories and properties emerge, develop in abstraction, and become related, their accumulating interrelations form an integrated central theoretical framework – the core of the emerging theory. The core becomes a theoretical guide to the further collection and analysis of data*" (Glaser and Strauss, 1967, p. 40).

In this early version of GT, phenomena create their own representations that can be directly observed by the researcher. The stated general goal of GT is to generate *theories* systematically obtained and analysed rather than verifying theory. This suggests an objectivist philosophy which claims that research should be value-free and objective.

Since then, Corbin and Strauss (2008) considered that GT "denote theoretical constructs derived from qualitative analysis of data" (p. 1), rather than "building" theory from data more broadly. While studying chronically ill people, Charmaz (1990) proposed a social constructivist version of GT:

> *(1) Ill people's creation of taken-for-granted interactions, emotions, definitions, ideas, and knowledge about illness and about self and (2) Researchers' sociological constructions which they develop, in turn, by studying chronically ill people's constructions.* (p. 1161)

In this view, participants experience their constructions as reality. Categories and theory about people's beliefs and actions do not emerge from the data and instead are constructed or interpreted rather than discovered by the researcher in interaction with participants (data). The main purpose of analysis is to identify and explain contextualised social processes, thus focusing on process and change. So the analysis is more reflective of the context in which participants are situated than traditional GT since the analysis needs to relate to the participants' stories and to the worlds in which they live (Charmaz, 2006). In either versions of GT, the aim of the emerging theory was to clarify and explain social processes, actions and interactions. These two viewpoints indicate that the researcher might not need to go through the full interpretative enquiry to produce theory, but rather that an abbreviated version of GT is also possible as a collection of methods of qualitative data analysis to produce theoretical constructs (Corbin and Strauss, 1990; Bryant and Charmaz, 2007).

So, it is possible to reposition GT from a methodology with positivist underpinnings to an approach that can be used within different theoretical frameworks. In my study, GAA was used as a systematic approach, a fundamental process for doing GT, in the context of SCT (Section 4.2.1). GA can offer methods for collecting, managing and interpreting data while an analysis through the SCT lens can depict how social situations are constituted by culture and context. SCT has a well-defined and flexible methodology which can work well with a GAA since SCT concepts can supplement a grounded analysis rather than limit its scope. In this study, I adopt GAA as a tool to build up a line of argument, i.e. identify how factors, theories and ideas connect and relate to each other to reach a conclusion.

GA was considered useful when considering the multi-layered or nested depiction of "context" (discussed in Section 4.2.1.5). Both GAA and SCT focus on social processes of change over contexts and time (Charmaz, 2014). Taking a GAA within SCT helped my analysis of the multi-layered nature of the teaching activity at macro and micro levels (individual lecturer within the lecture theatre/ department/ institution) and was able to focus on societal, cultural and historical changes in different contexts. In my research study, my aim to build a *model* of how lecturers teach statistics interacted with my own interpretations and understanding. From an ontological viewpoint, the way I constructed reality was not value free since my theoretical perspective is situated within SCT. By using GAA within SCT, I was able to expand context-analytic procedures, such as seeking new data sources and analytic strategies that facilitate a more context-rich analysis.

All versions of GT take into account theoretical sensitivity of data, treatment of literature and other secondary data sources, constant comparison of data, data coding, theoretical sampling methods, and the use of memos (explained in Section 4.7.2). GT offers a methodology for gathering and analysing data inductively for context and for bringing process into a characterisation of teaching. There are several ways in which a GAA can be applied. In my research design, I adopt the approach described by Corbin and Strauss (1990), Corbin and Strauss (2008) and Charmaz (2014).

4.3 Research design

Qualitative *research designs* are generally *inductive* and need to involve data analysis by coding for themes and patterns. Maxwell proposes that

> *"The activities of collecting and analysing data, developing and modifying theory, elaborating or refocusing the research questions, and identifying and addressing validity threats are usually all going on more or less simultaneously, each influencing all of the others."*(Maxwell, 2012, p. 3)

The research process Maxwell describes does not involve choices from a fixed 'menu'. Instead, the research design is a 'blueprint' of research and, as Yin suggests,

> *"A research design is the logic that links the data to be collected (and the conclusions to be drawn) to the initial questions of a study. Every empirical study has an implicit if not explicit research design."* (Yin, 1994, p. 18).

So the research design deals with (1) *what research questions to study*, (2) *what data are relevant*, (3) *what data to collect* and (4) *how to analyse the results*. The main purpose of defining the research design is to help me build up the argument of how the evidence in my thesis addresses the research questions. The object of the action are other people, the students in the lecture theatre. Teaching is a process of interaction between the lecturer and the students in which the lecturer *guides* the students' learning. The use of psychological tools in a goal-directed, purposeful teaching activity can explain relationships between micro- and macro-levels of context by integrating analytical findings from the different contexts. I present more specific details about how I apply GAA within SCT in my study in Section 4.7.

4.4 What research questions to ask

Within an interpretative paradigm, the research questions are generally open-ended, descriptive and non-directional (Creswell, 2003). As I mentioned in Section 4.1, within this paradigm, typically the research starts with a broad, generic question which is later refined (Guba and Lincoln, 1994). Towards the end of my Pilot study (Chapter 3), I turned towards *re-formulating* the *research questions* that would guide and focus my research in the main study.

Using the theoretical perspectives in Section 4.2, the code system emerging from my literature review (e.g. Figure 2.2, on page 16), and the first draft graphical representation in Figure 3.3 (page 95), I was able to develop further my conceptual model and capture it in graphical format, represented in Figure 4.4, to include some of the aspects I considered influential on statistics lecturers teaching practice. In the representation in Figure 4.4 on page 132, I grouped "who" and "what" influenced the pilot study lecturers' teaching practice on two levels, the micro-level and the macro-level "contexts". I considered that this graphical representation is a "messy" conceptual model using some of the aspects emerging from the Pilot study and my theoretical perspectives (Miles et al., 2013).

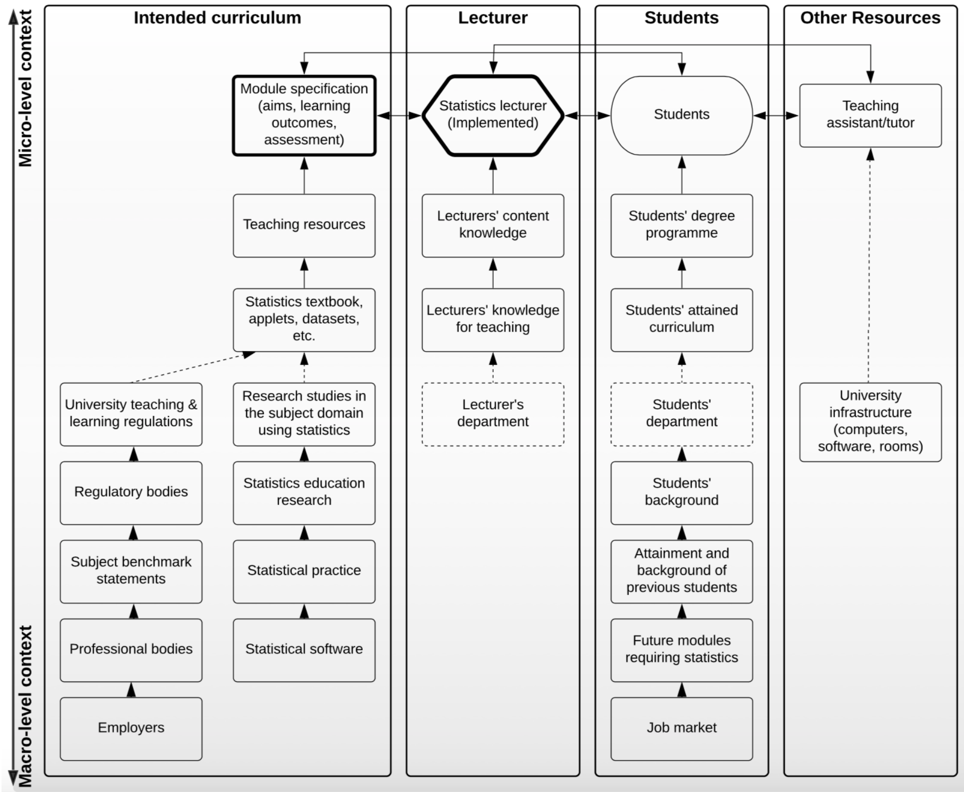


Figure 4.4: Main study initial conceptual framework

In the code system in Figure 3.3 (page 95), the aspects are grouped according to curriculum, lecturer, students and other resources. However, not all aspects have the same degree of influence on the lecturer. The module specification, the students and lecturer's research background appeared to influence the lecturers' teaching practice more than for example university teaching and learning regulations. Employers on the other hand were less influential *according to the data gathered from interviews.* For instance, one participant (participant 8) mentioned 'importance of statistical concept on the job market', yet at the same time linked this statement with 'easy statistics/methods'.

Another important aspect that influenced the lecturers' teaching in this data was the students' degree programme, not necessarily the same domain as the lecturers' research background. For example, participant 12, a mathematician, commented on the challenge of collaborating with the students' department in order to create curricula 'relevant' to students' Engineering degree programme. The lecturers' teaching experience and experience with previous student cohorts was another aspect which featured in the interview data. For example, in the Pilot study, participants 9, 10 and 13 contrasted 'easy' versus 'difficult' concepts or techniques for students to grasp, which relate to SCT concepts such as prior knowledge and the importance of preparatory stages (Vygotsky, 2012). Although regulatory and professional body requirements for accreditation feature highly in how intended curricula is designed in UK universities (Section 2.4.3), for *the participants in my study* these influences were only peripheral to their day-to-day teaching.

Based on this conceptual model, I therefore expected to study some, rather than all of these aspects and relationships in my main study depending on access to persons or documentation. The research questions were intended to implement the conceptual model. I therefore did not expect to include peripheral aspects in my research questions. Instead, I expected that contexts such as 'teaching practice', 'teaching process', 'content', 'lecturers' or 'teaching resources' would be featured.

Research questions represent 'facets of inquiry'. In the inductive grounded approach to my research design, I decided to formulate initial general research questions concurrently with the development of the theoretical assumptions for my study and refine or reformulate the research questions during the course of fieldwork (Miles et al., 2013). As I was adopting an SCT lens to my study, I was particularly focused on lecturers' roles, any points of tension/contradictions in their work, possible effects of (re)organisation of the elements in the activity system and on recognising incidents of contradictions that lead to change and transformation. Yin (1994) proposed to form research questions in terms of "what", "how" and "why" are an indication of which research strategy might be appropriate. The "what" research questions are exploratory in nature and hence require an *exploratory case-study.* Similarly, the "how" and "why" questions could be answered using a *case study* strategy that would include *multiple-case studies* relying on qualitative evidence (Yin, 1994). The choice of research questions based on my conceptual model (Figure 4.4) also guided my *sampling, data collection and data analysis* choices (Miles et al., 2013).

My overall research focus in my main study was characterising the **context** of teaching statistics by studying enacted teaching practices and processes. My focus on the role of "context" in teaching lead me to develop the following *generic* research question:

> ***How*** do lecturers teach statistics to students in the "context" of non-specialist undergraduate modules?

Early on, since I sought to describe and explain aspects of teaching practice on statistical modules, I decided to examine the process of teaching statistics using teaching resources within *the micro- and macro-level context*, in an empirical enquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 1994, p. 13). As a result, I was able to make my research questions more specific, as in Figure 4.5.

- **Purpose of study:** Represent the teaching of "in-service" statistical modules at undergraduate level from the perspective of the lecturers.
- **General topic area:** What insights can I gain about the teaching of statistics at university by studying the way lecturers use learning resources (examples, problems or exercises)?
- **Research questions**

1. *What characterises the "macro-level context" of teaching statistics at university?*
    - 1.1. **What** characterises the *teaching activities* on introductory statistics modules?
    - 1.2. **What** characterises the *lecturers' beliefs* about intended curricula and about students?
2. *What characterises the "micro-level context" of teaching statistics at university?*
    - 2.1. **What** are the statistics lecturers' *lecturing styles*?
    - 2.2. **How** and **why** do the lecturers adopt these teaching methods to teach introductory statistics?

Figure 4.5: Research purpose & questions (main study)

As I began to collect observational data, I also started to analyse the teaching I was observing by writing memos, making notes during teaching (e.g. about lecturers' planning and implementation in the lecture theatre) and observing the students in tutorials. In considering *teaching resources in context* to be fertile opportunities for lecturers to mediate (influence) students' statistical sense making (Rezat and Sträßer, 2012), I anticipated that for analysing the complexities of higher education statistics teaching practice, a GAA to data collection and analysis would be suitable (Charmaz, 2014; Corbin and Strauss, 2008; Wertz et al., 2011) within SCT. By using a grounded approach to my data collection and analysis, I aimed to characterise or conceptualise the two lecturers' teaching practice and reveal a *model* of the process of teaching in context. In my application of GAA, I started with line-by-line, open coding. As data was 'fractured' by the process of open coding, I started to assemble categories to group phenomena, while asking questions of the data such as "*how did a change in teaching context precipitate a (re)organisation of the activity system?*", "*In what ways and to what extent?*", "*Did these changes vary by case (module/lecturer)?*", "*How did the two lecturers shift from a concrete representation to another concrete representation or to an abstract one?*", "*How did the lecturers teach for statistical knowledge?*" and finally "*How did the lecturers engage and motivate students when teaching statistics?*".

4.5 What data are relevant

In my main research study, my purpose was to explain the teaching of statistical modules from the perspective of lecturers. Thus my focus in the main study was on the delivery of the modules but also the lecturers' planning and thinking. As such, my data is primarily in the form of words, language or text based on transcripts from *observations of teaching*, *interviews with lecturers*, *field notes* and *artefacts* such as handouts given in lectures or tutorials. To process this *qualitative data*, in my research, as discussed in Section 4.2, I use a grounded approach to frame my qualitative enquiry and as an approach to data collection and initial coding (Charmaz, 2014; Glaser and Strauss, 1967; Corbin and Strauss, 2008).

Throughout my research, I was guided by my research questions (in Section 4.4). As my aim is to understand how lecturers teach statistics and to build insights into the *context* of teaching of statistics, the *grounded approach* supported the data collection and analysis that focused on the teaching actions and processes and lecturers' intentions and strategies in relation to teaching within specific contexts (Charmaz, 2014).

As such, I used the sample of lecturers who participated in my *Pilot study* to select at least two lecturers (cases) for participation in my *main study.* I considered that multiple modules and lecturers (cases) are likely to offer a deeper understanding of the teaching processes I was interested in. I assumed that a study of "context" requires at least two different perspectives and that the *tools*, i.e. *domain* (e.g. Psychology, Engineering, etc.), *lecturers' teaching experience*, *lecturers' research background* and *students' background* influence a lecturer's teaching practice or activity. For this reason, I was interested to include lecturers with different backgrounds: teaching on different modules and departments, teaching within his/her department and outside their department, not trained statisticians yet using statistics in their research, teaching student cohorts with a strong background in mathematics or with assumed weak pre-knowledge of mathematics, with small/large student enrolment. Given the time-scales for my study, I also gathered this data within a specific time-frame.

Based on my sampling strategy, I contacted five lecturers from each domain of study (summarised in Table 3.1 on page 78). Three lecturers were planning to teach during the specified time-frame and were also willing to participate in my study. Guided by the content of the three modules, in the end I included two of the modules observed in the data analysis, Module A and Module B. My choice was justified by the content of the three modules, since the two modules were teaching similar introductory-level inferential statistics. Module A was an introductory statistical methods module aimed at Psychology students. The lecturer was a psychologist who had been teaching the module for five years. Module B was an introductory statistical methods module aimed at Engineering students. The lecturer was a mathematician who was teaching the module for the first time. I discuss the teaching-learning context of these two modules in more detail in Chapter 5.

4.6 What data to collect

The lecturers included in my main study data collection and analysis formed the primary *cases*. For each lecturer teaching a different statistics module, I collected data relevant to their teaching practice. The dataset for this study therefore comprised audio-recordings of lectures and laboratory sessions, interviews with lecturers at the end of the modules, memos and field notes, copies of student laboratory work and a sample of summative assessments (assignments and examinations).

I collected this data over one semester, as follows:

- *Audio-recorded observations*: Based on my research questions, I began to focus on the lecturers' teaching using examples, problems and tasks and I narrowed my literature searches towards the *teaching of statistics at university* and *context* in statistics education.

- *Memos and field notes*: during each observation, I kept field notes and at the end I wrote the key points and highlighted the parts of each lecture or laboratory session which seemed interesting at that time. I also captured whether students seemed attentive, the noise levels in the lecture theatre or laboratory, informal conversations with lecturers or students before or after teaching, which I was not necessarily able to audio-record.

- *Audio-recorded interviews with lecturers*: Interviews with the lecturers at the end of the modules.

- *Audio-recorded conversations with students* during breaks.

- *Student laboratory work*: Exemplars of student work during laboratory sessions.
- *Summative examinations*: Photocopied exemplars of students' assignments and examinations for summative purposes on one module. Although I used this data to make a judgement about student learning, it had a peripheral role in my main research study.

In order to gain insights into characteristics of teaching on statistical modules, I sought to explore the natural scene of the lecturers' teaching activities on three statistical modules by using participant or anthropological observation. My aim during these observations of teaching was to be as neutral as possible as a participant-as-observer (Gold, 1958). Since the lectures were attended by large numbers of students, I felt that my participation during lectures blended with the natural teaching activities and gave me access to the modules, documents and other insights into the delivery of the modules. During these observations, I was able to use an audio recorder and take notes from the back of the lecture theatre. Since I was not involved in the teaching of the modules, I was able to achieve a reasonable analytic distance from the module staff and students. Although I attended the lectures and tutorials as a researcher, I considered that my activity did not interfere with the expected working relationships between students and teaching staff in the lecture halls or laboratory sessions. In the case of classroom interactions, Delamont and Hamilton (1993) consider that observers cannot be entirely neutral even when they are not involved in the teaching activities. However, this research refers to classrooms in the context of small group teaching of children in schools. In the lecture theatre, where student group sizes and organisation are very different from what can be experienced in schools, I made the assumption that my actions as a researcher were largely ignored by students.

The focus of my observations was guided by the research questions. For instance, I contrasted the teaching methods of two modules, how the lecturers used teaching resources (e.g. examples, exercises, problems) and studied how they delivered particular statistical topics. Since initially my research questions were more general, I was interested in gaining insights into the teaching process rather than narrowly focusing on pre-determined categories. Similar to the pilot study (as outlined in Section 3.3), I did not use an observation schedule.

4.7 How to analyse the data

For analysing this *qualitative data*, I was concerned to build a set of **categories** guided by the mediational triangle in Figure 4.2 that could be tested against additional data, i.e. additional observations of lectures and interviews with the participants. My analysis was carried out on two levels, the macro and the micro-levels of context (Section 4.2.1.5). Eventually, my intention was to use these **categories** to create inter-relations between them towards an emergent model of teaching statistics in "context". Of first concern were however the ethical considerations from which to draw evidence of a good quality data collection process, which I discuss next. I then provide an account of the two modules included in the data analysis, my analytical approach for the analysis of observations data, documentation and the post-module interview data.

4.7.1 Ethical considerations during data collection

Before starting to collect the data, I wanted to define in some detail the expectations I wanted to build with the lecturers in order to maintain and improve the quality of my interpretations (Miles et al., 2013). Within an interpretative paradigm, I was aware of my role as a researcher in the data collection and analysis. As lecturers, my participants had control over what happened in the lecture theatres. As best as possible, I aimed to keep the purpose of my presence in the lecture theatre as unobtrusive as possible and maintain the anonymity of students and lecturers. Anonymity was important to ensure the participants presented their beliefs about their teaching and any other information about the modules in an unbiased manner, avoiding self-censored or defensive accounts.

From the start of my classroom observations, I negotiated with the lecturers how much time and effort would be required from them, what kind of data is involved, the researchers involved in the study, the participants' role in the data analysis and possible benefits to them as lecturers but also to the wider research and teaching community, including the teaching assistants.

The agreement between my participants and me was that I can participate in observations of live teaching and that they will provide documentation relevant to my study as well as participate in a post-module interview. In my research design, I tried to ensure my participants invested a reasonable amount of time. However, they did not receive a compensation for participation.

Burgess (1985) vividly describes some ethical dilemmas with carrying out ethnographic research in a school. During my research design, I was mindful of my role as a researcher in the lecture theatre or laboratory, boundaries of informed consent and data dissemination. The institution and the classrooms/lecture theatres where my detailed observations were situated were susceptible to improvisation, not stable or durable.

4.7.2 Analysis of observations of teaching and documentation

The analysis of the observations of teaching and related teaching resources taking a grounded analytical approach started with an initial overview of the *field notes and memos*, followed by *transcription* of the data and open coding. **Constant comparisons** *within* and *between* the two modules and lectures enabled me to start to build a set of categories and emergent themes that I could then test against other data. Within the grounded approach, I used the following ten iterative strategies to analyse the audio-recorded observational data of lectures.

1. Data collection of lecturers' teaching practice.

2. Memos and transcription.

3. Pilot analysis: defining the unit of analysis and open (initial) coding.

4. Sampling and further open coding: identify additional relevant data for analysis.

5. Comparative analysis of data: compare chunks of data (teaching episodes identified in the pilot analysis and through further sampling of the data).

6. Categories: start to formulate categories (ideas) based on the emerging codes.

7. Comparative analysis of data: further sampling of data.

8. Conceptual saturation: themes, what analytically different groups of concepts are indicating

9. Test the themes: develop a theoretical explanation of the teaching with context dimension of teaching practice, i.e. based on emerging categories and themes derived from the pilot analyses, identify further opportunities for data analysis to answer the research questions and delineate relationships between themes.

10. Inter-relationships among themes towards the development of an emergent model of teaching within a "context".

These ten steps are summarised in the tabular representation in Table 4.1 (Harry et al., 2005). In this analysis, I was interested to identify major issues and themes, and through constant comparisons gradually focus on the emergent issues. My aim was to generate a picture of teaching actions and processes using examples and problems in context. As it is common in GAA, the data formed the basis for the emergence of theoretical ideas relating to the lecturers' use of learning resources in undergraduate statistics education. My aim was to produce a model of teaching that embodied a theory (explanations) of teaching and represented the relationships between entities that formed the theory.

Table 4.1: Data analysis map

Levels of analysis			Codes
⟲	6	Model	*Increased levels of abstractions*
↑	5	Inter relationships & validation	...
↑	4	Test the themes	...
↑	3	Themes	...
↑	2	Categories	...
↑	1	Open Codes	*Based on observations and interviews*

In this section, I provide a rationale for each of these steps in my data analysis. In the next Chapter 5, I continue with details regarding my interpretations of the data, examples of analyses and findings.

4.7.2.1 Memos and transcriptions

Once I decided which modules should form the cases enquiry, the next step in the data collection was to keep memos and make notes during and at the end of each observation. During these stages of the data collection and analysis, my memos were characterised by crude evaluations of the teaching that I was observing. For instance, I often questioned whether students were making sense of the material lecturers presented to them. In one of my comments on lecture 10, module B on Experimental Design, I jotted down

> **Quote 3 (Memo, Module B, Lecture 10)**
> *'Student **noise levels** are rising sometimes when the lecturer is talking (e.g. during the section on threats to construct validity). I wonder what this might mean. Do the students understand this or is there a feeling that this is not **relevant** to them? There are no 'calculations' in this topic, it is **philosophical** in nature. Is there something to do with the amount of knowledge the lecturer delivered to students or with when this topic was delivered (at the very **end of the module**). It seems that students were less **engaged** today than in previous lectures'.*

In this example, I questioned a number of aspects of the teaching: the curriculum design (the sequences of learning during the module), student behaviour (elevated noise levels), student sense making (not necessarily visible during the lecture) and the lecturer's ability to convey meaning through this mode of teaching. Such memos gave me some insight into teaching practices on these modules and supported my next stage of the initial data analysis. However, later in the data analysis, I realised that during the early stages of this research journey I was *critical of the teaching* I was observing, while later on, especially in the comparative analysis of data, my main focus was to become *critical of my interpretation* and justify relating conclusions. Memos helped me deal with instances of confusion, doubt or insight during data analysis. I used memos as instances of data which helped me conceptualise and reorganise my analysis.

My analysis of observations of teaching and interviews involve transcription using software. For this study, I used Transana[3], a qualitative methods software which can be used as a tool for transcribing and coding the data.

I considered that coding full transcriptions of lecturer discourse in lectures and interviews could bring me a deeper level of understanding (Charmaz, 2014). Transana also allows the linking of pictures/slides to lecturing transcripts and memos, so I was also able to code slides for a better understanding of the context of teaching. I explain my coding process in the next section. Although transcriptions can make the data analysis process more efficient and 'preserve the rapid flow of ideas' (Charmaz, 2014), they also require skill in analysing. The extracts from transcripts or lecturing documents might not necessarily be *illustrative* of the

[3]Transana supports a qualitative analysis approach to coding which I used in this analysis. First, I selected analytically interesting or important quotes or episodes with transcripts. The next step was to create codes and sub-codes. The software also allows for writing memos for codes and episodes and create diagrams based on the coding scheme.

data. Instead, I regarded them as *instances* of the data itself (Wertz et al., 2011).

4.7.2.2 Unit of analysis and pilot analysis

My next step in the data analysis was to define the **unit of analysis**, i.e. the concept that will form the focus of my analysis, a "case" which Miles et al. (2013, p. 45) defines as 'a phenomenon of some sort occurring in a bounded context.' Defining the unit of analysis was a core challenge in translating the theoretical perspectives into an applied method of analysis. For Vygotsky (1987), the unit of analysis was word meaning, a holistic internal aspect of the word as a microcosm of human consciousness. However, some regarded word meaning as inadequate in explaining the relationship between "elementary" and "cultural" development in ontogenesis (Cole, 1985) and meaning is not "a genetically primary unit for the analysis of mind" (Zinchenko, 1985, p. 100). As a result, later research based on SCT approaches have searched for other units of analysis, such as mediated action (Wertsch and Stone, 1985; Zinchenko, 1985), activity (Davydov and Radzikhovskii, 1985), the object-oriented and artifact-mediated collective activity system (Engeström, 1987), activity systems (Engeström, 1999), the person speaking (Bakhtin, 1986) or community of practice (Wenger, 2000; Jaworski, 2008), to name a few examples.

Vygotsky (1987) argued that external relations among elements are not the same as internal relations of a unit:

> "*In our view, an entirely different form of analysis is fundamental to further development of theories of thinking and speech. This form of analysis relies on the partitioning of the complex whole into units. In contrast to the term "element", the term "unit" designates a product of analysis that possesses all the basic characteristics of the whole. The unit is a vital and irreducible part of the whole.*" (p. 44)

On the basis of Vygotsky's scientific activity, Zinchenko (1985) outlines seven requirements essential in a unit of analysis. Several of these requirements highlight that a unit must be a functionally integrated, holistic psychological structure and have different and even contradictory characteristics (or origins) that is the catalyst of development.

In SCT, the mediational triangle represents the idea that activity systems are a basic unit of analysis for understanding human activity, which includes the subject (or subjects) whose agency is selected as the point of view of the analysis, the object who is acted upon, and the dynamic relationships btween them (Cole and Engeström, 1993; Barab et al., 2002). As in Figure 4.2, the relations between the subject and the object of the activity are not direct, they are mediated by a range of factors, including tools. The "doing" is to transform or change something (Kuutti, 1996).

Working within SCT, I was guided by my research questions (in Figure 5.1) and the conceptual framework in Figure 4.4. More specifically, in defining my unit of analysis, I considered the mediational triangle in Figure 4.2 was my 'minimal meaningful context' for understanding teaching actions in the activity system as the unit of analysis (Cole, 1985).

To broaden the scope of my enquiry beyond observations and interviews, I also needed an integrated account that brought the macro and micro together, the lecturers' activity during teaching and the social situations and concerns before/after teaching (Lerman, 2001). In this study, the macro and micro tool-mediated activities as units for analysis allowed me to study the changes in activity in which the lecturers and the contextual transformed social reality.

In the **macro analysis**, I focused on lecturers' participation at three levels: the module teaching team, the university and the research/statistics community. The data for the macro analysis came primarily from interviews with lecturers in the Pilot study (Chapter 3) and Main study, documents relating to the module, teaching assistants and the literature review in Chapter 2.

For the **micro analysis**, the unit of analysis was looking at the detail of activity *during* teaching, focusing on particular *teaching episodes* (TE) in the modules. At the micro level, I analysed the observational data from the perspective of the two lecturers engaged in teaching actions in the context/setting of the lecture theatre mediated by teaching resources (tools).

Before starting to select teaching episodes from each lectures, I read the transcripts from beginning to end to get a feel of the lecturer's activities during the lecture. Using transcripts of data, my focus was on quotes or TEs which represented the statistics lecturers' teaching activity of inferential statistics on introductory modules to students in the lecture theatre using *tools*, such as examples, problems or exercises. Thus I defined as a TE a chunk of teaching, a lecture activity or event coded as example, exercise, problem or task. A lecture could then be chunked into temporal sequences of teaching as continuous process of ongoing real activity in a real setting that lead to the emergence of an exposition, the 'story' or 'narrative' of the teaching under investigation.

The initial, pilot data analysis involved four TEs which I selected for three reasons. First, they were the opening examples the lecturers used to introduce the topics of t-tests and analysis of variance (ANOVA), two of fundamental statistical models students learn in introductory statistics. Second, based on my initial coding, the first two TEs from the lectures on t-tests yielded a rich set of concepts (ideas) that could be compared and contrasted in useful ways. Third, the lecturers used the (everyday) "contexts" of these tasks throughout the module, when teaching both topics. I then proceeded to identify additional TEs in the lectures on *t*-tests and ANOVA across the two modules.

4.7.2.3 Sampling and open coding

For Charmaz (2014), coding is about defining "what is happening in the data and what it means". Coding allows the weaving of *generalisable* statements and *specific* contextual analysis of actions and events. My theoretical perspectives shaped my analytical approach by connecting fragments of data with the analytical abstraction of my analysis.

In the initial stages of the coding process, I took a first episode of data (L1 introducing t-tests) and used it as the starting point of the coding process, using the Transana software.

Based on the theoretical perspectives, during this period of 'focused coding', I was sensitive to issues of teaching process, content and context. I also focused on actions and emotions to preserve a sense of flow, precision and focus on making connections between fragments of data and on what was happening in the data Charmaz's (2014).

As I was analysing the data, I also wrote notes which were labelled with a **code**, a conceptual label representing my analytic tool and interpretations of the lecturer's actions in the episode. For example, if I coded an episode in one of the lectures as "context", I then sought to find other instances where the lecturer used (other types of) "context" and how this compared to examples from the other module. As new codes were emerging relating to "context", I created a category entitled "context' and added new sub-codes to it. The codes were quite concrete but close to the data and suggestive of possible processes in the data. The codes also illustrated the context/situations in which the codes occurred.

4.7.2.4 Comparative analysis of data

Based on my memos and an initial coding of the transcripts of these two lectures, TEs represented units of **comparative analysis**. Although within-case or **within-TE** analysis brings understanding in its own terms, cross-case or **cross-TE** analysis or synthesis can increase the *generalisability* and *transferability* of interpretations to other contexts and gives confidence that the processes identified are not entirely idiosyncratic (Miles et al., 2013).

My approach to the data analysis consisted of looking at two lecturers' teaching processes and content to observe how particular lecturers teach introductory statistics. I collected data about each individual lecturer, including context variables and lecturer beliefs in the pilot and main studies. I then compared these rich profiles for analysis to identify historically grounded patterns.

In my analysis of observations of teaching, after initial, open coding and further sampling, I proceeded to compare TEs that I identified in the previous step from different parts of the lecture, different lectures and different modules in terms of how similar or different they were conceptually based on the coding applied. Where I used the same code in different episodes, I considered whether there were new or additional meanings of the code. Similarly, I also considered how the codes were different or conceptually distinct from each other. I considered that a cross-TE analysis would allow for identifying patterns of context, process and content.

4.7.2.5 Identify categories and further comparative analysis

Coding involves attaching labels to chunks or segments of data that depict what each chunk is about which then allow to ask analytic questions about the data. Chunks of data represented by codes then can be compared with other codes/segments of data. Through coding, constant comparisons and writing memos, I was then able to produce **categories** which answer questions about the data (Charmaz, 2014). Categories are at a higher level of abstraction than codes and themes are at a higher level of abstraction from categories.

The relationships between categories allowed for further analytic refinement of the data into **themes**. Charmaz (2014) emphasises that grounded analytical approaches state that the aim is to 'construct theory' rather than generate themes. In my analysis, using the unit of analysis, I analysed actions and processes inductively rather than focus on generating themes.

In order to generate **categories** and then **themes**, I compared episode with episode from the same lecture, across different lectures and across modules for similarities and differences. I also identified incidents (smaller chunks within an episode) which were worth comparing. Codes that were considered to be similar in some way I grouped together under a higher level descriptive concept, or category. This enabled me to differentiate between different categories and themes and identify properties and dimensions specific to that category/theme (Corbin and Strauss, 2008). I also attempted to make connections between the codes, categories (assembled codes, grouped phenomena that represent increased levels of abstraction) and themes (ideas, representations).

4.7.2.6 Reaching conceptual saturation and identifying themes

This process of comparing teaching episodes within and across lectures and lecturers, revealed over thirty characteristics of these two lecturers' teaching which I further expressed in terms of themes or concepts. This grounded analysis was helpful in characterising these two lecturers' teaching, what learning resources these two lecturers were using on their modules how lecturers were teaching in the lecture theatre, their use of examples and other learning resources. Following my ten-step iterative data analysis process, at this point in the analysis I sought to carry out comparative analysis of data using additional TEs and produce further categories (ideas), substantiate and justify my interpretations of the data in order to move towards more robust conclusions.

The grounded data analysis allowed me to identify similarities and differences in what characterised the context of teaching on the two modules (first research question) and how the two lecturers taught using context (second research question) along three themes:

- Context and abstraction: shift from concrete context to another concrete context or abstract (mathematical/statistical) context.
- Statistical knowledge: emphasise procedures, build skills and concepts, discourse that includes statistical arguments or limited mathematical content.
- Lecturer intentions to engage and motivate students by emphasising knowledge from other domains (media resources, psychological studies), the utility of statistical techniques or simulating the process of statistical enquiry.

Comparisons and additional data also clarified for me particular themes, such as what I meant by 'constructing' and 'de-constructing' context. The main theme that was starting to emerge was centred around contextual knowledge and its relevance to statistical knowledge. In the background, lecturer's motives and beliefs started to emerge as a theme from codes such as knowledge from other domains (engineering, psychology, management) and what I perceived to be their teaching objectives. However, since my observational data did not allow me to fully tap into lecturers' motives and believes, I needed to turn to the interview data to further the analysis.

4.7.3 Analysis of post-module interview data

At the end of the module, two to three weeks after the final week of teaching, I interviewed the lecturers about their experiences teaching the module. I collected the interview data to reveal the extent to which participants individually identified or abandoned the tools used during teaching. I aimed the interview questions towards the meaning participants assigned to the teaching process. Transcripts of naturally occurring talk (as in the lecture theatre) revealed *how* the lecturers taught, *who* spoke and allowed for comparisons to stated beliefs in interviews - *why they taught in the way that they did.* By comparing instances of naturally occurring talk with formal/informal interviews, it was possible to identify unique 'speech genres' and situate them within institutional/cultural contexts (Wertsch, 1993).

The comparisons between observations and interviews can then reveal teaching styles situated in institutional/cultural context. Such comparisons between different data sources could reveal contradictions and how they were resolved. In this way, I was able to in the analysis lecturers' tacit beliefs about teaching/ students/ statistics, local situations of the module which might be had been apparent with analysing only observational data. In this analysis, I focused in particular on how lecturers' experiences changed during the course of the module, how, what, and when they noticed such changes.

4.7.3.1 Post-module interview research questions

The three research questions that directed the data collection and analysis are as follows:

- What are the lecturers' beliefs about teaching/ statistics/ students?
- What learning resources do lecturers use on their modules?
- Why do these lecturers use learning resources in their teaching?

In these interviews I was therefore interested in what ways and why, the use of learning resources was similar or different among the two modules.

4.7.3.2 Data collection

I scheduled interviews with each lecturer at the end of each module, once the teaching and assessment activities had finished. Although I had informal conversations with the lecturers during the module, the interview was an opportunity for me to gather systematic data about the lecturers' views on specific issues relating to my research questions. The aim of the data collection and analysis was to make comparisons between the two modules. As my initial data analysis was done inductively throughout the data collection phase, I used a an interview guide, in Appendix C.1, to capture some of the topics I saw in the data at the time.

4.7.3.3 Interview format

For this study, I chose an approach to interviewing my participants which could be described as 'structured interview'. At the beginning of the interview, I explained that I was interested to find out more about the participants' experience in teaching the module and the influencers in their choice of particular learning resources such as tasks, problems and examples. I then proceeded by asking them some background information about their academic and research background, teaching experience and experience in teaching this particular module.

After asking the lecturers an open-ended question, the participants were encouraged to tell their experience of teaching their module and only when they had finished their narrative I asked questions about points they brought up in the interview that I felt needed further elaboration, but also about other areas which I was interested in exploring and had prepared for the interview. Although I had prepared an interview guide, I allowed the interview to flow freely.

For example, I asked the participants to tell me whether, as they looked back on this module, any lectures stood out. This was followed by questions relating to their teaching, such as describing the most important ways to teach statistics and how they designed/chose the tasks and examples they used. Other areas which I explored during the interview were about the lecturers' use of context, student evaluation and assessment, statistical and mathematical content, statistics as subject and the importance of software.

In order to obtain rich data about the lecturers' views about the way they used learning resources in their modules, I prepared copies of such resources from the three modules I observed. The interviewees were encouraged to reflect on the use of the resources, how their resources differed from or were similar to the resources used on other modules and why.

At the end of the interview, I asked the lecturers whether there was something they'd do differently with the module or whether there was something else they would like to ask or add to what has already been said. The interviews lasted approximately sixty minutes.

4.7.3.4 Data analysis

The first step in the data analysis was to transcribe the interviews. Before starting to code the interviews, I read the interviews from beginning to end to get a feel of what the interviews were about. At the start of the coding process, I took a chunk of data (e.g. the beginning of the interview) and used as the beginning point of the coding process, using the Transana software. As I was analysing the data, I also wrote memos which were labelled with a code, a conceptual label representing my interpretation of what the lecturer said in the episode.

Chunking the data

In order to make the data analysis manageable, I broke down the interviews into manageable episodes (chucks) with the aim to examine each episode in more detail. I used this approach to also compare episodes from different interviews in terms of how similar or different they were conceptually (i.e. in terms of the coding applied). Where I used the same code in different episodes, I considered whether there were new or additional meanings of the code. Similarly, I also considered how the codes different or were conceptually distinct from each other.

For example, if I coded an episode in one of the interviews as 'informal teaching', I then sought to find out this lecturer's beliefs about teaching and how these beliefs compared to the other participants. As new codes were emerging relating to the 'teaching beliefs' theme, I added them to it.

Linking codes

Once I started to break the data apart and delineate codes (concepts) to stand for episodes of data, I also attempted to make connections between the codes and relate them to themes.

For example, using the next episode, I related the broader code 'informal teaching' to the more specific 'teaching statistics as an everyday conversation' and 'student numbers'. I then hypothesised that the interviewee's beliefs about teaching and his actual approach to teaching in the lecture theatre were influenced by the context of this module. I was then able to check in the data the lecturers' beliefs about teaching and what they perceived to influence their teaching practices. As these codes were connected, I could also elaborate on them and explore links between how a lecturer's 'teaching beliefs' influence his 'teaching strategies'.

Analytic strategies

The next step was to ask questions of my data which I could apply across datasets, such as for example:

- What are these lecturers' beliefs about teaching?
- How does the classroom situation influence their teaching practice and choice of strategies?
- What tools do the lecturers use in their teaching and why?

The interview data I was then able to include in my analysis of observations and learning resources. Integrating coding schemes from different sources of data was possible through using the Transana software.

4.8 Summary of chapter

This chapter depicts the process of defining my research paradigm, the theoretical perspectives and analytical approach and the research design. Here, I sought to provide evidence about 'how' I carried out my main research study, and also 'why' I did the study in the way that I did. Within an interpretative research paradigm and SCT of teaching and learning, l designed a qualitative study of observations of teaching statistics at university. My theoretical and analytical choices were motivated by my focus on teaching process, content and "context" and guided my data analysis and interpretations. Further, I adopted a grounded analytical approach to my data analysis. In the next chapter, I present the outcomes from the data analysis and aim to describe a model of university teaching of introductory statistics in context.

5 Main study Data Analysis and Findings

Chapter contents

One of the outcomes of my Pilot study in Chapter 3 was a 'rough' conceptual framework of teaching practice, which I further refined in Section 4.4 to formulate the research questions. As a reminder, my research questions are as summarised in Figure 5.1. In Chapter 4, I described in some detail the methodology for my main study, which includes my research paradigm, theoretical perspectives and research design. A summary of the data analysis process is in Figure 5.2. My goal ultimately is to understand the teaching of statistics in the lecture theatre – *how* and *why* lecturers make and implement decisions in the course of lecturing.

In my data analysis, I am interested in the **teaching** of introductory statistics modules at university in **context**. More specifically, adopting a grounded analytical approach within a

1. *What characterises the "macro-level context" of teaching statistics at university?*

    1.1. **What** characterises the *teaching activities* on introductory statistics modules?

    1.2. **What** characterises the *lecturers' beliefs* about intended curricula and about students?

2. *What characterises the "micro-level context" of teaching statistics at university?*

    2.1. **What** are the statistics lecturers' *lecturing styles*?

    2.2. **How** and **why** do the lecturers adopt these teaching methods to teach introductory statistics?

Figure 5.1: Research questions (main study)

Research methodology *(Chapter 4)*
$\updownarrow$
Data analysis (Methods and interpretation)
$\updownarrow$
Report findings (Integrate findings)
$\updownarrow$
Relevance and contribution *(Chapter 6)*

Figure 5.2: The data analysis process

SCT framework (as in Section 4.2), I seek to integrate my findings across macro- and micro-levels of context in order to examine teaching practice. In my representation of lecturing practice and processes, I attempt to provide a description of the properties or components of the teaching process and content and relate *context* (micro and macro, physical and intellectual), *lecturer beliefs*, *lecturer contextual variables* and *lecturing tools*. Fundamentally, I seek to understand how these components of teaching process and content change, transform and interact with each other during the delivery of the module. In my interpretations, I use the Pilot study findings (Chapter 3), the post-module interview data and the observations of lecturing.

5.1 Ethical considerations in data analysis

The lecturers granted me access to their lectures and laboratory sessions, which I audio-recorded and wrote memos (as described in Section 4.7). At the beginning of my study and throughout the data analysis process, I contemplated on a number of ethical issues which might have implications for data analysis (Miles et al., 2013). In Section 4.7.1 I described how my participants gave full consent to participate in my study. In order to ensure the trust of my participants, I also agreed with them regarding the type of data analysis I was doing, which was transcribed, annonymised and coded using software (as in Section 4.7).

In the literature review and the Pilot study, I defined in some detail the worth of the project and the contributions I intended to make to statistics education research. My competence as a researcher was an important aspect of the data analysis (see Section 4.1). During the project, I also improved my research competence through training, discussions with supervisors and attendance at conferences. As the study progressed, I made my interim findings public in presentations in workshops, with the agreement of the participants through following the University's Ethical guidelines and procedures.

5.2 Macro versus micro analysis

In SCT, the social context is affected by cultural and historical factors (in Section 4.2.1.5). Relationships between micro- and macro-levels of context, seek to understand context in terms of psychological and cultural tools (Section 4.2.1.3) that are provided by the *sociocultural* context of university teaching and learning and used by individuals in their *interpersonal* context. In the *micro* analysis I studied instances of lecturers teaching their students using lecturing tools in the lecture theatre. With the macro analysis, I intended to shed light on the more complex issues that account for what I was observing at the micro level (Lutz, 1993).

Macro analysis (Section 5.3)
↕
Micro analysis (Section 5.4)
↕
Model of teaching statistics in context (Section 5.5)

Figure 5.3: Macro-micro data analysis

Further, I conceptualised micro as *small* scale and *macro* as large scale. The lecturers adapted their teaching styles or methods to the students at both micro and macro levels. For example, on a small scale, in the lecture theatre, lecturers provided tailored representations of statistics depending on their goals and beliefs about the students. On the large scale, lecturers adapted the content of the modules to allow for students to achieve the learning outcomes.

To examine the teaching of statistics at university, in my study, I attempt to integrate findings from across multiple contexts taking a SCT perspective on teaching (Section 4.2). The theoretical framework used in this analysis emerged through data analysis, which started using the grounded approach and then applying the framework to subsequent analyses using the mediational triangle (Figure 4.2, p. 114). In my analysis, I regard the macro-micro as a dimension rather than two separate components of analysis (Hammersley, 1993b). For clarity, in this chapter, I present my analysis of macro and micro separately, yet being aware of possible overlaps among components of analysis, perspectives on the data which are half-way along the macro-micro dimension and relationships between them. Engeström (1987) further suggests that in a complex society, there are a multitude of relatively independent activities that are represented as sub-triangles but with the same internal structure. One cannot assume the existence of a singular activity system. In my interpretation, I also sought to make connections between macro-micro. To build an understanding of teaching statistics in context, I conceptualised the mediational triangle in Figure 4.2, p. 114 for each layer of analysis, macro and macro with inter- and intra-relationships between the nodes (corner of a triangle such as 'subject') of two systems, as in Figure 5.4.

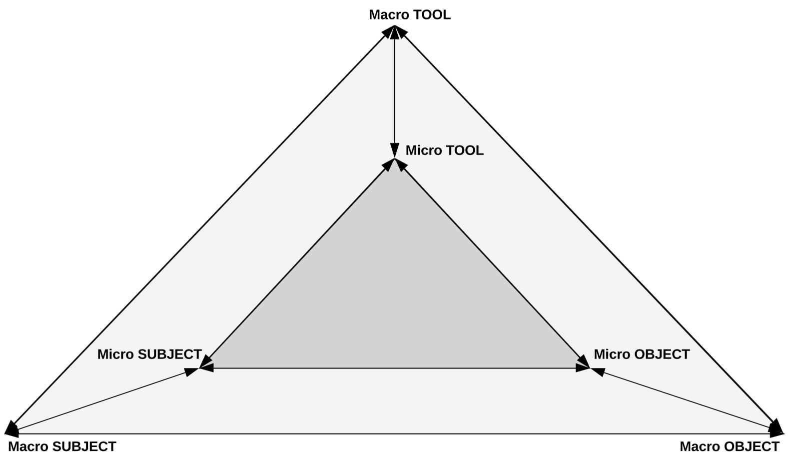


Figure 5.4: Integration of macro and micro analyses

The macro-level analysis is detailed in Section 5.3, followed in Section 5.4 by the micro-level analysis. The factors included for each node of the mediational triangles emerged during data analysis taking a grounded analytical approach and with a basis in the theoretical perspectives in Section 4.2 and the conceptual framework in Section 4.4. Next, in Section 5.5, I represent the macro-micro factors as a model of teaching statistics in context.

5.3 Macro-level analysis of context

A first step in the analysis was to look holistically at the two modules' teaching-learning environment and start identifying contextual factors and relationships. In Section 4.7.2.2, I explain my approach to the data analysis and how I selected the unit of analysis for the macro level. Using the components proposed by Gage (2009, in 4.2.2) and concepts of my theoretical framework (Section 4.2.1), my interpretative analysis started with a description of each module's teaching and learning context, teaching approaches and content.

As a precursor to the analysis, I filled in each node of the mediational triangle in Figure 4.2 with empirical data from the Pilot study (Chapter 3) and literature searches (Barab et al., 2002). I depict my analysis in graphical format in Figure 5.5.

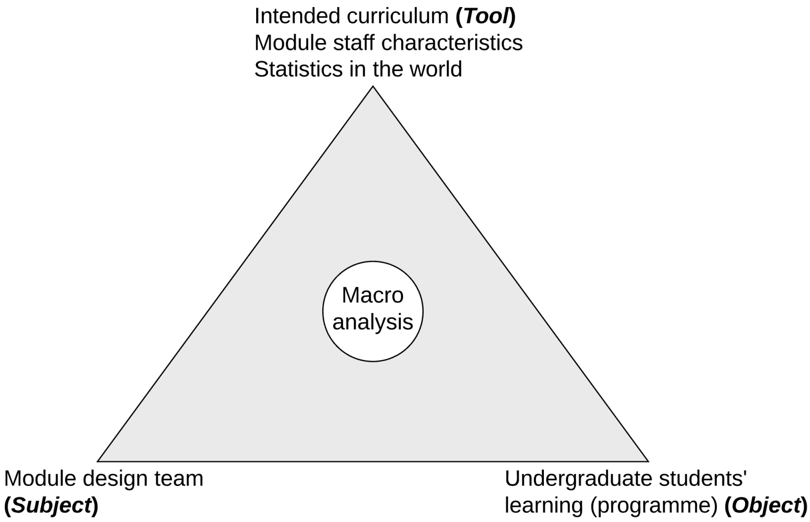


Figure 5.5: Macro analysis of teaching

Subject can refer to an individual or a sub-group whose agency is chosen as the point of view of the analysis (Engeström, 1999). Under **subjects**, I placed the module team in charge of designing the statistics modules (Section 5.3.1). In higher education, the module team might include the lecturer in charge of designing a module, in collaboration with the programme team, heads of school and a Teaching and Learning committee who ultimately approves the programme and module specifications[1]. However, the lecturer teaching the module might not have been part of the design process.

Under **objects**, I included the undergraduate students's learning during a programme of study as the protagonists of the activity (Section 5.3.2). Under **tools** I placed items such as the intended curriculum and content, module staff characteristics (beliefs, goals, knowledge of teaching, knowledge of statistics, knowledge of students' degree subject, knowledge of previous cohorts of students), knowledge of statistics practice in the field, Psychology or Engineering (Section 5.3.3).

My analysis of the **intended curricula** at the macro level focused on the lecturers' curricular and pedagogic planning, the lecturers' goals and beliefs about teaching statistics, about students and the teaching resources. Next, I was also able to link to my data some

[1]A specification provides a concise summary of the main features of the programme and the learning outcomes that a typical student might reasonably be expected to achieve and demonstrate if full advantage is taken of the learning opportunities that are provided.

of the institutional issues such as university regulations and practices. The literature review provided some information about how employers, statisticians or professional bodies might influence the teaching practices in higher education. In this section, I characterise the components of teaching process and content for each node of the macro-level mediational triangle (Figure 5.5).

5.3.1 Macro-level *subject*

Based on my sampling strategy (Section 4.5), I analysed data from two introductory statistics modules, one from a Psychology programme and another from Engineering (presented in more detail in Section 5.3.3). In Gage's linear model of classroom teaching, presage variables, such as teacher characteristics and context variables to do with the teaching situation are the building blocks of teaching (Section 4.2.2). In the macro analysis, I was interested in similar changeable factors of the teaching situation. Along the macro-micro dimension, the wider (large scale) sociocultural context was towards a macro world-view of the situations I was dealing with. Module staff characteristics, being *outside* the classroom, were concerned with institutional organisation rather then the micro lecturing practices of the actions inside the lecture theatre. My analysis of the teaching of Modules A and B started before I went into the classroom, as I gathered information about the module background information, staff and documentation from the virtual learning environment and the pilot interviews (Chapter 3).

The two lecturers had contrasting backgrounds, one being a Psychologist teaching in the Psychology department and the other a Mathematician teaching in an Engineering department. A key member of staff on module A, Lecturer 1 or **L1** was a Psychologist and an experienced lecturer who at the time of this study had been the module leader for over five years. L1's research background was in qualitative methods rather than quantitative. At the time of my study, L2 was contemplating teaching other modules in the future, possibly not quantitative methods.

The lecturer on Module B or **L2** was a mathematician who, due to unexpected circumstances, had to step in to teach the module in the second week of teaching rather than week 6 as it had been planned and was delivering the module for the first time: As a result, L2 had no plans to amend the module materials, which had been developed by someone else in the module team, in any significant way during this delivery of the module. The sudden involvement in the teaching seemed to influence how L2 felt about his preparation with the delivery of the module (as in Quote 4).

> **Quote 4 (L2, Interview 2)**
> "I did not expect that I had to do it [teach the module] which meant that some of it was slightly less well prepared than it might have been, but anyway, life goes on. It was adequate." (0:01:36.9)

In Quote 4, L2 provides his own holistic post-module evaluation of how he *thought* about the circumstances of teaching the module. In this case, macro-level constraints such as availability of staff, timetabling and the module content determined what went on in this classroom, at the micro-level. I became familiar with the content and structure of Module B in the pilot study as I had the opportunity to observe two lectures and two laboratory sessions in the previous academic year (in Section 3.3). Although the lecturer changed during these observations, I was familiar with the current lecturer as I interviewed him in the pilot study (Section 3.2).

Drawing from the module specification and administration, my codes for macro context were therefore about institutional practices and regulations which determined the rules and structure for teaching and also about the unexpected unfolding of the activity in the real setting of teaching on the module. The changes in teaching staff on one module as well as the apparent stability of the other highlighted the responsiveness to the environment and the improvisatory nature of human activity (Lave, 1988). My approach to looking for actions in the teaching activity might mean that my interpretations of the changes and transformations in space and time of the phenomena I was observing and thus unstable and less descriptive of enduring factors. In my study of real, everyday teaching actions, I understood that I could not cling on "objective" and pre-planned actions. The analysis demanded a flexibility in interpretation necessary to understand how the two lecturers engaged in the teaching process (Suchman, 1985).

5.3.2 Macro-level *object*

L1 was not planning any changes to the intended curriculum from previous years since he believed the design was suitable to the students' expected achievement and background:

> **Quote 5 (L1, Interview 2)**
> "at this stage, I have struck the right balance for what I am trying to achieve [...] I am really happy when [the students] are just doing the basics, understand the basics" (0:09:03.8).

The **intended curriculum** was based on L1's personal beliefs about the attainment and experience with previous cohorts of students on the module. Over the years, the module had experienced several transformations and changes as a result of students' backgrounds and attainment. At the time of my observations, the lecturer believed that he had found a suitable approach for his students.

The significant thing about these adaptations to students' background is their history over the life of the module. The lecturer picked up some cues from students' background (e.g. A levels record) and the achievement of previous, comparable cohorts but also cues from informal feedback from students in the lecturer's social interactions with the current students. The lecturer formed an impression of the students from the cues accumulated during a longer period of time. The lecturer's module design is based on these (perhaps fragmented) cues. The lecturer seemed to have formed a representation of the 'average student' using his creativity and intuitions about the students for the benefit of the students.

From this point of view, the object of L1's planning actions had as object "the average student" on that particular Psychology programme.

Module B was aimed at second-year Engineering students from a range of disciplines (Automotives, Materials, Product Design, Management, Manufacturing, Sports Technology). The module was delivered at the same institution and enrolled over 120 students. Each lecture was meant to be delivered by a lecturer and lasted approximately 120 minutes. The statistics module followed a Mathematical Modelling module, which would have covered first and second order differential equations, Fourier series and Fourier transform, so the students on the programme were expected to have A level qualifications in Mathematics, potentially with a statistics component.

The subject of the activity can be studied in relation to the object of the activity to determine how and why the module curriculum (tool) at the macro level is influenced by the programme of studies defines the students' background and the achievement of previous cohorts of students. Although the Psychology programme aimed to give students a grounding in statistics, the 'average' student, as conceptualised by L1, did not in his view match the student profile defined in the specification. For instance, L1 believes that students find the statistical content challenging and as a result the content of teaching needs to be adapted to suit their profile.

> **Quote 6 (L1, Interview 2)**
> "even the good [students] were less comfortable on some of the issues we were introducing" (0:02:52.7) and that "[students] know it's science but they, perhaps expect it to be light as it were [...] because it can cause quite a lot of anxiety. [...] If you look in other psychology department, stats would be taught in a very different way, or maybe more mathematical." (0:18:57.2)

L2 however, teaching the module for the first time, had a much broader conceptualisation of the object of curriculum design. In his own words (Quote 7), L2's planning activity had as goal students' transformation and change mediated by knowledge of how statistics is used in the world (statistical literacy).

> **Quote 7 (L2, Interview 2)**
> "It would be almost desirable from a statistician's point of view because you would be educating these various people who would go and apply statistics, they would turn out more statistically literate then they otherwise would be." (0:45:17.1)

Interpreted through a Vygotskian lens, Quote 7 relates L2's teaching activity towards students' statistical sense making to the concept of ZPD, Section 4.2.1.4. Here, L2 focused on the differences between the students' present statistical knowledge and skills and the outcomes the students might attain as a result of attending the lectures and successfully progressing through the module. The outcomes might be short-term such as module grades and also long-term, such as how students might use statistics in their future employment and lives. There may be different outcomes for a statistics education from a statistician or a lecturer's points of view, which may give rise to developmental possibilities of the activity (Engeström, 2000). In planning his teaching, L2 identified contradictions and tensions in the activity of teaching versus a potential future activity of using statistics in the workplace, discussed in further detail in Section 5.3.4.

5.3.3 Macro-level *tools*

Intended curricula as defined in the module specification and in lecturers' planning and beliefs about students were the tools in my macro-analysis. I considered curriculum and statistical practice to be specialised psychological tools as specialised language (speech), writing systems or number systems and so used as a means of achieving change and transformation by module staff (Foley, 2009). Technicality and abstraction are tools curriculum designers use to explore statistics curricula. The undergraduate students at the programme level (object) are expected to learn this technical and abstract language appropriate for statistics education. The 'new' reformed curricula in statistics education (Section 2.4.2) proposes that curricula needs to reflect how statistics is used in practice within the students' field of study as knowledge of 'specialised registers' (Daniels, 2001).

5.3.3.1 Intended curricula

In the data analysis, I chose to focus on Modules A and B as they covered similar content in univariate inferential statistics. Table 5.1 compares the structures of Modules A and B. As my focus was on characterising the teaching of statistics at university in "context", similarly to my Pilot study, during these observations I focused primarily on the lecturer. This meant that, at least initially, the data from laboratory sessions was less relevant since the lecturers were not present. Table 5.1 briefly summarises some of the characteristics of these two modules. Further, in Table C.1 in Appendix C.2 (page 290), I summarise the lectures I observed for the modules included in the data analysis.

Module A

Module A was an introductory level, first year twelve-week statistical methods module aimed at Psychology students at a research intensive university. The module was delivered in the second term enrolled around sixty students. Each lecture was delivered by a Psychology lecturer, **L1**, and was planned to last sixty minutes, followed each week by a practical session and *stats clinics* that were run by two teaching assistants. Following institutional regulations and the module specification, the module was taught over twelve weeks.

In the first semester, the Psychology students would have covered concepts in inferential statistics such as variation, chi-square, correlation and regression. L1 designed the module in the second semester to include two further topics in inferential statistics (week 1), Student's t-test (two lectures weeks 2-3) and analysis of variance (ANOVA, five lectures weeks 4-8,

Table 5.1: A description of Module A and Module B

	Module A	***Module B***
Domain	Psychology	Engineering
Year	1	2
Lecturer	Psychologist (L1)	Mathematician (L2)
Teaching the module	5 years	First time
Years of teaching experience	8 years	3 years
Students	60	130
Gender	approx. 80% female	approx. 75% male
Students	Future Psychologists	Future Engineers
Lectures	11 (Weekly)	11 (Weekly)
Duration of lectures	1h	2h
Laboratory sessions	11 $\times$ 1 hour	4 $\times$ 2 hours
Stats clinic/tutorial	11 $\times$ 1 hour	None
Lectures on t-tests	2	2
Lectures on ANOVA	5	1
Lectures on other topics	0	6.5
Revision sessions	2	1.5
Assignment	50%	20%
Exam	50%	80%
Main lecturing resources	slides	slides/print-outs
Teaching own resources	yes	no
Room	Laboratory	Lecture theatre

as in Appendix C.2). The emphasis in the module was on experimental design, inferential statistics using software and examples of applications in real-life or realistic applications. The remaining three sessions were spent on revision (weeks 9 and 10) and a class test (week 11).

The linear ordering of content through the module was accompanied by two summative assessments, a mid-term assignment and a final multiple choice test. In this case, I had codes relating to the macro context for the intended curriculum: module specification, teaching resources planned for the delivery of the module and sequencing of content. At a macro level, I was mindful of the beliefs of module staff about the students who normally enrolled on the programmes. For example, Module A's scheduling was from 9am to 1pm each week "to suit the first year university students" and ensure their attendance (L1, Interview 2).

Module B

Module B was a twelve-week introductory statistics methods module delivered at the same institution and enrolled over 120 students. Each lecture was delivered by a mathematics lecturer, **L2**, and was planned to last 120 minutes. The module offered three one-hour laboratory sessions run by a laboratory assistant where students had the opportunity to solve statistical problems using software, but there were no tutorials.

In Module B, each lecture sequentially aimed to cover a new topic (also see Appendix C.2). The content of the module focused on a range of topics, starting with probability, distributions, descriptive statistics, hypothesis testing, t-tests, contingency tables, regression, correlations, ANOVA and finally experimental design. Lectures six and eleven were revision lectures. The emphasis during these lectures was on mathematical explanations of concepts and hand calculations with answers provided during the lectures. The last lecture was spent on revision.

5.3.3.2 Module planning and lecturers' beliefs

For analysing module staff planning and beliefs, I relied primarily on the Pilot study interview, interview 1 and the post-module interview, interview 2 with L1 and L2. For Module B, since I observed lectures and tutorials in the Pilot study, I also had access to documentation, such as handouts, before starting my observations of teaching. Here, I discuss each module's planning and beliefs about teaching, content and the students in turn.

Module A

L1 was mindful of students' feedback and planned his teaching at 'a medium high' level for the "average student" (in Section 5.3.2), meaning that students were expected to learn *basic* principles (Quote 5, p. 155). L1 saw the strength of his teaching in offering very limited mathematical challenge to the students but emphasising experimental design principles relevant to Psychologists and how these principles linked to statistical concepts.

In the pilot interview, one of the constructs elicited (Section 3.2) by L1 contrasted "knowledge underpinning the [statistical] analysis" *versus* "decision making", as in Quote 8. For L1, *knowledge* seemed to be about justifying an approach to a type of data analysis and a tool for students to use. My interpretation is that "knowledge underpinning the analysis" referred to scientific or schooled concepts which may remain tacit and only partially revealed to the students (Section 4.2.1.5). On the other hand, "decision making" was about statistical reasoning, the creative narrative of the analysis based on statistical rules. There seems to be a distinction here between traditions of doing statistics in school or as statisticians do and everyday experiences Psychologists might have.

Quote 8 (L1, Interview 1)
Knowledge underpinning the analysis - *It is theoretical knowledge that feeds into using the techniques, the tools*

versus

Decision making - *It involves justifying your approach, style of presenting (according to a pre-defined structure or template). It involves both procedural and creativity, there is a choice that the student makes, fit narrative into conventions; justification, putting the narrative together.*

Bernstein (1999, Section 4.2.1.4, p. 123) distinguishes between two forms of discourse, vertical and horizontal, according to the forms of knowledge realised in the two. In my interpretation of Quote 8, 'decision-making' is a form of horizontal discourse and 'statistical knowledge' is a form of vertical discourse. Further, this also reflects some of the purposes with a statistical analysis of describing, deciding and predicting from data.

L1's main teaching goal was that the students should have an understanding of experimental design and how to select a suitable statistical technique and interpret the output: 'we are trying to put statistics in a new context, trying to see that when faced with a 'context' or scenario, students can use a t-test or ANOVA' (Interview 2, 0:08:09.8). The module team replaced certain components of the statistical process, such as mathematical calculations and data collection by the exclusive use of statistical software using existing data-sets (Quote 9).

Quote 9 (L1, Interview 2)
"I find mathematics relevant, I know that it is all about mathematics at the finer level, but in my teaching, personally I do not feel the need to include it. And as a group we agree on this, the teaching strategy of doing it this way." (0:13:37.9)

L1's intentions appeared to be to show students a conceptual way of understanding statistical concepts and topics (Section 2.4.1.5) in order to help students approach statistics like other psychologists would: 'logically not mathematically', by using software and a deep understanding of their research design and research question. In both interviews 1 and 2, L1 perceived the statistical analysis as 'easy' if students engaged with the practical exercises, but not when students were required to apply different techniques in different contexts, as illustrated in Quote 10.

> **Quote 10 (L1, Interview 2)**
> "We feel it is important that [students] get transferable skills, [...] trying to think why you would use [a statistical test] in context, rather than automatically pressing buttons. At least I hope that this is what we do." (0:18:57.2)

L1's short term goal for the students was for them to be able to carry out a t-test and an ANOVA when presented with a dataset, but also, in the longer term, to link different statistics curricula to future Psychology projects or employment. In interview 2, Quote 11, L1 expressed the importance he placed on experimental design principles as a prerequisite to making sense of statistical concepts.

> **Quote 11 (L1, Interview 2)**
> "What I am interested in is the logic; if [the students] get the logic of design first and then the logic of how you have to think of variance that makes it easier for the student to understand what is going on or the value of t and the significance level." (0:11:48.2)

The teaching style on the module was expected to be 'formal' (Quote 12). Using the contrasts suggested by Bernstein (1999, p. 158), my expectations for L1's enacted teaching discourse was a dualist focus on operations versus principles, statistical (outside) context versus everyday (inside) context, a social discourse that is between intimacy and distance and a lecturing voice that alternated between dominated and dominant.

> **Quote 12 (L1, Interview 2)**
> "Even the communication between the lecturer and the students, even some of the communication, some of the lectures tend to be formal." (0:34:59.1)

In Quote 12, L1 reveals a contradiction between his aim to use active learning as a teaching tool and the reality of the module which imposed a more traditional methodology.

Module B

Similar to L1 (Quote 8), in interview 1, L2's main issues for teaching statistics at university revolved around mathematics and everyday context. In Quote 13, L2 contrasted "understanding the real world problem" versus "doing statistics", i.e. much of the scientific knowledge that is the focus 'inside' university teaching with statistics 'outside' the module that is used in the workplace to solve everyday problems. The contradiction between "university" and "workplace" statistics suggested in Quote 13 reflects Bakker et al.'s (2008) finding that formal statistical inference is about populations and focuses on induction while workplace statistics is pragmatic and focuses on processes and action. In the workplace, contextual reasons might need to overpower statistical ones. The tension for L2 is about how to teach contextualised, everyday horizontal knowledge in the context of the lecture theatre, focused on vertical discourse. One important aspect of the macro context in the teaching of statistics is the lecturers' position in relation to the object of the activity mediated by knowledge as tool. L2 expected a change in the relation between 'informal' statistics that students might have and formal knowledge.

> **Quote 13 (L2, Interview 1)**
> **Understanding what the real world problem is** – *translating [the real world problem] into maths equivalent; ability to think about the world in a mathematical way, deciding what to apply. Some people may be able to do this, some may not. Real world interacting with the maths world, applying statistical ideas to real world context*
>
> **versus**
>
> **Doing the stuff** – *going behind what test to use, more of the detail, what to do.*

Vygotsky suggested that the educational process should be considered not only in terms of curricular content but also as possible sources of students' progress (Kozulin, 2012). L2 desired for the students to develop the 'ability to think about the world in a mathematical way' and 'apply statistical ideas to real world contexts' suggests to me Vygotsky's idea of evaluation of ZPD rather than already formed abilities (Section 4.2.1.4). Vygotskian theory transcends educational stages since processes of learning and development of statistics (and mathematics) transcend developmental and phase boundaries. In the data analysis, the interview data suggests that the lecturers believed that students come from school with particular mathematical and statistical culture (e.g. procedural) and expect them to develop towards statistical literacy, reasoning or thinking (e.g. Quotes 7, 13).

L2 expected students to regard statistics as 'a normal thing to do' and 'be part of a relatively normal everyday conversation' (L2, Interview 2, 0:03:45.9), thus indicating the focus in his teaching on 'statistical literacy' (Ben-Zvi and Garfield, 2004). For Vygotsky (2012), the teacher's role was to organise and prepare a "natural" transition from the everyday drawing (or play) to mastering meaningful writing (or the scientific). L2 similarly observed in Quote 13 the challenge faced by statistics educators in scaffolding (p. 119) the learning for the students to help them progress from "everyday" concepts to mastery of a system of statistical symbols and signs.

Vygotsky (2012) explains the teaching of writing as uniting drawing and play to the written and spoken symbols in a way that is relevant and meaningful to students (Section 4.2.1). In this view, the teacher needs to bring the students to an inner understanding of statistics. The lecturer needs to ensure that statistics is *organised development* rather than learning. Teachers organise all the preparatory stages or actions and the entire process of transition from one mode of understanding to another, e.g. from drawing objects to writing letters to drawing or as in Quote 13 from understanding what the real problem is about to 'doing the stuff', i.e. going through the statistical investigative cycle. Although the expectation was not that students would come onto the module as 'tabula rasa' ("blank slate") and that they had some preexistent knowledge of statistics or the context of their own disciplines (e.g. Engineering), formal teaching needed to account for students' current development and help them continue make progress through both concrete and abstract representations. Haenen et al. (2003) suggests that in 'good' classroom conversations 'the students experience the boundaries of a concept and have the opportunity to specify the concept further'. Statistical concepts thus might emerge through 'conversation' or reflections on the scientific knowledge experienced in the lecture theatre.

L2 regarded statistics as a life skill, but not necessarily a subject students used immediately after graduation, as in Quote 14.

Quote 14 (L2, Interview 2)
"You can never predict just because you have not used something in your first three months of a placement it does not mean it is not going to be relevant in the next three months or three years or however long." (0:27:28.5)

In this case, L2's experience with and knowledge of statistical practices determined his views on the curricular processes on the module. It seemed that L2 was interested in expanding students' knowledge and understanding into the future, beyond the immediate present needs of the module or programme, as part of a general academic education. Thus, considering the students' mathematical background (having studied at least two other mathematics modules previously), he expected them to engage with hand calculations.

A previous member of the module team and a participant in Pilot 1 interviews (Section 3.2) described t-tests as 'tedious detail' and less important. In this participant's view, important concepts for students to learn were understanding what is a p-value, understanding difference between main effects and interactions in an analysis of variance, the difference between inferential and descriptive statistics and how to interpret results of statistical tests. These strategies appeared to form the basis for the module specification at a macro level.

Both L1 and L2 agreed that having too many examples from outside the students' programme of study may not help students see the relevance of statistics to their studies or future profession. As a mathematician, L2 tended to believe that giving students a good background in the underlying statistical methodology was more important than an emphasis on real-world problems. When I asked L2 whether he could see a situation where one statistical module would meet the needs of very different disciplines, he answered that it would be desirable from a statistician or mathematician's point of view. Quote 15 suggests that such a module may however not necessarily be a popular choice for departments.

> **Quote 15 (L2, Interview 2)**
> "[Students] would get taught only useless theory and they won't know how to apply things in their discipline they are meant to be at university to learn about." (0:48:15.0)

This idea of 'useless theory' in Quote 15 was striking to me, so I went back to my interview data from the two participants. The participants seemed to imply that some content may or may not be relevant to the students or to their disciplines, which was reminiscent of Vygotky's view of 'good' teaching which needs to be necessary, relevant and taught naturally (Section 4.2.1). To L2, it was important that students learnt statistics and mathematics for practical purposes, have awareness of different methods and their limitations (Quote 16). In his view, a mathematician has an understanding of the underlying mechanisms, whereas a non-mathmatician/statistician would not have. However, his assumption was that knowing the required mathematics can help dealing with the everyday "context". So the category 'useless theory' I expanded to a 'black box'. A 'black box' seemed to include the underlying mathematical principles which are not required when applying statistics in practical situations using software.

> **Quote 16 (L2, Interview 2)**
> "[Students] just need to know that there's a **black box** there that they can use. And if they want to compare two different groups or whatever, then they can apply a t-test to that situation. And if they want to look at flows of rivers, and whatever, then a statistical differential equation that they can use, they don't need to know complicated existence and uniqueness theorem about differential equations, just the same as they don't need to know about the derivation of why this test statistic follows a t distribution or why it follows another distribution." (0:41:12.2)

In interviews 1 and 2, L2 believed that there should have been more emphasis on using software for analysis. In Quote 17, L2 recognised the challenge with teaching statistics which, in his view, should involve both a mathematical understanding of the material and also working with larger datasets using software and focusing on interpretation of statistical outputs. It seems that in his view, the challenge for curricular planning is what statistical content and how to make the statistics visible to the students and turn 'black boxes' into 'open boxes' for the students, who, at least initially, are perceived as outsiders (Williams and Wake, 2006).

L2 challenged the belief that students, as novice statisticians, could use software as a 'black box', with no understanding about what goes on in the analysis. In my interpretation, it seems that some understanding of the mathematical bases of statistical models could complement the use of software 'as if it was a black box'. In this case, I noted a tension between the planned, intended curriculum on Module B, and L2's beliefs about teaching and about statistics.

Quote 17 (L2, Interview 2)
"If I was to be involved with the module for longer, then there should be less emphasis on hand calculations and more emphasis on getting the computer to do something for you. But obviously being able to interpret the output as well. Even so, I think there is still some advantage in doing some hand calculations so you understand what the computer is doing for you. You do not want it to be a complete *black box*, the computer, but you need to be able to... Perhaps an ideal outcome would probably be that you can use the computer as though it is a black box, but equally, if you are in the right mind set, or the right circumstances, you can lift the lid of the black box and understand a bit of what goes on inside if circumstances were appropriate." (0:12:30.4)

L2 was also mindful of students' misconceptions and difficulties which he would have liked to address in his lectures. For example, L2 believed that students found difficult and needed more practice in turning a real world problem into a mathematical problem, as in Quote 18. Here, L2's focus was on helping students discern relevant from irrelevant information in a statistical analysis and interpretation.

Quote 18 (L2, Interview 2)
"Recognising what statistical procedure needs to be applied, what bits of information [the students] are given, they need to plug that into a formula and then they very often get extremely confused if there is any spurious information given as well. [The students] want to know, they want to pick out what information is relevant. And it is probably something that students need a lot more practice at. But is is difficult, I know." (0:23:22.7)

L2 also recognised the challenge in helping students achieve this aim, given current module curricula (as in Quote 19), structured around statistical procedures, and the teaching sequences through the module. Here, there seemed to be a contradiction between doing statistics in real-life, outside the module at the macro level and implementing the statistics curriculum in the lecture theatre, at the micro level.

Quote 19 (L2, Interview 2)
"Because you [the lecturer] only present a situation which is appropriate for the procedure that you want to model in five minutes' time." (0:26:14.6)

However, given the unexpected planning of Module B, L2 believed that he followed the same strategies as for previous cohorts. In Quote 20, L2 provided his own holistic post-module evaluation of how he *thought* about the circumstances of teaching the module. In the circumstances of Module B, L2 was using tools (teaching resources) designed by module staff over a number of iterations. While the planned teaching and learning resources were similar, I was interested to see how L2 implemented the curricula during the module (in Section 5.4).

Quote 20 (L2, Interview 2)
"I just took over the module and ran it again as much as possible in exactly the same way it was run last year" (0:10:40.7).

5.3.4 Macro-level *contradictions*

My qualitative data analysis was an iterative process involving several layers of analysis of the transcripts from interviews and lectures with their respective teaching resources. In the final stages of the macro analysis, I aimed to extract broad themes from the categories that emerged through my analysis. The analytic strategy was consistent with my approach in the Pilot study of reporting broad themes and categories (Section 3.5.4, p. 92). By weaving together the outcomes of the Pilot study and the Main study macro analyses, I was able to pay attention to and carry out a comparative analysis of the contradictions and tensions (Section 4.2.1.4, p. 118) that were revealed in the lecturers' accounts of their experiences with teaching statistics on the two modules. The analysis of contradictions also aimed to identify opportunities or circumstances for change and transformation in the teaching activity (Engeström, 1999). To explain lecturers' development in context, I present this part of the analysis with reference to the three themes that emerged from the Pilot study.

5.3.4.1 Teaching in "context": statistical versus contextual knowledge

From a SCT perspective, lecturers' beliefs (Section 2.5.1) about mathematical, statistical or contextual knowledge represent mediational tools that shape the teaching activity (Section 4.2.1.3). Thus the teaching activity needs to be analysed in terms of the tools employed by lecturers (subjects of the activity) towards students' statistical sense making (objects of activity). Teachers' beliefs about the relationship between mathematics and statistics was researched in the case of school mathematics. Begg and Edwards (1999) found for example that teachers believed that mathematics is not needed for a good grasp of statistics and that statistics gives meaning to mathematics. By contrast, L1 and L2 planned their teaching to address their students' perceived level of mathematical preparation for doing statistics.

In Chapter 2, I discuss how Shaughnessy (2007), Cobb and Moore (1997) and Wild and Pfannkuch (1999) emphasise the importance of the contextual in statistics and a view of statistics as fundamentally different from mathematics. L1 similarly believed that a lack of interest in mathematics should not prevent students from grasping 'basic' statistical concepts, although he agreed that the module content can cause high levels of statistical and mathematical anxiety in students. Context knowledge as a tool was used to demonstrate to students that is is possible to think about data without mathematics and as a motivating factor (Langrall et al., 2006). L1 took into account students' prior knowledge of mathematics, which from a vygotskian viewpoint, showed that what students learnt on Module A had a previous history and started with students' *everyday knowledge* and experience with statistics within Psychology (Section 4.2.1.5).

L1 conceptualised context as being about how one uses statistics in real life. In his view, understanding statistics in *context* needed to include how one uses *context*, the background information required to carry out a task, the dataset, the mathematical content, experimental design issues and their impact on interpreting the statistical information presented (Interview 2, 0:05:58.0). In his view, statistics was a tool for use in another subject area and students were required to build an understanding of statistics *outside* of statistics (Gal, 2000): “students should understand the general logic and apply the context, rather than formulae” (Interview 2, 0:13:37.9). In L1’s view, students first needed to understand the dataset, the research question, followed by which statistical technique to apply. In this approach, L1 suggested that students being guided by the lecturer, gradually become more able to abstract and generalise from concrete, everyday perceptions of objects. Further, students required ‘a questioning attitude’ that involved statistical knowledge and also contextual knowledge (Wild and Pfannkuch, 1999). Statistics was therefore a tool for understanding Psychology, since with statistics, students “got more than just statistics, got more about design, about putting things in context” (0:31:14.4).

For L2, ‘context’ was perceived to be a teaching device “to make the lecture a little bit like an informal conversation” (Interview 2, 0:03:45.9) and to motivate students while carrying out hand calculations (0:12:30.4). For L2, one way to conduct and interpret the statistical information was therefore to present a basic understanding of statistical terminology, language and formulae while embedded in a ‘very brief’ context. L2 believed that students’ difficulties with statistics are due to the movement from a real-world context which might contain spurious information, recognising what statistical procedures need to be applied and carrying out the statistical calculations. In this view, with statistical knowledge, L2 challenged the students to recognise *relevant* information within the real-world *context.* In order to improve students’ sense making, L2 believed that the focus in introductory statistics should be on“turning a real world problem into a mathematical problem” (0:23:22.7). From a SCT lens, L2 believed that statistical or “true” concepts represent scientific knowledge that needs to be systematically learnt or developed during teaching and learning experiences in the lecture theatre or laboratory (Vygotsky, 1987).

Both lecturers focused on students’ (objects of the teaching activity) statistical understanding in context as key to interpreting the statistical information, interpreted as scientific knowledge, planned for the module and relevant to students’ lives or subject-domain, e.g. engineering and statistics (Watson, 2000). This view of understanding concrete context versus understanding abstract statistics supported the finding from the Pilot study which identified a tension between the theoretical basis of statistics and the application of statistics in context.

5.3.4.2 Teaching of the statistical process components: knowing about versus doing

A contradiction also emerged between the students’ sense making (the object) and the statistical process components (mediational tools). The tools employed in the teaching activity appeared to be inadequate in helping students with diverse learning needs make sense of the statistics. In dealing with context, both lecturers referred to the activity of carrying out statistical analysis, reflected in the process components, from data collection to interpretation. However, there seemed to be a contradiction between the teaching activity and the statistical analysis activity. For L1, there was a contradiction between interpretation, decision making and *doing* the analysis. L1 stated that tasks were similar in terms of the "calculations in the background" and what differs are the contexts in which problems are set. In the case of the Psychology module, the contexts were about people rather than machinery, as for the Engineering module. This is a view that data analysis seamlessly is a process of "changing representations to engender understanding" (Wild and Pfannkuch, 1999, p. 227). However, the need to understand the statistical process components lead to a tension between moving students from using computers to do the statistical analysis to thinking in a statistical way.

For L2 however, it was important to first think in a statistical way and only then use computers to assist with the statistical analysis. L2 used a metaphor to refer to the use of software "as a black box" (Quote 16, p. 164) and believed that students should understand *something* about the statistical calculations before using software. The challenge with transitioning from abstract concepts to concrete applications is that statistics is taught in concepts isolated from each other and also from the context and statistical process components in which they can be applied (Bakker and Derry, 2011). Vygotsky proposed the idea that it is possible to use a concept before fully grasping its meaning (Section 4.2.1.5). Concepts can be developed through activities in which the concepts gain meaning. The learning of concepts need not be formal since it can take place within the domain of activity in which they function (Bakker and Derry, 2011).

In a vygotskian sense, word meaning is acquired in activities mediated by others gradually, through a series of actions and part of a system of coordinates (Section 4.2.1.5) in which the concept is used (Derry, 2007). Although L1 focused on everyday knowledge before statistical, he also recognised that two concepts may apply to different areas of reality but may be comparable in degree of abstraction or generality (Vygotsky, 2012). Although L1's teaching aimed to focus on applications of statistics to Psychology, statistical processes could be expressed and interpreted in a number of ways, depending on the context in which they were used.

By contrast, L2 believed that teaching statistics should be an accumulation of statistical concepts, starting with the most simple, as a transition from the abstract towards the concrete, similar to the *representationalist approach* that follows a topic-by-topic approach in which statistical concepts are taught in separate sections, from simple to more complex (Bakker and Derry, 2011). In L2's view, students required an understanding of terminology to be able to then embed statistical language and concepts in a wider context or question claims made based on the analysis (Watson, 1997).

L2's view of learning statistics was different from L1 who promoted a transition from conceptual (with software) towards interpretation (without software). In L1's approach to planning curricula, the order of teaching was governed by the increasing complexity of statistical models and also by the reasoning steps that lie in students' ZPD (Section 4.2.1.4), in what Bakker and Derry (2011) termed the *inferentialist approach.* L1 promoted the use of spontaneous concepts to support students' interpretations even in the absence of the scientific concepts (Quote 11, p. 161). In Section 5.4, the analysis will show how the implemented curricula on the two modules differed in how the lecturers transitioned from the abstract ('knowing about' statistics) towards the concrete ('doing' statistics) or the reverse, from the concrete towards the abstract.

5.3.4.3 Student learning: what students *should* do versus *can* do

Beliefs about teaching as ascending or descending towards the concrete were also motivated by lecturers' beliefs about the learning outcomes students were able to achieve in the teaching and learning time available. A third contradiction in the teaching activity was between the lecturers' intended curricula as a mediational tool and the attained curricula, interpreted as the object of the teaching activity. *What students should do*, I interpreted in terms of lecturers' beliefs about intended curricula. Also, *what students can do* represented the attained curricula, students' statistical sense making as a result of engaging in the teaching activity.

L1 made his longer term goals for the students explicit (e.g. be able to use statistics for a final year project or as Psychologists). L1's longer term goals were in contradiction however with shorter term goals of addressing students' low prior attainment in mathematics and anxiety with statistics. Similarly, L2's longer term goals for his students were to think about the world in a mathematical way and to apply statistical ideas to real world contexts (statistical literacy). L2's view of statistics as a relevant academic education with students' activity in the lecture theatre. For L2, statistics was relevant regardless of students' programme of study since it was 'a normal thing to do'. However, L1 believed that what statistics is relevant depended on the programme of study, students' backgrounds and motivation or their future employment needs. These apparent differences in lecturers' beliefs may have been linked to their different teaching approaches for motivating students (Section 2.5.3.4) discussed below in Section 5.4.3.2.

From a SCT viewpoint, statistics needs to be relevant or 'necessary for something' in order to motivate students (Section 4.2.1.2, p. 112). L1 considered that relevant context was a preparatory stage in students' sense making, while L2 viewed mathematical knowledge as a preparatory stage in students' development of statistical literacy. For L1, the teaching of statistics involved the guiding students to view context as relevant to current and future contexts rather than mathematics, while L2 was closer in his beliefs about statistics to Gordon's (1993) position (Section 4.2.1.2). Since actions and interactions across contexts affect the teaching activity, the analysis of the interview data indicated that different approaches to teaching statistics are motivated by lecturers' beliefs, norms and values in interaction with students' learning of statistics to create ZPDs (Bakker et al., 2008; Akkerman and Bakker, 2011). Finally, these beliefs about teaching as ascending or descending towards the concrete were linked to lecturers' views of how to address the contradiction between what students can do based on their perceived motivation and preparation for doing statistics and what 'relevant' statistics students should do, the intended curricula.

5.3.5 Macro-level model of teaching

Based on Vygotsky's work, Cole (1996) and Cole and Gajdamaschko (2010) distinguish between notions of context defined as nested within each other and also as weaving together (Section 4.2.1.5). The context of behaviours and ways in which cognition can be related to that context are created simultaneously by the combination of subjects, objects and tools in a setting (Cole, 1996). Throughout my macro-micro analyses, I looked to interpret relationships between the nodes of the activity system (Figure 5.5). The aim of the analysis was to identify implications of the active construction of context in teaching actions to do with planning curricula (Daniels, 2001).

At the macro-level, I considered what teaching actions were taking place before and after teaching the modules from the lecturers' point of view, but considering other members of the module planning team who, at some point, might have affected the planning actions. Some of the historical changes of the two modules became apparent during my pilot observations and interviews, conversations and post-module interviews with the two lecturers. In my interpretations of this data, I concluded that the two lecturers, as subjects of the activity system, had complex representations of the body of students (the object of the activity) and their history with them mediated by tools such as the lecturers' (and module staff) knowledge of statistics and also of aspects of teaching and beliefs about teaching, students or statistics.

On Module A, L1 informed his planning of the module over a long period of activity. L1's *micro-experiences* with previous cohorts of students and with implementing the curriculum, influenced the macro-planning. The tools at L1's disposal were not new, they changed over a number of iterations. Engeström and Middleton (1998) maintain that tools 'come to embody the stable and structural work practices', however that they are 'not just there' (p. 4). Engeström and Middleton propose that tools are invented, borrowed (purchased) and put into use, then they 'wear out' and are disregarded and replaced by new ones. In the interviews, knowledge of statistics is constructed by the relationships between the object of the activity system (undergraduate students on the programme) and other objects (past cohorts of students).

On Module B, interviews from L2 suggested that he was planning to teach 'in exactly the same way' as in previous iterations of the module. However, on Module B, there seemed to be a dilemma of teaching action: of stable routines and skilled development of teaching tools and the disruption, innovation and change in module team. A participant in the Pilot study interviews, who had contributed to Module B's development, was able to articulate in his own terms his intentions for the teaching of statistics. However, previous expertise, beliefs and intentions, might not have been immediately obvious or known to L2. A module team's ongoing collaboration and construction of teaching tools might be disrupted when team members change or when the body of students, the object of the activity, changes. In the analysis of local module planning practices, I noticed 'a whole social network of resources beyond the confines of the setting' (Laufer and Glick, 1998). The intended curricula on these modules can be explained by looking at how subjects, objects and tools connect to other different subjects, objects and tools in space and time.

Although module specifications are intended to specify 'what' to teach, the lecturers formed their own representations of 'how' to teach, the process of teaching. For L1, a Psychologist, everyday concepts within the real-world context of statistical problems and how to apply statistical models rather than their mathematical underpinnings were more important. L2, a mathematician, was more ambivalent in 'how' to teach to 'lift the lid of the black box'. In L2's case, a conceptualisation of the object of activity was broad, based on macro-experiences (previous teaching of other modules aimed at different students, the module specification). L2's knowledge of the history of changes on the module and more limited social interactions with the module staff and previous cohorts appeared to influence his belief that he had *less* control over the content of teaching and learning than L1.

The link between the nature of knowledge and how to acquire it is important in explaining the lecturers' pedagogy and their creation of forms of teaching. Vygotskian ideas of developmental teaching considered that teaching can provide information and concrete skills but also promote cognitive development through a good quality intended and implemented curriculum. The lecturers seemed to believe that the purpose of statistics education is to create in the student a system of statistical and mathematical concepts and also incorporate statistical literacy, reasoning and thinking.

In university statistics, the subject matter may be used as a tool to affect changes in students, whereas in the workplace statistics is used to solve everyday problems. For L1 and L2 (Quotes 8, 13), 'doing statistics' was about translating the horizontal everyday discourse into vertical scientific (statistical/mathematical) discourse as a process of (re-)contextualisation. The emphasis in the interview data is on the teaching and learning of statistics within authentic and meaningful events. The importance of the interplay between the scientific concepts in statistical theory and the spontaneous, everyday concepts in the real-world "context" of the statistical problems seem central to the lecturers' conceptualisation of teaching and learning of statistics. If the scientific cannot connect with the everyday concepts, then learning of statistics breaks down. For the two lecturers, it was important to plan module content that connected meaningful real-world contexts with theoretical bases of statistics. The challenge for planning the teaching of statistics was to collate their own idiom/language within their own discourse, which results from lecturers' attempt to impose multifaceted sense on a particular word/statistical concept. The lecturers' inner speech, where one word can stand for a number of thoughts and feelings, can be substituted in intended as well as implemented teaching for a long and profound discourse.

5.4 Micro-level analysis of context

The second step was to analyse the enacted or implemented (Figure 2.2) teaching on the two modules. For the micro analysis, I also filled in each node of the mediational triangle with empirical data from the pilot observations and I refined my interpretations during the data analysis. This process of analysis resulted in the items in Figure 5.6.

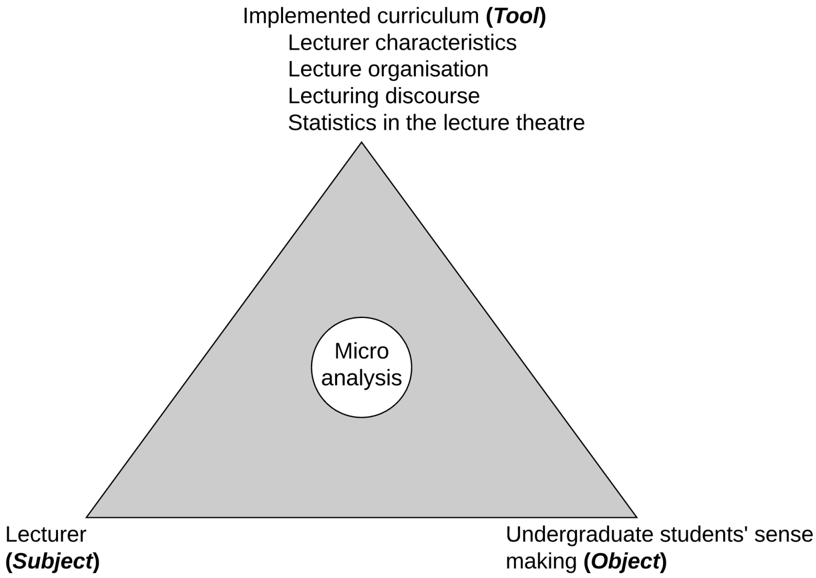


Figure 5.6: Micro analysis of teaching

Units of analysis (Section 4.7.2.2) are the means by which the analysis undergoes functional reorganisation and needs to be a functionally integrated whole while simultaneously allowing for internal contradictions and heterogeneity that is the catalyst for development, change and transformation (Lee, 1985; Zinchenko, 1985). The micro analysis was concerned with the tool mediated activity of the individual lecturers acting in the context of the lecture theatre. The micro unit of analysis therefore was therefore composed of individual lecturers and the tools they used to allow a focus on different aspects of the activity of teaching from the macro analysis. Thus, the micro unit of analysis included important links between the subject's activity towards the object mediated by tools. Under **subject**, at the micro level, I included the lecturer teaching in lectures (Section 5.4.1). The **object** included the students' sense making inside the lecture theatre (Section 5.4.2). Under **tools**, I placed the implemented curricula, the cultural practice of using curricula in the teaching activity to represent meaningful statistics (Section 5.4.3).

5.4.1 Micro-level *subject*

At the macro-level, the module team occupies the subject position (Section 5.3.2). In the next action of teaching in the lecture theatre at the micro-level, members of the module team can be the subject. In my analysis, L1 and L2 were the subjects of the teaching actions in the lecture theatre, each with specific goals and intentions for dealing with and implementing the intended content. In the macro-analysis, the lecturers revealed a complex set of beliefs about the purpose of teaching and learning at university and the body of students. Inside the lecture theatre, at the micro-level, the lecturers came prepared with specific teaching plans and goals for each lecture.

5.4.2 Micro-level *object*

The students learning and making sense of the statistics in the social context of the lecture theatre and during social exchanges were included as the object of lecturers' activity. Students' attendance in Module A was fairly stable in lectures at around 45-50 students each week. The lecturing hall was organised with round tables of about eight students. L1, reflecting on the teaching of the module, expressed the view that a good lecture is when 'students make an effort, rather than the lecturer talking at the front of the class' (L1, Interview 2). At a micro level, L1 expected students to engage with the module through exercises, assignments and attend regularly, even when students required additional support with the mathematics.

L1 showed a deep concern for the students' views of statistics and was keen to help them succeed. However, L1 also acknowledged the challenge of turning short term goals in the module into longer term effects for the students' learning. For example, from my observations, L1 would spend from one to five minutes in each lecture advising students on study skills and attendance to improve their performance on the module.

In Module B, out of around 130 students registered, about 80 students attended each week. In interview 2, L2 felt that he could not engage the students in a 'normal conversation' given the number of students and hoped that the students gained 'a bit of an appreciation that doing some statistics is a normal thing to do' rather than 'spouting stuff at them'. L2 considered important to stretch and challenge the students, "get students off balance", and ask them questions that they were not expecting in order for students to get "an education when they are at university rather than just turn up, learn things and get a qualification" (Interview 2, 0:26:14.6).

In my data for both modules, the students were present in lectures, were taking notes, making eye contact, smiling, nodding, whispering to each other or getting distracted by social media. Verbal interactions lecturer-students was mainly in informal conversations outside lecturing time. Direct interactions between lecturers and students were also taking place using electronic mail or in one to one meetings in their offices. As I was observing the lectures and tutorials only, it is possible that I could not discern all the elements of these interactions in my analysis.

5.4.3 Micro-level *tools*

In my observations of lectures, in the SCT perspective, tools featured prominently since learning was seen as a process of "appropriating tools for thinking" (Rogoff, 1982; Renshaw, 1996). In lectures, tools were made available by lecturers who (initially) acted as interpreters and guides in the students' cultural apprenticeship. Also, the actual means of social interaction (language, gestures) were appropriated by the students, internalised and transformed to form tools for thinking, problem solving and remembering (Wertsch and Stone, 1985). L1 delivered all the lectures in Module A using slides, with the lecturer standing at the front of the class and students sitting at round tables, resourced with a laptop computer for the practical session only. L2 delivered the lectures in a similar fashion to L1, using slides and standing at the front of a large lecture theatre, occasionally writing on the whiteboard.

Similar to the macro analysis, curriculum was interpreted as a psychological tool (Section 5.3.3). In my dataset, I had access to the lecturing resources (slides, handouts) and the transcripts of lectures. Thus the tools in my analysis were lecturers' implemented curricula as represented in lecture resources – written slides or handouts and in lecturers' teaching discourse – the lecturer talking about statistics. The micro tools included chunks of lectures interpreted as temporal sequences of content, pedagogic routines and statistical routines. These tools thus mediated a movement from the social place of functioning to the individual place of functioning.

I considered that a *lecture* is a ritual or technique at lecturers' disposal, and hence a tool in the lecturing activity (Vygotsky, 1978b). Bligh (1972) draws attention to the importance of lecture organisation. In my **first analysis** therefore, to characterise the types of lectures used on these modules, I considered the organisation of lectures in terms of how the content was structured and the timing of each routine (5.4.3.1).

Other lecturing techniques or routines discussed by Bligh (1972) include how lecturers go about giving explanations, using handouts, different styles of lecturing and ways of obtaining feedback about student learning or student opinions about the teaching. Viirman (2014) further described and analysed mathematics lecturers' presentation of mathematical routines or discourses. I use a similar strategy to categorise lecturers' pedagogic routines in my **second analysis** (Section 5.4.3.2).

Significant differences between mathematics and statistics (Section 2.4.1.5) mean that the routines used in Viirman's study are not specifically relevant to the teaching of inferential statistics to non-statisticians. In my **third analysis**, taking a SCT perspective on teaching and learning (Section 4.2.1), I investigate statistical routines, the ways in which lecturers construct statistical meaning in their statistical discourse (Section 5.4.3.3).

5.4.3.1 Lecture organisation

In this first analysis of lectures, I used implemented lesson/lecture plans I re-created using lecture resources (slides, handouts, writing on the board), lecture transcripts and my own notes during lectures (Saroyan and Snell, 1997). My unit of analysis was lecturers mediating students' sense making mediated by the organisation of lectures in terms of content, time, context and sequences of learning. The analysis of how lecturers organised the teaching resulted in three types of lectures: new topic lectures, integration lectures and revision lectures, each with associated forms of lecture organisation, routines and purposes.

In summary, new lectures were 'hierarchical' in form, which is considered to be a useful organisation for complex and difficult subject matter and to aid students' memory (Bligh, 1972). In hierarchical lectures, different points of information are grouped together with a unifying feature, such as a heading. Since each item is grouped with another idea, hierarchical lectures are considered useful for the presentation of facts and for dealing with introductory and difficult topics, aimed at students with limited prior knowledge of the subject matter.

Integration lectures used lines of argument or 'chaining' to arouse motivation, synthesise what was taught up to that point and suit the students' level of attainment since by this point the students were assumed to have had some understanding of the subject matter. In integration lectures, the lecturer would therefore make assumptions about students' prior knowledge and abilty to handle the concepts. In chaining lectures, each story consists of a chain or sequence of events. The two lecturers synthesised what was taught up to that point and made assumptions about students' prior knowledge and their ability to handle the concepts.

Revision lectures assumed some prior knowledge of statistics and used either chaining (L1) to maintain students' attention or a hierarchical representation of the module content and a focus on statistical knowledge rather than involving reasoning processes (L2).

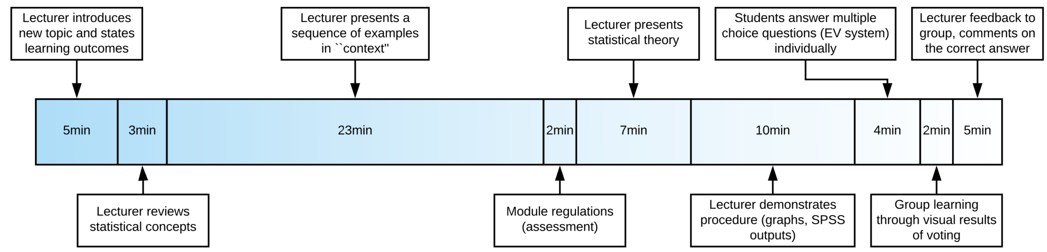
Module A types of lectures

Based on my analysis of the ways in which L1 organised the lectures, I concluded that there were three types of lectures: *new topic* (lectures 1, 2 and 4), *integration* lectures, consolidating material relating to the introductory lectures (lectures 3, 5-7) and *revision* that aimed to prepare students for the module assessments (lectures 8-11). Figure 5.7 shows the timelines of L1's lecture activities for the three types of lectures[2,3].

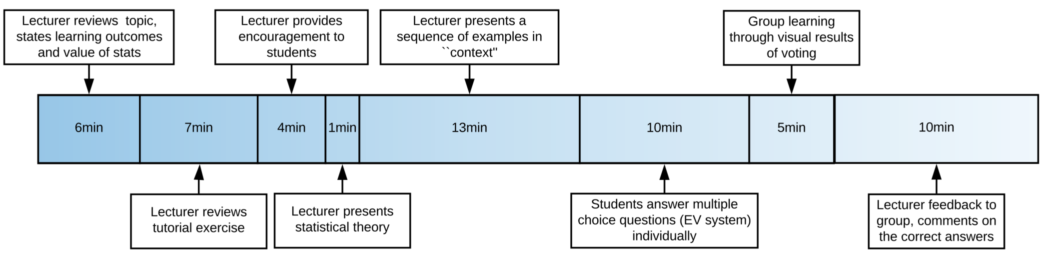
New topic lectures. L1 would start a lecture with stating the aims of the module and of the lecture, a brief comment about module regulations (e.g. attendance, summative assessments, deadlines), a brief introduction to the topics to be covered, e.g. 'today I would like to start with a very common inferential statistic test which is a t-test. [...] I will go into [t-tests] in a bit more detail, understand how it works and what we can do with it' (Lecture 2, Module A, 0:01:56.8). Bligh (1972) classified lecture organisation into two types: "hierarchic" forms and "chaining", with variations of the two in more complex forms. The 'new topic' lectures I broadly characterised as hierarchic in form since L1 presented the statistical theory first followed by examples linking to the statistical theory.

[2]Note that the timings in Figure 5.7 are approximate and that lectures did not last exactly sixty minutes since they were immediately followed by tutorials.

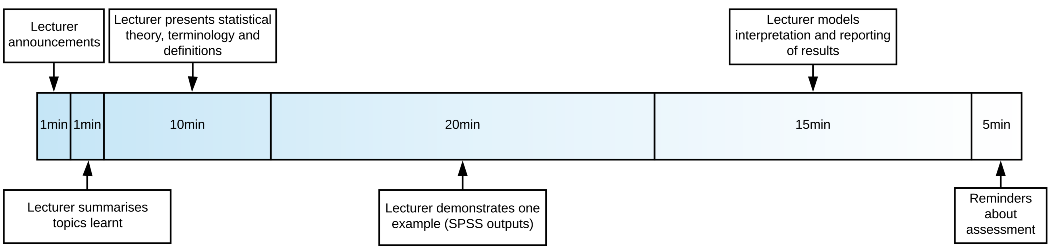
[3]The content of lectures is summarised in Section 5.3.3.1 and Appendix C.2 (page 290).



(a) Lecture 2, Independent t-tests (New topic)



(b) Lecture 3, Dependent t-tests (Integration)



(c) Lecture 8, Revision, ANOVA (Revision)

Figure 5.7: Module A: introductory, integration and revision lectures

For example, in a new topic lecture, L1 would spend five to six minutes on reviewing previous sessions and talking about the value of statistics to students' professional training and the value of student engagement in the learning opportunities provided by the module. The review included concepts the lecturer considered to form the basis for the new topic (e.g., hypothesis testing, statistical significance, principles of experimental design), a restatement of the learning outcomes ('by the end of this session you should be able to run an inferential test using SPSS', Lecture 2 0:05:23.3). The ways in which L1 linked his learning outcomes to the lecture content implied a view of statistics that was relevant to students and also necessary for something (Vygotsky, 2012, see Section 4.2.1.2, p. 112). L1 then went on to explain the new statistical model (e.g. independent samples t-tests) using sequences of examples of "context", real-world situations or representations. For instance, examples of differences between two conditions included the scores of males and females on a statistics test, driving under the influence or not, comparing two drugs.

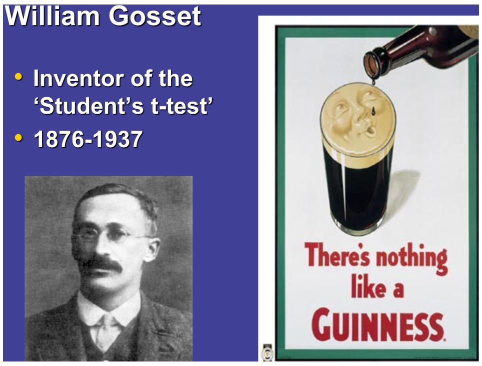
For example, in TE:1 L1 introduced t-tests by referring to the history of how Gosset invented the t-tests and showing pictures of a pint of Guinness and of W. Gosset on screen. In total, a 'new topic' lecture included four different representations ("context"). With Gosset's picture on Slide 1, L1 represented a cultural icon for the statistics' community and aimed to shape students' identity (Rodd, 2003): "[Gosset example] gives the students a sense of first how statistical techniques have come to be, as it were, what kind of inventions they are" (L1, Interview 2, 0:04:49.1).

The everyday representations of situations when statistics might be used as a tool were followed by statistical theory. L1 reluctantly used mathematical symbolism and in the case of the lecture on t-tests, L1 presented a 'basic' formula (Slide 2, TE:1) 'for [Students] to understand the general principle', and at the same time L1 emphasised that students 'will see how this overall principle works in practice' (Lecture 2, 0:30:30.4). I refer to this teaching episode again in Section 5.4.3.2 and 5.4.3.3 when I discuss L1's discourse in terms of pedagogic and statistical discourses.

Module A, Lecture 2, Title of slide: Basic formula for t

TE:1 "Now let's get into t-tests. [...] What do you think (.) is the relationship between beer and statistics? Do you think there's any relationship between beer [stout] and statistics at all? Beer, in this case Guinness, Hmm? (2) Well, there is one, and a very interesting one. What we refer to as the t-test was invented by William Gosset. (( )) work at the Guinness brewery in Dublin. [...]

Slide 1: William Gosset



This is the basic formula for t-test. What do you have? You have mathematical averages. Two groups, M_1 minus M_2, divided by the measure of spread within groups. The measure of spread or variance is an important aspect in a t-test." (0:28:27.2)

Slide 2 title: Basic formula for t

$$t = \frac{M_1 \text{ minus } M_2}{\text{measure of spread within groups}}$$

At different points during the lecture, L1 also spent approximately five to ten minutes encouraging the students to persevere with the practical exercises. For example, in TE:2, L1 gives students advice on how to approach their studies on the module by carrying out the exercises provided in class.

Module A, Lecture 5

TE:2 *"It's so easy statistics, it's always here, never goes away, never moves. It's always there. [...] Statistics is easy, do the same practical, allow a bit of time and then come back to the same exercise and see if you can do it again. Sometimes you may get a false impression. Consolidate. When you are in a position to do it in a different context, it is not so easy." (0:11:48.2)*

At the end of a lecture, L1 would present about ten multiple choice questions to students on slides using an electronic voting system (Kay and LeSage, 2009). Students answered the questions individually and submitted their answers via the clickpad. L1 than presented the spread of answers in the class for each question as well as the correct answer. Finally, L1 gave feedback to the whole group with comments about the correct answers, followed by further encouragement.

Integration lectures. While in the introductory lectures L1 provided theoretical statistical and mathematical knowledge, approximately spending seven to ten minutes in the introductory lectures (Figure 5.7a), in the *integration* lectures L1 emphasised *practical* applications of statistical models at the expense of *theoretical* content (Figure 5.7b).

For example, L1 would often state that 'what is important is the practical side' (Lecture 4, 0:00:44.7). It appeared that the main aim of these lectures was 'to put [statistics] as much as I can into a broader *context*' (Lecture 3, 0:46:55.3). As such, L1 would present three examples within a "context" and involve students in answering multiple choice questions using the electronic voting system at the end of the lecture.

For example, the aims in TE:3 stated the problem students needed to deal with (e.g. two-way analysis of variance, ANOVA) which were followed by a definition of ANOVA (slide 2) and two examples of studies using ANOVA (slides 3 and 4). L1 preferred to narrate different statistical situations or "contexts" (as in slides 2 and 3) before commenting on the research design (variables), the results of the analysis and the interpretation of results. The integration lectures were organised in a sequence of stories and events so chaining of information in my view best described L1's approach to organising such lectures (Bligh, 1972). Chaining therefore was a natural choice of exposition for these types of lectures.

Module A, Lecture 6, Slides

TE:3

Slide 1: Aims

- To consolidate key issues related to two-way analysis of variance
- To discuss in more detail the notions of main effects and interaction

Slide 2: Two-way ANOVA

- Involves two IVs [independent variables] and a single DV [dependent variable]
- It has the potential to indicate the extent to which the two IVs may combine to influence scores on the DV

Slide 3: Context-dependent memory study (Example 1)

- Godden & Baddeley (1975)
- Context dependent memory = memory for information is best when the surroundings at recall match the surroundings at encoding.
- 2 IVs = place of encoding (on land vs under water); place of recall (on land vs under water)
- Findings: word lists that were first encoded under water were best recalled under water; those encoded on land were best recalled on

land

Slide 4: (Example 2)

- We might want to see if... Students revising for exams should try to match the conditions of the exam as much as possible when doing the revision
- Questions:
  1. Will there be an overall difference in exam performance between those participants who revised in the sitting room and those who revised in the exam hall?
  2. Will there be an overall difference between those who took the exam in the sitting room compared with those who took it in the exam hall?
  3. Will exam performance be best when revision conditions match exam conditions?

Because L1's lecturing involved a chain of argument, L1's teaching was accompanied by detailed oral and written explanations (on slides). An example such as slide 3 in TE:3 would be explained only briefly (approximately 5 minutes), as an example of a real study. L1 would then spend 15 minutes on the second example and provide much greater level of detail and linking the "contexts" of the two studies. The slides supported L1 in making clear to students when moving from one stage of the analysis to the next and for summarising the 'story' L1 was telling at that time.

Revision lectures. The *revision* lectures primarily aimed to review how to interpret statistical outputs produced by software (SPSS) in one example and to revise abstract, theoretical statistical underpinnings as preparation for the end-of-module test (Figure 5.7c). Within the framework formulated by Bligh (1972), revision lectures are a combination of hierarchic and chaining forms of lectures. L1 would make a few brief points referring to theory he wanted to emphasise in the revision lecture from a number of topics covered in the module. For each topic, L1 would then present an example and give practical advice to students on how to perform the analysis (TE:4).

Module A, Lecture 7, Slides

TE:4

Slide 1: SPSS Exercise

- 1 way ANOVAs: This is a one-way ANOVA with 3 levels, W, X, and Y. Scores are measured on Z
- 2 way ANOVAs: This is a two way ANOVA with a 3x4 between subjects design, Factor A relates to (whatever) and has 3 levels, X, Y and Z, Factor B relates to (the other thing) and has 4 levels, A,B,C and D
- Report 3 F Values!

Slide 2: SPSS Exercise continued

- Is there a significant difference?
- Answer the question, + report appropriately e.g. giving F/t, df, p values
- Visual displays
- Graphs with appropriately numbered label, e.g. Fig. 1, 2, 3 etc throughout your document.

Slide 3: Tables

- For each study, paste in your output tables and give each one a label – e.g. 'Table 1, 2, 3 etc, Table showing. . . '
- . . . THEN do a description! Make sure all descriptions illustrate your understanding of the findings as clearly as possible
- Golden Rule: More is more!

Revision lectures made assumptions about students' prior knowledge and assumed students had some ability in handling the material. However, the presentation was for students to use independently, in tutorials or after the lecture.

Module B types of lectures

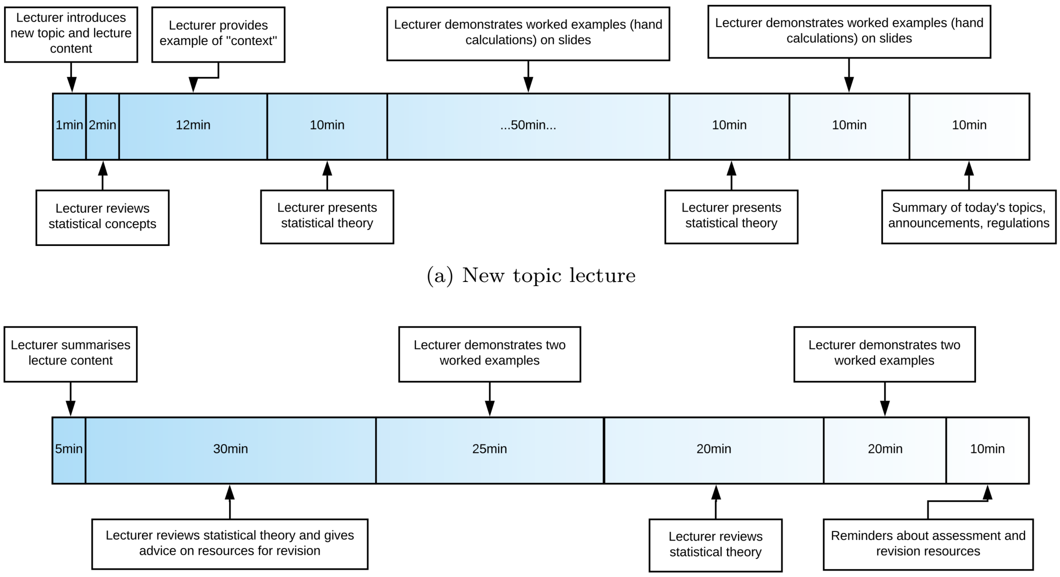
L2 also used lectures and slides as the main teaching tools. My analysis of the Module B lectures revealed that there were two types of lectures, new topic lectures and revision lectures (Figure 5.8).

New topic lectures. Lectures 1-5 and 7-10 started with L2 providing a brief review of the previous week's topics (e.g. contingency tables, conducting χ^2 tests), stating the new content and relevant textbooks, in Figure 5.8a. This brief introduction to the lecture's content (rather than learning outcomes) was then followed by definitions of theoretical statistical concepts (e.g. ANOVA). Using Bligh's (1972) framework, L2's lectures used a "classification hierarchy" which consisted of a title of the lecture, a topic such as Analysis of Variance and different points of information, such as the slide in TE:5

Module B, Lecture 9, Slide title: Today

TE:5

1. Coursework
2. Analysis of Variance (Workbook 44).
   - What is it for?
   - When is it appropriate?
   - How do you do it?
   - When can't you do it?
3. One-way ANOVA
4. Two-way ANOVA



(a) New topic lecture

(b) Revision lecture

Figure 5.8: Module B: new topic and revision lectures

Once the mathematical bases of the statistical tests were presented on slides and explained to students, L2 proceeded with presenting around four **worked examples**, e.g. the relationship between whether preference for Marmite depends on gender or road sign legibility and age, as in TE:6. In lectures 3, 4 and 10, L2 included newspaper articles and published studies to teach particular concepts (TE:7).

Module B, Lecture 7, Driving Age Example

TE:6 A Pennsylvania research firm conducted a study in which 30 drivers (of ages 18 to 82 years old) were sampled and for each one the maximum distance at which he/she could read a newly designed sign was determined. The goal of this study was to explore the relationship between driver's age and the maximum distance at which signs were legible, and then use the study's findings to improve safety for older drivers.

In worked examples, L2 showed students calculations on slides accompanied by explanations on the whiteboard using graphical displays. L2 focused on structuring the material and providing "step-by-step instructions" on how to carry out an analysis as "it might be argued it makes it easier [for the students]. You might call it spoon feeding, I call it making things clear" (L2, Interview 2, 0:15:42.8). To achieve this, L2 offered small steps at a time combining theory, calculations and interpretation of data so that the students could follow the lecture and "think about what they were learning and keep up [with the lecturer]". Thus, the solution to the problem appeared over several slides, each covering answers to one of the steps in the analysis, moving the students progressively towards answering the questions.

Module B, Lecture 4

TE:7 "So, very often in statistics we want to go and use some data to answer questions like those presented on this ER taken this example, taken some time ago now from BBC News website 'Are women drivers better than men?'. the whole idea is that this article on the BBC website claims that ((the old adage)) that women drivers are better than men, they are a bit more cautious, is the common explanation provided for that. It has apparently been proved by the Advertising Standards Authority. The results of that nature, the claims of that nature, claims of that nature can be properly backed up. And they can have an awful lot of influence. This somewhat more recent BBC news article (2) shows us the kind of influence that those kind of findings can have, apparently regardless of whether, ER, men or women are better or worth drivers ((some amongst us)) have decided that this is sex discrimination, so maybe it's unfair on the basis that there are underlying differences between men and women. Maybe it's unfair on the basis that we ought not to be allowed to distinguish between men and women. There are obviously other, this is this is a more popular example, ((but)) looking for differences between groups, looking of evidence between groups is one of the main applications, ideas behind hypothesis testing. We'll see examples through this lecture how we might go about collecting and analysing data that addresses questions of this nature. "

The practical examples were then followed by further detailed statistical theoretical bases of the statistical tests covered, including variables, hypotheses, test statistic. L2 relied on the theory-example-theory (statistical rule or approach) strategy for explaining links within a lecture in which statistical theory was followed by an example, followed by further theory and examples, as in Figure 5.8a. L2 conveyed the mathematical bases of statistics with little reference to statistical software in lectures. L2 also aimed to demonstrate to students the link between variables, the research questions that were answered using statistical experiments, applying formulas and interpretation of results.

Revision lectures. Lectures 6 and 11 provided an opportunity for the lecturer to emphasise the module content and topics students needed to be aware of most. In Lecture 6, statistical theory was followed by worked examples printed on handouts with gaps for students to fill in the answers while the lecturer provided the calculations. In Lecture 11, there was an emphasis on the summative exam structure, on providing general advice on how to answer the questions in an exam situation and on the resources students should use to prepare for the exam, as in Figure 5.8b. Unlike Module A revision lectures, L2 showed students a hierarchy of module content, move on sequentially through the theory relating to each topic and showing students 'what they might be asked in the exam' as in TE:8.

Module B, Lecture 11, Slide title: What might you be asked in the exam

TE:8

1. T-tests
   - Calculate and interpret one sample t-test.
   - Calculate and interpret independent sample t-test (when equal variances are and aren't assumed; use F-test to check this assumption.)
   - Calculate and interpret paired t-test.
2. χ^2-test
   - Analyse contingency tables with χ^2-test of (i) independence or (ii) goodness of fit.
3. Regression and correlation
   - Calculate, interpret and use regression equation.
   - Calculate and interpret Pearson's correlation.
   - Conduct appropriate hypothesis tests.

However, the revision lectures were more complex than the new topic lectures since L2 conveyed a more complex argument through comparing and contrasting the different statistical models and principles. For example, L2's general advice at the end of the lecture was to 'make sure you understand not just how to do the various tests, but also when they are appropriate and how to interpret the results' (0:04:33.6, Module B, lecture 11).

In summary, in Modules A and B, new topic lectures were characterised as hierarchical in form since the object of the lecture was to introduce new concepts and then show how they are applied in practice. Integration lectures on Module A were characterised by chaining, with the lecturer presenting a sequence of examples in different "contexts". The integration lectures were aimed at keeping students engaged with the module, present connections across the curriculum or different lectures and help students 'consolidate'. Finally, revision lectures were a combination of chaining and hierarchical forms in Module A and hierarchical forms for Module B but with the addition of complex comparisons across the module content.

5.4.3.2 Pedagogical routines

The literature review (Section 2.5.3) revealed very limited research carried out on how lecturers provide explanations to, motivate or question students in statistics education. Vygotsky's (2012) account of how children 'explain' words (the development of word meaning) involves what object the word designates, its attributes or what can be done with it. In this view of thought versus meaning, the inner planes of verbal thought, abstract concepts can be translated into the language of concrete action. The description of lectures from the two modules shows that the lecturers planned and organised their teaching using a particular combination of teaching methods and decisions *during* lectures, some of which I summarise in Table 5.2. My second analysis of the lecture observations data investigated the pedagogical routines to do with how lecturers mediated students' sense making mediated by explanations, motivators and questioning (Bligh, 1972; Viirman, 2015).

Table 5.2: Pedagogical routines

Strategy	**L1**	**L2**
Explanations		
with narratives and objects on slides	+	+
with narratives and whiteboard	-	+
with narratives of "context" from another domain	+	+
with narratives about history of statistics	+	-
with metaphors	+	+
with theory-example-theory	+	+
with equations	+	+
with definitions	+	+
with graphical representations	+	+
with statistical methodology	+	+
how to design study	-	-
how to collect data	-	-
how to analyse data using hand calculations	+	-
how to analyse data using statistical software	-	-
how to interpret statistical analysis	+	+
Motivators		
by showing the utility of statistics to students	+	+
with "contexts" relevant to students	+	+
by providing handouts	-	+
by making the learning outcomes clear throughout	+	-
by making the content of the lecture clear	+	+
by showing the nature of statistics	+	+
by commenting on study skills	+	-
by referring to summative assessment	+	+
by using humour	+	-
Questions		
control questions	+	+
facts	+	+
formative assessment	+	-
rhetorical questions	+	+

Note: + = *yes*, - = *no*

Explanations

The issue of how lecturers explain statistical ideas to students is particularly complex since statistical ideas involve knowledge and understanding from multitude of domains, including mathematics and the students' programme of study (Section 2.6). In experiments of micro-teaching, Gage (1978) for example captures the importance of teacher explanations to student learning (Section 2.5.3.1). Bligh (1972) summarised the challenge with providing explanations in lecturing:

> "*Effective explanations link what is to be explained to two or more ideas already in the students' minds. If these two ideas are not already in the students' minds, they too will need to be explained. If so, at least one of them will need an explanation of a different type.*" (p. 119)

Since numbers in statistics are numbers with a context (Cobb and Moore, 1997), the lecturers' discourse built up explanations and links through the curriculum presented in lectures using slides showing statistical theory and examples of applications of those models in different "contexts". Lecturers explained not only the statistics but also the 'external' contexts, i.e. from domains other than statistics. I come back to this point in more detail in Section 5.4.3.3. Bligh (1972) recorded eight types of explanations used in lectures. Relevant in my study were *mental*, *regulative*, *analytical*, *functional* and *spacial* (Section 2.5.3.1). In this section, I provide examples of explanations used by the lecturers during the course of the modules.

Lecturers' narratives about the history of statistics is an example of a *mental* explanation, which is a type of metaphor (Martin, 2003). A strategy used by L1 was to explain to students why and how some statistical concepts were invented and how they have evolved, which are considered to be important to conceptual understanding (Derry et al., 1995; Bakker et al., 2008). For example, L1 introduced a new topic, Student's t-test, using the example of how Gosset invented t-tests nearly a century ago. In TE:1, p. 178, L1 provided details about the statistical model and links relevant statistical concepts (e.g. mean, sample, population), explained the problem Gosset solved using statistics while he was employed at the Guinness factory in Dublin thus giving statistics a purpose outside the module and finally presented a formula for t-tests. L1 did not go further with this example but rather used it as a general motivator (discussed next) for introducing the theory of hypothesis testing.

Another strategy used by lecturers was to use *regulative* explanations by presenting basic statistical concepts first, before engaging students in complex examples for which in-depth understanding is required such as in a theory-example-theory sequence. In lectures, both lecturers sought to build curriculum from a few, basic yet central abstract concepts. This approach to teaching statistics is reminiscent of the sociocultural top-down approach to teaching developed within the context of school mathematics (Newman et al., 1995; Renshaw, 1996). The top-down approach emphasises the knowledge of the teacher and an ascent from abstract, theoretical knowledge to concrete (Davydov, 2010). Signs and symbols are treated as pre-given tools that can be brought into the classroom/ lecture theatre[4].

Within the university setting, the lecturers 'moved' from statistical concepts to statistical activities (e.g. explaining a "context" or showing a worked example) to understanding of statistical models (t-test, analysis of variance, regression). In this way, the general and abstract concepts appeared to connect to concrete and empirical experiences under the lecturer's guidance. However, following from the Pilot study findings (p. 92), university statistics curricula presents a contradiction between the theoretical knowledge students *should* learn versus the knowledge they *can* learn as well as the theoretical knowledge relevant to students' longer term goals. As discussed in Section 5.3.4, curricula that is relevant or necessary for something can differ depending on lecturer, students or domain of study.

Worked examples (with hand calculations or statistical outputs created with software) were instances of *analytical* explanations. For L2, a teaching strategy was to show hand calculations and use examples that enabled students to apply the statistical models. The problems were manipulated to show straightforward relationships between elements, as in TE:6 for example. In this type of analytic explanation, L1 showed a general rule (statistical model equations), an instance/example of the variables in the model (intermediate steps in calculations) and finally how the general statistical model is applied. Similar to L1, this strategy required the presentation of statistical theory first, which could then be used in calculations. The step-by-step approach to presenting the worked examples suggested L2's wish to provide 'correctly organised' teaching in order to enhance students' learning of statistics.

L1 on the other hand chose to show details of the problem (e.g. Slides 3 and 4, TE:3) followed by graphical representations and outputs of analysis produced using statistical software (SPSS). L1's emphasis was on explaining the research questions that are answered using statistics and the possible interpretations of statistical analyses, i.e. the first and the last steps of the statistical investigative cycle (Wild and Pfannkuch, 1999). L1 was keen to establish empirical generalisations which the students then could apply in further situations (e.g. Quote 11 and TE:2). L1's implemented curricula appeared closer to what a psychologist does. L1's focus was on the "context" of the problems presented on slides, captured in a sequence or cycle of cases through the module (Malone et al., 2012).

Functional explanations sought to emphasise relationships between the statistical process components of the statistical investigative cycle. The lecturers' step-by-step, 'scaffolded' presentation of the worked examples or statistical outputs were examples of functional explanations, intended to make sense of the purpose of the statistical activity for the students and served the purpose of conveying the thinking behind the steps of a statistical analysis (actions) in the example. In either hand computations or interpretations of statistical outputs, the lectures' strategies were to ensure students possessed the tools to make the connections between "context" and statistical model.

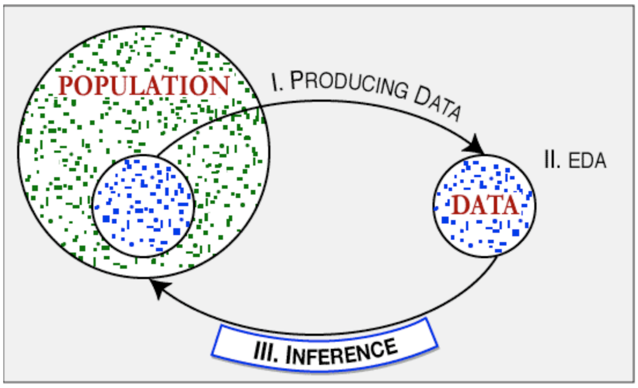
To introduce new concepts, the lecturers often made use of visual aids as *spacial* expla-

[4]By contrast, within an approach generally promoted in bottom-up teaching, the taken-as-shared ways of using symbols grow out of children's activity under the teacher's guidance (Cobb et al., 1996).

nations. For example, L1 used pictures of thin and overweight people to explain differences in means. L2 used the graphical representation to explain inferential statistics. In this case, the spatial explanation uses an abstract visual aid of intersecting circles and arrows. At first glance, it is hardly comprehensible in the absence of a deep understanding of inferential statistics, in TE:9. L2's explanation however, rooted in a "context" assumed to be familiar to students' future professions, makes reference to everyday, concrete objects, cement, most adults would be familiar with. In explaining ANOVA in Slides 3 and 4 in TE:3, L1 explains the psychological concepts involved in the two studies first and then moves on to explaining the statistical model. Such examples suggest the challenge in combining statistical and non-statistical concepts in explanations.

Module B, Lecture 11

TE:9 (0:52:18.4) In a manufacturing type situation, we might want to use information about the constituents of a particular kind of a cement maybe in order to predict the strength of that cement. In order to actually measure the strength of the cement, we have to do some sort of test where we have to actually destroy a bit of cement. So obviously we want to be able to make accurate predictions based on the input without actually having to test and destroy actual cement we want to use for some building project.



However, such representations of abstract statistical concepts, such as statistical inference in TE:9, can result in contradictions when their meaning is taken for granted. In the context of school mathematics, Seeger (1998) considered that representations and manipulatives can hinder students' sense-making. For the students in my study, the representation of statistical inferencing might require an explanation of the statistical concept and an additional explanation of the visual representation. In this TE, the explanation involves an additional external context of a manufacturing situation. In *Activity, consciousness and personality*, Leontiev (1978) noted that a scientific explanation of the relationship between an image and its features can be subjective:

"in order for a sensible, visual, or aural image of an object to appear in a man's

> head, it is necessary that an active relationship be established between the man and this object" (p. 49)

In my analysis, I use a common reformulation of Vygotsky's model of mediated act (Figure 4.2) to fit my data analysis in which the unit of analysis is individually focused (rather than to understand dialogue, multiple perspectives and networks of interacting activity). In a Vygotskian interpretation, the individual needs to be understood with his/her cultural tools and the social context needs to be interpreted with the agency of individuals who use and produce tools. This means that the objects of the activity, students' sense making, become "cultural entities" and the object-orientedness of the teaching action in TE:9 becomes key to understanding human cognition (Engeström, 2001). So if an image is used as a teaching tool, its function is to build up the internal actions of the student. Its function depends on the culture in the lecture theatre that teaches multiple perspectives, both abstract and concrete.

Using the "symbolic" representation of the visual image in TE:9 and L2's speech about a "concrete" manufacturing situation, L2 intended to teach students a statistical concept rather than enrich the sensory experience in the lecture theatre. The representation in this case reflects L2's explanations of statistical inference using a metaphoric representation using visual aids in the lecture theatre and not statistical inference. As in the case of Vygotsky's (1978b) example of tying a knot in a handkerchief to remember something, L2 creates a mechanism of "reverse action", a tool that is operating on the individual, not on the environment.

Motivators

Reminiscent of Vygotsky's (2012) view of learning which is required to be relevant to students' lives and necessary for something (Section 4.2.1.2), Lave (1988) has contrasted the success of learning situated in 'everyday' contexts as purposeful and motivating with the inert knowledge of formal schooling. Lave was referring to the acquisition of workplace-based skills, not necessarily applicable to the learning of statistics. Instead, the two lecturers in my study often emphasised to students the value of statistics to their lives and professional training. L1 aimed to keep students motivated during lectures and planned for the examples to have 'personal meaning' for the students. Davydov and Markova (1982) propose an approach to learning in the context of educational activity which assumed that a person does not necessarily become 'submerged' in activity. For the development of concepts, Davydov and Markova (1982) suggest a direct link between the experience of object-directed activity and the mental development of the student (child). Vygotsky (1997) and later Davydov and Markova (1982) were concerned with the apparent contradiction between social practices and personal meaning.

> "*In the course of development of educational activity, it is necessary to ascertain and create conditions that will enable activity to acquire personal meaning, to*

> *become a source of the person's self-development and comprehensive development of his personality, and a condition for his entry into social practice.*" (p. 55)

Davydov (2010) describes a teaching intervention in which students began with counting of concrete objects followed by abstract, algebraic notations. The assumption is that when mastering a concrete operation, the child will also master some general structural principle which can then be applied broader than the particular operation. Similarly, both L1 and L2 extended the content beyond statistical theory and included everyday concepts which they believed were familiar to students. Since their main goal was to provide students with the key aspects of statistical models and the situations in which such techniques might be useful in Psychology or Engineering, the presence of real Psychological studies, newspaper articles or engineering scenarios seemed to be a powerful pedagogical tool for bringing statistics into the micro-context of the lecture.

Thus, a teaching strategy used by L1 and L2 to accomplish their learning objectives was to integrate everyday concepts and statistical, scientific concepts. First, the everyday concepts used in examples to describe statistical models were elaborated and made more precise and general by lecturers. Second, students' understanding of Psychology, Engineering or other 'everyday' situations, assumed to be personal to the students, was used by the lecturers to represent statistical models, inteded to be relevant for something (see Section 4.2.1.2, p. 112).

L1 would also address the students directly and keeping them engaged with words such as 'let's recall', 'remember', 'coming back to our example'. L1 would also appeal to the students' empathy by making his discourse personal. In TE:10 L1 made past references to himself as student. In this way, L1 encouraged to students to 'imagine' themselves as psychologists and showed the utility of the statistical model to students' lives outside the module.

Module A, Lecture 4, Introduction to ANOVA

TE:10	"I remember when I was a student, [I was thinking] "why don't we do multiple t-tests?". We know t-test is a very good instrument, a very good tool. Why don't we do t-tests? Is there some rule against it? Is it because you just say so? But it's a good statistical reason. Imagine we wanted to design a course and we wanted to see which teaching method would work best for our students. You're an educational psychologist for instance and you wanted to design a course."

Staats (2007) contends that the "context" of a problem or application, even the thin "context" of a textbook word problem 'invites students to participate in an imaginative act or a fantasy world' (p. 7). In my case, L1 compared the utility of ANOVA, the topic of the lesson, with a model taught in a previous lesson, t-test. L1 also referred to students' futures, thus trying to make the statistical concepts memorable. Statistics on this module was not just numbers and datasets since L1 tried to connect to people and life in different macro-contexts. Staats (2007) uses the concept of 'imagined world' to refer to the context of mathematics. L1 actively promoted this type of thinking about 'imagined worlds'.

"Contexts" were also used a motivator and stimulating tool on Module B, as for instance in TE:9. In these TE, lecturers demonstrated how they can use 'dynamic contexts' in socially contextualised learning experiences for the students to develop a sense of the importance of statistics to their professional and personal lives yet still focus on statistical learning objectives. In mathematics education, Boaler (1993) maintains that approaches to context that develop both personal meaning and a deeper understanding of mathematics improve transfer of learning to new situations. A challenge for lectures is to find 'external' contexts which are real to the students, not just the lecturers. In Memo#3, Module A, I noted that L1 assumed that his students had some understanding of psychological studies as well as familiarity with a range of statistical concepts in order to be able to engage with this example. For instance, in Lecture 2, L1 encouraged his students to connect the module's activities to the psychology literature they were familiar with and use statistics as psychologists do.

On Module B, there were very few instances of interaction with students and there was no component to evaluate learning during lectures. Instead, students were encouraged to go back to the lecture slides and textbooks after the lecture. Similarly to Module A, L2 also aimed to influence students in accordance with his own educational goals (as discussed in TE:7 and Section 5.3.3) and considering students' motivations, interests and willingness to take part in the lecturing activities. It seemed that the focus on Module B was on formal learning in communication and cooperation with the lecturer and after having done this joint activity, the students were encouraged to carry out similar activities independently and approach these exercises in 'statistical' ways (Section 4.2.1.4).

Questions

Vygotskian developmental teaching (Section 4.2.1.4) emphasises teaching as guiding or as assisted performance. The teacher, seen as the more capable adult and organiser of learning, provides or negotiates 'hints', 'supports' or 'scaffolds' to or with the student. Since the focus is on change within the ZPD, the teacher is expected to lead the student from informal to mastery and beyond through mediational means such as modelling, leading questions or by providing initial elements of the task solution. Vygotsky's distinction between everyday, pseudo and scientific concepts means that a scientific concept is achieved when the everyday, followed by pseudo concepts have merged with the scientific (Lave and Wenger, 1991). Questions asked by teachers have an important role to play in the process of transformation from everyday concepts to scientific concepts (Davydov, 2010).

My analytic approach was based on the grounded methods used in the initial phase of data analysis (open coding and categorising, Section 4.7.2). The categories or concepts were further refined based on previous research by Bligh (1972) and Viirman (2015). Other models were considered at this stage such as Bloom's (1956) taxonomy or the categorisations proposed by Pedrosa de Jesus and da Silva Lopes (2011) (Section 2.5.3.2). However, these were not suitable in the context of the lectures observed as these frameworks focused on cognitive levels not revealed in the data.

In this analysis, I noticed four types of questions (Table 5.2) control questions, facts, formative to check student learning and give students feedback about their own learning and rhetorical. Both lecturers used *control or comprehension questions*, such as 'any questions up to now?', 'what does it tell you/us?', 'right', 'yes?', 'remember?' (Viirman, 2015). In the situation of the lecture theatre, students did not generally give a verbal response. However, in my notes, I observed that students were nodding and making eye contact with the lecturers. Thus students were able to give non-verbal cues to the lecturer. These questions were generally asked at key points in the lecture at the transition lines in Figures 5.7 and 5.8. The lecturers would then move on to another part of the lecture.

The lecturers also asked *factual questions*, such as 'how do I do degrees of freedom, where do I look?', 'are you familiar with SAT scores?', 'but which group differs from which?', 'what does quasi mean?', 'what are the two main types of hypothesis?', 'is it positive or negative?', 'does the other variable tend to increase or decrease?'. On Module A, only on one occasion a student answered such a question, 'what's the level in psychology we use for statistical significance?'. However, rather then moving on, lecturers would offer additional explanations and factual information about the content. For example, in Lecture 4, L1 asked 'how many conditions do we work with in ANOVA? Hmm? (3) Two, two conditions, two groups." (0:05:56.9).

Again, lecturers were able to receive feedback from students with factual questions through eye contact, level of noise in the lecture theatre, pauses which lead to informal brief comments (often inaudible). Although the questions involved simple facts relating to the content of the lecture, often the factual questions were about key statistical concepts. Only once lecturers were satisfied that students were familiar with such concepts they moved on to the more difficult part of the lecture. Emphasising key concepts was also important for lecturers in scaffolding the learning.

In Module A, feedback from students, either in informal social exchanges or using electronic voting, meant that students could be active participants in the lecture, as managed by the lecturers. The lecturer and students checked the learning that took place together, mediated by the multiple choice questions (related to the end-of-module assessment) and the voting system. Histograms of student answers to the four multiple choice questions were means of making students' understanding visible to the group. L1's feedback directed at the students (objects) in relations to the questions asked (tools) was intended to convey the thinking behind the assessment tasks. Instead of using electronic voting, L2 asked students twice to show hands for one of two choices of statistical tests (e.g. paired versus independent t-test).

A fourth type of questions were rhetorical questions which were very quickly followed by an answer from lecturer such as 'why is it the case?', 'what does the null hypothesis do?', 'what statistical techniques come to mind?', 'how might we go about describing that relationship?', 'here is the data, what can you tell from it?', 'is there any evidence to support this concern?', 'what kind of procedure should we be using?'. These rhetorical questions seemed to require a higher level of cognitive involvement from the students and marked fundamental issues for the students to take away from the lecture. Rhetorical questions were also opportunities for lecturers to open an important sequence of learning.

Through this analysis, I concluded that at the micro level, the lecturers provided context in the lecture by incorporating macro-level factors such as students' previous knowledge and experiences of studying other subjects at university, relevant to their lives or of a well-known situation (e.g. drug controlled trials, machine break-down), as well as through visual or gestural support (e.g. pictures of thin and overweight subjects, spontaneous examples from human resources, students' heights). In my analysis in Section 5.4.3.3, I therefore aim to characterise the nature of lecturers' discourse by looking at the ways in which lecturers explained the content of the examples, problems and exercises with the aim to characterise the lecturers' statistical routines or discourse.

5.4.3.3 Statistical routines

In this analysis, I assumed that the statistical discourse 'coordinates participants and determines what thinking, speaking, and writing counts as meaningful, correct, good thinking, and acceptable understanding' (Blanton et al., 2016, p. 440). In the process of analysing the transcripts of observational data and teaching resources (Chapter 4), an initial stage in the data analysis was to identify a way of splitting the data into teaching episodes (TEs) based on naturally occurring breaks in the teaching such as a new topic, example or exercise. In this analysis, the unit of analysis was the "context" used by lecturers in their lecturing discourse. When comparing different teaching episodes, I started to notice differences in the ways the lecturers focussed on statistical concepts as well as on particular 'external' situations (e.g. TE:1, TE:3 and TE:6). In this section therefore, I develop the concept of 'context', the stories, situations, scenarios and narratives used in examples by the two lecturers.

In my initial coding of the data, I noted that the TEs showed an intertwine between statistical and non-statistical 'context' or representations. Borrowing from Vygotky's geographic analogy (Section 4.2.1.5), I represented the lecturers' discourse using a system of two axes which would allow for depicting the relationships between concepts with a "longitude" and a "latitude". Further, Bernstein (1999) proposes a model of vertical scientific and horizontal everyday discourse (p. 123). Statistical problems deal with information which can involve non-statistical objects, equally difficult to understand or deal with in the analysis as everyday, spontaneous knowledge (in Section 5.3.3.2). Bernstein's theory of vertical and horizontal knowledge/discourse is a way of explaining the interdependence between the development of scientific and everyday concepts which transcend educational stage or context. In my analysis, I develop Bernsteinian ideas by showing that statistical discourse prescribed for engineering or psychology students incorporates numerous examples that model everyday situations, intended to include the students into the discourse of university statistics. The recontextualization of everyday real-life situations into curriculum is a two-dimensional representation between 'real statistics' and 'real-life'.

Decisions regarding the non-statistical context during a statistical analysis can lead to various interpretations of the statistical outcomes (output). In this dataset, I therefore decided to represent the two types of knowledge in a TE as two-dimensional axes: statistical S representations or "contexts" on the vertical axis and non-statistical NS representations on the horizontal axis, in Figure 5.9. The axes range from 'rich' or more (+) of the statistical/non-statistical context to 'thin' or less (-) "context". Vygotsky's framework to concepts (Section 4.2.1.5) includes everyday, pseudo and scientific concepts. A pseudo concept is somewhat in between an everyday, concrete and a scientific, abstract concept. Therefore, in Figure 5.9, a pseudo concept used in teaching might involve a hybrid of 'some' statistical and 'some' non-statistical "context". Abstract, statistical represent a degree of generality presented by a given mode of representation. By contrast, non-statistical representations have a concrete, everyday and objective reference of a concept and relate to the reality to which they apply.

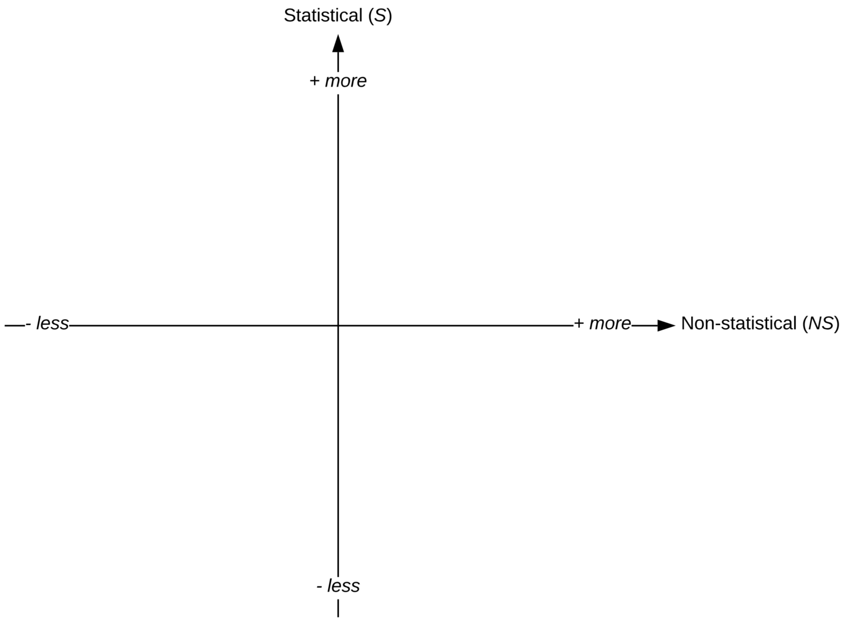


Figure 5.9: Lecturing discourse: statistical versus non-statistical

The diagram in Figure 5.9, through its depiction of the two axes, statistical and non-statistical, is also reminiscent of Cobb's definition of statistics as the "interplay between pattern and context" (Cobb and Moore, 1997, p. 802). For Cobb and Moore, statistical knowledge and context knowledge are quintessential components of a statistical investigation, and similar to Ben-Zvi and Aridor-Berger (2016) use of context and data 'worlds'. The context world is *subjective*, introspective analysis since it might include for instance cultural knowledge about the data, the norms by which it was produced, while the data world is all about *objective* real scientific data manipulation (Vygotsky, 1978b). In my analysis, while the coding revealed the two contrasting yet inter-linked *worlds* of the situational and the statistical, I sought to interpret the ways in which the lecturers were *interlacing* the statistical and the non-statistical types in their teaching (Cobb, 1999).

To characterise the "context" in the examples given in teaching, in this analysis I included 234 teaching episodes (detailed in Table 5.3). Next, I *placed* different TEs on the two axes in Figure 5.9 and categorised five broad types of discourse, defined according to levels of statistical and non-statistical context: real life, realistic, symbolic, prototypical and parabolic, represented in Figure 5.10.

1. **Authentic** : actual, real life, published and referenced studies.
2. **Simulated** : realistic "context" which may be based on real-life studies but is adapted for specific classroom purposes.
3. **Parabolic** : a narrative about an informal, non-statistical, real-world context (e.g. a historical example) that is compared to a formal representation of a statistical concept.
4. **Prototypical** : idealised "context" such as worked examples, with sample answers and calculations.
5. **Symbolic** : statistical theory with minimal reference to a contextual situation.

These TEs embodied characterised macro conditions, such as authentic statistical applications in the real-world or intended curricula of each module. They also referred to micro conditions about how the lecturers implemented these resources in the classroom.

I arrived at the differentiation of the two forms of representation (Wertsch, 1992) alongside the two dimensions by comparing pairs of episodes and theorising how TEs were similar or different from each other. The representations in my data connected the world of statistics to the world outside the lecture theatre by using scenes, situations from the students' domain of study, from everyday life. These representations resulted from an understanding of the necessity of getting into the structure of statistics (Seeger, 1998). Therefore, here I show that the placement of these TEs on either axis (or dimension), as representations of everyday tools and of statistical tools, is of course not absolute. Instead, the order of these types on each line was about how the *statistical* or the *non-statistical* "context" in the TE related each other. Also, particular examples might shift along one of the dimensions during the teaching process or when used for different purposes or in different situations.

In Figure 5.10, **authentic** and **symbolic** TEs are similarly rich in statistical conceptual information, but differ along the non-statistical axis. A **simulation** would perhaps be less embedded in a non-statistical "context", but be used by the lecturer to focus on specific statistical concepts or procedures and give students a view of how particular types of statistical analyses are to be carried out. **Parabolic** TEs are pedagogical tools used as metaphors, often at the start of new topics or lectures. Although parabolic examples were embedded in

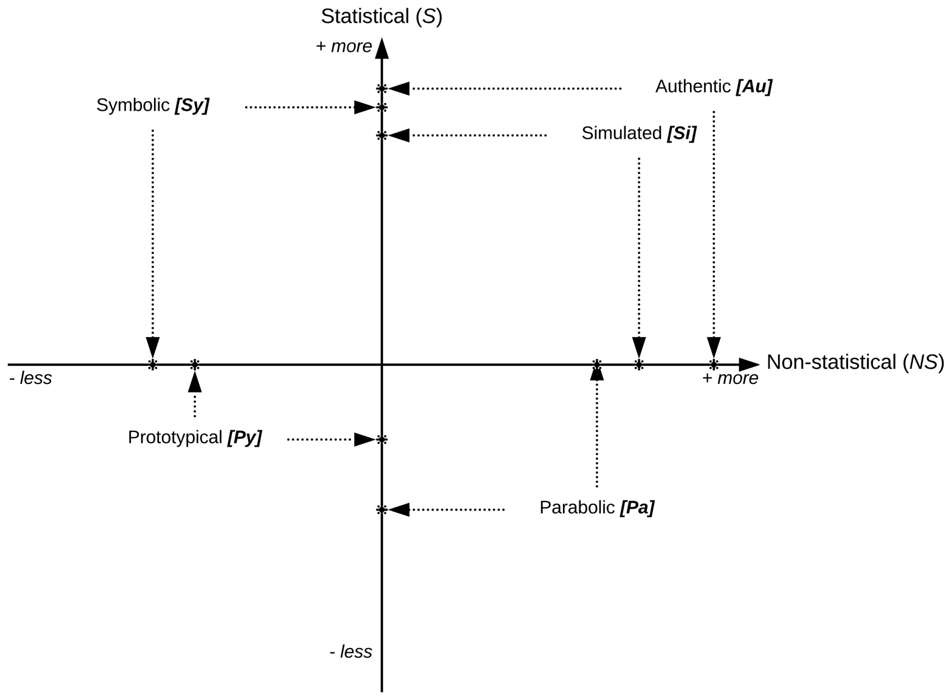


Figure 5.10: Lecturing discourse: statistical versus non-statistical forms of representation

an everyday 'context', they were used as a resource in initiating the translation of informal language into more formal, statistical language. From this point of view, I considered them closer to a statistically *thin* (less −) 'context' and also to being contextually *embedded* (more +) in a scenario. The process of categorising 'context' in these teaching episodes also lead me to consider a fifth type of episode, **prototypical**, which characterised the lecturers' use of worked examples.

In Table 5.3, in module A, I identified 80 teaching episodes in eleven lectures, of which 17 covered periods when L1 made his learning outcomes and aims explicit or made announcements. In the remaining 63 TEs, the majority of the lecturing time, L1 spent talking about authentic (30%), simulated scenarios (30%) and symbolic episodes (20%). In Module B, I identified 193 teaching episodes across eleven lectures. Of these 193 teaching episodes, five covered periods when L2 gave advice to students (e.g. about module assessment) and twenty-one when he made his aims explicit (mainly the content being covered). The remaining 167 teaching episodes were different types of "context". L2 spent the majority of the lecturing time on prototypical (35%) and symbolic (35%) TEs, with only a small number of simulated, authentic or parabolic TEs. The table also shows that the same scenario was used on several occasions to make different points during the module.

In what follows, I discuss each of these types of episodes, bringing in examples of data which I considered paradigmatic in the sense that the lecturers used them a number of times or for particular purposes.

Table 5.3: Types and number of teaching episodes

	Teaching episodes					
Module	**Authentic**	**Parabolic**	**Simulated**	**Prototypical**	**Symbolic**	*Total*
A	9	7	10	(2)	41 (39)	67
("Contexts")	(4)	(7)	(5)	(NA)	N(A)	
Time spent (approximate)	30%	10%	30%	10%	20%	
B	3	24	8	38	96	167
("Contexts")	(1)	(13)	(8)	(29)	N(A)	
Time spent (approximate)	1%	18%	6%	35%	35%	

Authentic

Authentic TEs were generally used on these two modules to (1) show students the importance of contextual knowledge in statistics, (2) statistical conceptual understanding and the role of contextual knowledge within it and (3) reveal details of authentic statistical practices and finally (4) for pedagogic purposes as motivators (Table 5.2). Authentic and simulated examples used by the lecturers were thus a higher level of statistical reasoning since they were about applying the methods to the specific case of the problem (Chervany et al., 1977; Garfield, 1998). In instances where the lecturers' discourse considers authentic interplay between data pattern and context, the non-mathematical substance of such examples, the statistical discourse is intended to develop deep statistical thinking (Cobb and Moore, 1997).

L1 for instance used Everitt's (Hand et. al. 1994[5]) study of effective therapies for anorexia six times, from lectures three to seven. The study used as independent variables a control condition and two treatment conditions, cognitive-behaviour therapy, and family therapy. The dependent variable in the study was the gain in weight over a fixed period of time. L1 adapted this context to teach both topics, t-tests and ANOVA (Example 1, Appendix D). to explain issues with research and factorial designs, validity and reliability at the end of lecture 10.

With authentic TEs, the lecturers first highlighted ***the importance of contextual knowledge in the statistical analytic process***. For example, L1 pointed out that the change in means of weight at time one and weight at time two might not necessarily be due to the family therapy intervention since other factors might affect the measurements of weights. The task relied on students' non-statistical knowledge about anorexia (even at a superficial

[5]In this Anorexia study, Everitt reported on a study of the effects of family therapy as a treatment for seventeen anorexic girls. Girls were weighed before and after several weeks of treatment using family therapy.

level) and an understanding of the experimental design of the study. The lecturers used the authentic TE (e.g. TE:3 and Example 4, p. 297) to show students how the concepts taught in the lesson were applied in real research practice. I therefore considered that the lecturers were *constructing* context, starting from explaining particular concepts, e.g. allocations into groups, then moving to show practical examples.

The second aim with these authentic TEs was to ***emphasise statistical [S] knowledge and how contextual [NS] knowledge supports this***, e.g. when manipulating data, selecting the appropriate test to use and interpreting the findings. For example, the students were required to consider the [NS] elements such as the research question, experimental design and also which statistical test was suitable to the [NS] context.

In laboratory sessions after the lecture (not attended by the lecturers), the task then required the students to use this non-statistical [NS] knowledge to understand the design of the study (types of variables) and how to enter the data into SPSS. The students used their understanding of the statistical [S] analysis learnt in the lecture (micro condition) to set up the study and run the analysis using software.

A third aim with authentic examples was to ***reveal details of authentic statistical practices***. The lecturers' experience with statistics in their research activities were relevant macro conditions in this case. For example, L2 felt he was "reasonably familiar with [this study], so can explain it better" (L2, Lecture 10, 01:38:41.1, in Example 4, p. 297). L2's main aim was to use this *research-based* example to emphasise to students the challenges with providing evidence for the validity of the study's claim, that participants were more persuaded by an argument if they knew the author's identity. L2 encouraged his students to 'anticipate these kinds of attacks' (e.g. whether the questionnaire was worded appropriately) and to 'design your study to allow for these kinds of explanations' (i.e. for the validity of inferences). Such learning aims could only be achieved using an example embedded in a contextual reality that L2 was familiar with.

Fourth, authentic scenarios were used in these modules for ***pedagogic purposes***, to provide analytical explanations through comparing and making sense of different statistical models, introduce new concepts, link statistical concepts through the module, review materials and make their learning outcomes explicit. For example, both lecturers used authentic scenarios to compare t-tests with ANOVA by manipulating the same scenario in different analyses. In this way, the lecturers also linked concepts through the module and revised content.

Further, Quote TE:11 shows how L1 used the *Anorexia* scenario to make his learning outcome clear to students and give them advice on how to approach their learning, evidencing his belief that revision and practice with examples in context aid students' comprehension and retention. In this case, L1 did the re-contextualisation for the students to meet the module's learning objectives ('define the role of inferential statistics for psychological research', Module A specification).

Module A, Lecture 5, ANOVA

TE:11 "When you are in a position to do [the analysis] in a different context, it is not so easy. What we need is transferable skills, don't we? I'm not just teaching you Everitt. I'm not just teaching you this example. I'm teaching you something that you can take into various other contexts, with your own data, with the data in your report for instance, you should be able to recognise the one way ANOVA and conduct a test and report it." (0:31:17.3)

In authentic examples, I considered that, although the lecturers embedded their discussions of experimental design or statistical models in an authentic scenario, these TEs moved from being authentic, contextually bounded to a more abstract understanding of statistical principles. In summary, with authentic TEs, the lectures showed students the relevance of statistics to their education or future employment and also connected authentic, everyday statistics to the module intended and implemented curricula.

Simulated

The use of authentic and simulated contexts characterised L1's style of teaching. Although simulated TEs included plausible or realistic scenarios, such TEs contained sufficient complexity (e.g. about research design and plausibility) necessary to achieve a particular pedagogical objective. Second, simulated TEs emphasised statistical knowledge and sense making but sometimes visualised these concepts using informal imagery that replaced formal, abstract statistical representations. When compared to symbolic, I judged that authentic and simulated TEs were more focused on the link between a situation and statistics. Third, simulated TEs showed students the steps of a procedure in a way accessible to students. For instance, the lecturers could ignore other issues with the data if they did not form the focus of the TE (e.g. ignoring measurement error when explaining regression).

A first aim with simulated was to ***get students thinking about the statistics*** presented to them. For example, in Lecture 1, L1 introduced a simplistic scenario of male and female test scores on a statistics exam to introduce new concepts in statistical inferencing (Example 2, p. 294, Appendix D). First, L1 showed two computer-generated analysis outputs of descriptive statistics using two versions of data A and B (Figure D.1a). Across the two versions, the means for the two groups were the same but the standard deviations (SD) differed in order to exemplify the effect the standard deviation SD has on t. To aid comprehension, these two situations were also plotted using bar charts (some with error bars, Figure D.1c). This *statistics exam* example was a means for L1 to develop students' statistical sense making and familiarity with statistical outputs (TE:12).

Module A, Lecture 2, t-tests

TE:12 "what is important is that you understand the number of participants, where you can derive the degrees of freedom, you have significance, you have the value and before we do anything else, you look here. I need to decide whether this is not significant" (0:43:46.7)

Further, L2 used four *real* political parties and three football clubs in the UK against a measure of happiness in order to stress experimental design principles and how to use graphical representations when carrying out an ANOVA (Example 5, Appendix D).

Another way of revealing statistical procedures to students was for the lecturers to **replace abstract statistical formulations with more concrete representations**. For instance, L1 chose not to go into exact formulations of abstract statistical models and used simulation as concrete images to 'surprise' the students and emphasise conceptual understanding. In L1's view, defining models using vivid imagery rather than formal statistical representation was an appropriate way to engage this cohort of students.

In Lecture 9, L2 presented a news article introduced previously in Lecture 4 on t-tests (TE:7, p. 183 and Example 6, p. 299) to teach the notion of interaction using informal imagery since the formal statistical process has been replaced by statistical software. L2 used the same variables *driving ability* and *gender* as in TE:7, p. 183 and added *education* (basic education, college education, higher education) to compare the *driving ability scores* of more than two groups.

Simulations were also used to ***show the steps of a procedure*** and provide students with an understanding of what goes on in the statistical analyses reported in research or media such as TE:3, p. 179. For instance, the intention in Example 2, p. 294 was not to use the 'context' of comparing test scores on a statistics exam as an application of the formal statistics, as it was the case with the authentic *Anorexia study* in the previous section, but to show the process of data interpretation. In Lecture 3, L1 expanded on the descriptive statistics from the previous week using the statistics exam study with the use of the t statistic to compare the two independent groups. In this way, the context of the statistics exam ***linked different concepts through the module and provided continuity for the students***.

Similarly, in Module B, Lecture 7, regression, L2 examines the concept of residuals, 'a number that we've come across in stats, those $\hat{e}_i$, these errors that we've come across a few slides ago, the difference between the independent variable and the predicted value of the dependent variable' (L2, Lecture 7, 1:23:30.6), by showing a series of scatterplots. L2 emphasised the value of graphical representations in both understanding statistical concepts (e.g. the definition residual $=$ observed$y$ $-$ predicted$y$ $= y - \hat{y}$) and in the process of data analysis. However, as I discuss below, Module B relied more on prototypical worked examples to emphasise detailed steps of statistical procedures.

In summary, simulations are examples of sufficient complexity for achieving particular pedagogic objectives. With simulations, lecturers aim to get students thinking conceptually by showing for example the link between contextual and statistical analyses and the mechanics of the models, in particular, the interpretation of plots and outputs. Depending on the lecturers' beliefs about the module objectives and experience with the students, simulations might also replace abstract statistical representations.

Parabolic

Parabolic TEs were specific types of examples used at the beginning of a lecture or topic. The focus in these examples was on the interaction with the students rather than the example itself in the sense that the lectures were first introducing new concepts, unfamiliar to the students, by using situations drawn from informal, authentic situations (e.g. history of t-tests TE:1, news article TE:7), but simplified to ***make unfamiliar statistical concepts familiar*** for the students. The lecturers talked about informal, authentic scenarios that they compared to a statistical model. In this way, with parabolic, lecturers were able to *explain in which situations particular statistical tests are used.* Finally, the lecturers provided a **formal representation** of statistical 'context' which was the point of the comparison.

The example of how Gosset invented t-tests nearly a century ago was an opportunity for L1 to (1) provide details about the statistical model and link relevant concepts, (2) explain the problem Gosset solved using statistics while he was employed at the Guinness factory in Dublin thus giving statistics a purpose outside the module and (3) present a formula for t-tests (as in Quote TE:1 on page 178). L1 did not go further with this example but rather used it as a ***general motivation for introducing the theory of hypothesis testing***. L1 also commented on key concepts associated with *t*-tests: sampling ('[Gosset] came up with a statistical way of comparing small samples') and statistical inference ('[Gosset] wanted to extrapolate to the whole population') using an authentic situation (barley, Guiness, quality control). L1 provided further *interesting* background to the *t*-tests and appeared to use *beer* (stout) as an aide-mémoire for the students: "when you'll have a pint of Guinness you will think 'Oh, *t*-tests" (0:26:18.7). Further, L1 emphasised the relationship between an authentic object (*beer*) and statistics. To me, this was an example of how L1 organised the learning from 'drawing letters' to 'writing meaningful text' (Vygotsky, 2012) and the movement from the metaphor using the history of t-tests, what L1 perceived to be everyday concepts, to the theoretical generalisation of the definition of t-tests. This may indicate that at university, much of the statistics that is taught to students has the potential to go beyond non-scientific, empirical generalisations, unlike in school mathematics (Kozulin, 2012).

The episode ended with the lecturer emphasising a formal definition for t-tests, "what is the question that a t-test is asking? the one that I put here" which precedes the formula for *t*-tests on the next slide. The closing lines invited the listener to actively think about the relationships between the story and the elements of a statistical model. The lecturer showed a simplified formula (as in TE:1 on p. 178) in order to facilitate the ***comparison between the statistical model and the non-statistical contexts***, with an emphasis towards the non-statistical. The lecturer appeared to me to use an abridged version of a factual reality to stimulate students to make sense of the formal, theoretical statistics by linking non-statistical and statistical types of 'contexts'.

In my memos alongside this episode, I observed the necessity of comparing different types of 'contexts' in order to provide an authentic characterisation of this lecturer's teaching. L1 reinforced my interpretation of his teaching as contextualised in Interview 2 when he suggested that he aimed to emphasise statistical sense making through non-statistical 'context' rather than statistical 'context': "[students] should understand the general logic of statistics and apply the context, rather than a formula." (L1, Interview 2, 0:13:37.9). L1 expressed his view of 'context' in Interview 2, Quote 21 as an essential learning outcome for these students. In this TE, the historical context aims to facilitate the transfer of meaning towards students' statistical sense making by which the properties of one object (the history of t-tests) are transfered to another (the scientific definition of t-tests) (Vygotsky, 2012).

> **Quote 21 (L1, Interview 2)**
> "The Gosset example is kind of trivial, but it is more than that. It actually shows how statistics can be applied to a real life context. That is the lesson for me at least, that Gosset used it for other purposes, to get good beer and psychologists are using it now to do good experiments in a way. I hope that gives [students] a sense of this idea that statistics is fun and useful and you need to find the right context, you need to understand what it does and how you use it." (0:05:04.1)

In TE:7, on whether men are better drivers than women (variables: driving ability and gender), L2 modelled to students the *claims* or hypotheses which could be made in this example. This TE was also structured in three parts: (1) provide details of a statistical model, 'the philosophy behind hypothesis testing, what is trying to do' (L2, Lecture 4), (2) explain the problem that was under investigation and (3) replace informal representations with formal statistical models. What is important to notice in this TE is the way in which L1 uses ***informal language and terminology as a precursor to formal hypothesis testing*** to highlight the importance of contextual knowledge to understand potential sources of variation and error (Gal, 2002). Although the contextual knowledge presented to students originated from real-life, L2 *adapted* and *simplified* it for the purposes of constructing hypothesis tests from the news article (H_0 and H_1) which he initially termed 'claims'. L2's main focus here was on showing students the benefits of the t-test to measure the extent to which the two groups are *different* from each other. So *concrete* situations are used here to ***make sense of a new statistical concept***, hypothesis testing. Further, L2 compared the null hypothesis H_0 with 'life being as simple as possible' or that 'everyone is the same' and the alternative hypotheses H_1 with a 'more complicated version of events', that one group might be *better* than the other (directional hypothesis) or *different* (nondirectional hypothesis). L2 then proceeded to relate these properties (variables, hypotheses) to more formal definitions of t-tests. Using statistical discursive routines in authentic, simulated or parabolic TEs, the lecturers appeared to *acculturate* the students into 'how statisticians reason and work within the statistics discipline' to develop new ways for students to view the world (Pfannkuch and Wild, 2004).

In summary, parabolic served as motivators when introducing statistical theory. Also, [P] were aimed at helping students transition from everyday, informal, intuitive understanding of statistics to more scientific, formal concept formations through the use of language made accessible for the students.

Prototypical

Relating to the first stage of statistical thinking as defined by Chervany et al. (1977), prototypical examples support students' comprehension of a problem. At the centre of Module B were worked examples illustrating the key concepts and methods using small datasets. L2's declared intention, as in Quote 22, was to show students 'prototypes' or *typical examples* of problem situations, procedures or concepts that students could *replicate*. In Modules A and B such 'worked examples' were also a focus of the summative assessment, counting for 50% and 80% of the module grade respectively.

Quote 22 (L2, Interview 2)
"So that the students hopefully have got a decent **prototype** to go and well, not quite to copy from, but replicate the procedure if they have seen it used. And hopefully when I present those kinds of calculations on the [board], in the lecture, I am doing it in a way that is similar to how you would if you were handwriting it down as well. I tried to get it [the presentation] to just display one little bit of the equation at a time so that they have to write down one bit and then write down the next bit and think about it in the same order as they would have to if they were writing it and doing it themselves." (0:08:49.9)

First, these 'prototypes' were intended for the lecturers to **model procedures and calculations**. In L2's view, the module content did not focus on the *transfer* of contextual problems to other contexts, as in Module A (TE:11 on page 199) but on providing 'prototypical' worked examples for students to *replicate* calculations. In Module B, prototypical were characterised by a reliance on hand calculations using small datasets, so that they could be computed in the classroom. In Example 7, Appendix D, students were expected to manipulate a small dataset of twelve pairs of mother-daughter heights to answer five questions. The example came at the end of the regression and correlation lectures and L2 intended it as an exercise to 'put those things [regression and correlation] together', i.e. how to apply the lecture content. L2 also pointed out in the lecture that it was 'an exam type question' (Module B, Lecture 8, 0:45:22.6). As a result, students were asked to apply formulae taught in the lecture (e.g. Pearson's product moment correlation coefficient r, the regression equation $y = \alpha x + \beta$, significance testing for correlation). The intention seemed to be for the students to be able to 'generalise' from the mothers/daughters heights situation and algorithm to other correlation and regression problems (similar to 21).

The choice of variables, mother-daughter pairs, was also *prototypical* since it relied on students' intuitive, situated and non-statistical understanding of the real-world situation where a child's height is predicted from that of the parent. So this "context" was meant to have a broad appeal to students in the classroom, at a micro level, based on macro level assumptions about this scenario. The explanation of regression relies on a personal context for its sense, seen as contextualised rationality (Wertsch, 1992). Regression is represented here in terms of its concrete particularity for which a different "voice" was required. For Derry (2013), decontextulised rationality characterises a 'voice' or social speech type, in a similar vain as Vygotsky's (2012) abstract reason. In the case of the regression example, the mother-daughter pairs represent objects and events (semantic content) in terms of formal, logical and quantifiable categories. The categories' meaning can be derived from their position in abstract statistical models independent of the lecturer's speech contexts. On the other hand, contextualised forms of representation (students' knowledge of the world which associates mothers and daughters' heights) represents events in terms of their concrete particularity (Wertsch, 1992). However, in the context of the lecture theatre, decontextualised models of discourse were privileged.

In Example 7, the calculations were carried out using calculators or by hand, with a focus on substituting 'appropriate numbers to appropriate places' or drawing a conclusion 'that there is a significant correlation between the heights of mother and daughter' (L2, Module B, Lecture 8, 1:06:20.5). I observed a similar aim in Module A where students were given step-by-step instructions in how to run an analysis using software (Example 3) and how to interpret the output, e.g. 'a one-way ANOVA revealed a significant effect for length of revision on SAT results' (Module A, Lecture 8).

A second important aim for prototypical was to **illustrate concepts and methods** for the students and **introduce 'new statistics'** for the students, including new signs, terminology (e.g. how to report or interpret results) and *rules of thumb* that are accepted within a particular module. In Example 7, p. 300, L2 used a *scaffolded* pedagogy in which he proceeded to show students the answers to each question on slides, using tables and the whiteboard (e.g. to explain significance testing using the normal distribution), while students copied answers using handouts. L2 interpreted $r_{test} \approx 3.126$ (Slide 3, Example 7) as 'moderately strong', thus showing students how to apply common rules of thumb for interpreting the size of a coefficient and offer a qualitative interpretation of the statistical output. In this case, statistical principles and practices outside the module are brought into the classroom to help interpret the output of this particular example.

Both lecturers aimed to provide some ***explanations for the principles and interpretations*** of the statistical examples. Lecturers' comments about statistics might be seen as 'discursive contributions' that support students' 'increasing participation in mathematical speaking/-thinking' (Lerman, 2001, p. 89). The challenge was to use a sufficient range of examples to help students recognise the correct statistical test for a scenario and to make sense of statistical concepts and models (Ryan and Williams, 2007).

In summary, prototypical TEs had a role in illustrating and replicating statistical procedures and calculations and influencing students' way of thinking about statistics. Ultimately however, I considered that prototypical were 'typical examples' in each module's setting and culture that encouraged the lecturers and their students to participate in joint statistical practices.

Symbolic

Symbolic TEs captured abstractions or generalisations of statistical models using formulae, definitions and graphical representations of variables. On these two modules, statistical theory preceded, followed or was introduced within a scenario that presented material in a general framework.

Instances where the **theory was introduced in the 'context' of examples** included TEs where L1 discussed how to interpret an SPSS output as 'abstract knowledge' ("because you have two rows here. Which one do you read? There is a rule", Module A, Lecture 2), despite it being linked to the context of a particular exercise. This practice was in accordance with L1's intention to teach statistics without mathematics (Quote 9, page 160). Similarly, L2 introduced decontextualised statistical theory in the context of prototypical as in Example 7 and Example 8 in Appendix D. In the latter, L2 integrated symbolic content and prototypical calculations using a brief scenario (comparing the performance of four machines in five observations) to interpret the context. In this example, L2 did not give the students the opportunity to consider reasons why the measurements for the different machines might be different from each other (between groups) or why the measurements might vary for each of the machines (within-groups). The aims here were to explain particular statistical tests and manipulate the data rather than search for patterns in the data or gain insights into a particular scenario. Instead, such questions about the data were explored within realistic and authentic TEs.

Further, L2 attempted to show students how they could use the statistics presented in the module to **address statistical questions in concrete prototypical examples**, e.g. finding the standard error of the mean or the proportion of samples that give estimates of the crushing strength of a load of bricks within a particular value of the mean. In this way, L2 made a link between symbolic with decontextualised statistical terminology and contextualised prototypical TEs. L2 also emphasised the 'nitty gritty' of statistics (L2, Lecture 3, 0:30:11.9) that included **statistical terminology as being different from *everyday language***. For example, L2 differentiated between 'estimate' as a statistical procedure, about collecting data from a sample and *estimating*, about applying a formula for estimating the population parameter based on the sample data (Lecture 3, 0:32:07.6).

Further, L2 **clarified differences in notations** for population and sample means and standard deviations in order to then explain the *central limit theorem* (*clt*) in Example 9, p. 304. Fundamental in *traditional* statistics, L2 also considered *clt* to be 'a key theoretical tool' (Lecture 3, 0:37:05.2) which gave him reasonable and useful approximations and later supported explanations of confidence intervals and the inferential techniques used in the module.

An alternative approach would have been to use software to obtain other, better approximations using simulation-based methods, including permutation tests and bootstrapping (Rice, 2006). L2 first chose to briefly state the *clt* ('if we have lots of observations then the sample mean approximately follows the normal distribution'). Other versions of the *clt* present various degrees of abstraction and generality (e.g. Chihara and Hesterberg, 2011), but in this case, since L2 was interested in its use for practical purposes, he presented a *less mathematical*, not as formal version to justify the use of the normal distribution, depending on two parameters, the mean μ and standard deviation σ of normal density.

The second step was to further explain *clt* conceptually (e.g. 'if we have large numbers of samples of independent observations, each of these samples being of size n, these samples can then be considered as random variables...'). Third, L2 explained the formulae on slide 3 in Example 9 identifying the Normal model but not the normal density equation. Forth, L2 encouraged students to use an interactive internet applet after the lecture to manipulate the population mean and sample size to obtain visual representations of the distribution of the population and how that compares to the sample distribution.

By contrast, L1 provided a description of statistical models using mathematical symbolism, which was simplified using non-conventional notations or words (TE:1, page 178). L1's beliefs about the student's mathematical level and perceived anxiety with the statistics influenced his curricular design and teaching approaches. L1 did not offer explanations about the process of solutions, how to solve equations or how to perform the necessary calculations. Instead, L1 focused on how to interpret outputs from statistical analyses and how to **work with models without mathematics**: 'what I am interested in is the logic of design first and then the logic of how you have to think of variance for instance. [...] I personally do not feel the need for mathematics. We just use SPSS' (L1 Interivew 2, 0:10:39.8). I noted how L1 showed the mathematical formula only after he had introduced the students to a model's applications and interpretations (TE:1). For example, in TE:13, L1 introduces the principle of ANOVA with a *contextual* example (predict memory scores on the basis of age and gender using an ANOVA model) rather than by showing a mathematical representation of the model.

Module A, Lecture 4, ANOVA

TE:13 "If you have more than two groups, you will use something else, analysis of variance. I am going to get to one way in a moment. A t-test allows only one independent variable. Analysis of variance allows for more than one independent variable. For example we are interested in the effects of gender and age on memory scores. We want to see if gender, one independent variable, and age, second independent variable, have an effect on how people score on a memory test. We have gender, for example males and females and we have age, let's say over fifties and less than fifty. And I want to see if it has an effect on memory scores. How many facts they remember from their holidays, for instance. That is different from what we have done before. So analysis of variance allows for more than one independent variable." (0:10:21.2)

In this case, L1 represented statistical knowledge in terms of its concrete particularity which further highlights L1's teaching "voice", dominated by contextualised representations rather than the "voice" of decontextualised rationality (Wertsch and Stone, 1985; Wertsch, 1992). In Module A, abstract, symbolic formulations are not the highest goal of statistical meaning-making (Lemke, 1997).

In summary, on these modules, symbolic involved informal versions of definitions, theorems or models to justify the use of statistics. Symbolic encouraged reflection on the relevance of models with informal graphical methods, particular situations and examples. Finally, formulae and statistical symbolism defined *general* models (abstract, deductive properties of the model but without formal proof) and showed a standard way of carrying out statistical analyses with prototypical TEs.

Overall, the characterisation of TEs in this section showed a weaving between micro and macro conditions that allowed these lecturers to make the module material *accessible* to students. The challenge in both Modules A and B was for the lecturers to present materials conducive to 'relevant learning', that resonated with students' past and future cultural experiences as well as knowledge assumed to be required outside of university.

The analysis of these TEs in which I sought to identify the synthesis between macro and micro conditions, highlighted the significance of the localised statistical practices within much broader considerations about the statistical analysis. Together, these multitude of 'contexts' portrayed statistics as a research, scientific practice and also as 'everyday' knowledge and activity, although perhaps with an emphasis towards the former, scientific practice. The analysis suggests a view of curricula in terms of its relevance for the lives of students and for "doing" statistics. Lecturers' teaching process is influenced by the contradiction between abstract and concrete representations as sources of change and development. The data shows how the two lecturers take into account the context in which teaching and learning activities take place - within institutions, modules, lecture theatres, etc. In this analysis, I looked to bridge the abstract rationality of statistics to the situated nature of the teaching activity.

5.4.3.4 Cycles of "context"

Analysing the teaching processes in the lecture theatre through the SCT lens, I was interested to explore how lecturers organise students' learning of statistics and why. To this end, I looked at transitions during teaching, across lectures (horizontal analysis) and how each lecture was conveyed (vertical analysis). In this part of the analysis, I present how the lecturers used TEs throughout the lectures in what I called "lecturing cycles of context". These "cycles of context" was another way of defining lecturers' teaching styles/approaches. To ensure a *"context"* conductive to achieving the learning outcomes, the teaching process was characterised by individual (of lecturers) or collective (lecturer, students, wider teaching community and statisticians) actions. At the micro level, the teaching I observed in Modules A and B used lecturers' discourse and prepared slides as the main mediational tools. In lectures, I observed very few and brief dialogic interactions with students on either module, with 95-100% of lecturing time being spent by the lecturers explaining, demonstrating and showing students sequences of examples. The use of slides during and after lectures played an important role in lecturers' teaching activity towards students' sense making.

In both modules, the slides used during teaching sequenced and structured students' learning. The slides were also a key reference of the teaching that took place for students, in addition to notes or voice recordings of lecturers' speech during teaching. So slides were used in both modules to scaffold or support the learning so that students could accomplish certain tasks otherwise not being able to on their own, often after the lecture, e.g. in laboratory sessions. By making the slides available during and after the lectures (following institutional rules), they were intended to aid students become independent learners. The slides used by students after lectures were intended to guide the students, help and support their learning in the absence of the physical presence of the lecturer but as an additional "voice" of the statistics being taught on these modules. In this way, the slides could highlight the importance of written language, statistical signs and symbols in the teaching that took place and emphasise the importance of social interactions through lecturers' direct support of students' statistical sense making. In Module A, the slides were projected on screen during the lecture to present statistical concepts, problems and multiple choice questions (linked to the electronic voting system). During lectures, students could take notes to capture L1's speech in addition to the summary presented on screen. In Module B, the slides were projected on a screen and were also made available on printed on paper but with gaps for students to fill in during lectures, e.g. with their own computations. In this way, L2 sought to engage the students in the communicative teaching activity that was taking place during lectures through the use of "gappy" slides.

Lecturing cycles - Introductory and integration lectures

1: Announcements/Aims/content/learning outcomes $\longrightarrow$ Revise concepts

2: *symbolic* $\longrightarrow$ *parabolic* $\longrightarrow$

3: *symbolic* $\longrightarrow \cdots \longrightarrow$ *authentic* $\longrightarrow$ *simulated* $\longrightarrow \cdots \longrightarrow$

4: *symbolic* $\longrightarrow$ *Multiple choice questions (Student feedback)*

Lecturing cycles - Revision

1: Announcements & summary of module **content** $\longrightarrow$ Exam/coursework structure

2: *Symbolic* $\longrightarrow$ *authentic* OR *simulated* OR *Multiple choice questions (Student feedback)*

3: Summary/Advice/Resources

Figure 5.11: Cycles of context in Module A lectures

Through interconnecting different types of "contexts", the lecturers sought to actively engage and ensure the participation of their audiences, the students. For L1, the contextual meaning of these episodes was motivated by his aim to provide students with a 'basic' understanding of introductory inferential statistics based on his knowledge about the students and his extensive teaching experience on this module (Quote 5.3.2, p. 155). For L2, these episodes were mainly about providing challenge to students, and also help them with what can be a difficult subject to study (Quote 18, p. 165). The use of these resources were therefore at the heart of the curricular planning for these modules as well as guided by lecturer's beliefs about teaching statistics.

In his teaching, L1 alternated between statistical theory and practical applications. Figure 5.11 shows that L1 first introduced or reviewed abstract concepts in the introduction and learning outcomes (the first five minutes of a lecture). Only after L1 discussed two to four contextually bound examples he drew out the theoretical principles, followed by further examples. So, the pattern in which L1 used stories following the schema in Figure 5.9 (page 194) showing types of TEs, moving from highly symbolic, statistical content to examples with limited, narrow statistical content and from highly embedded to reduced contextual stories (narratives). This transition from symbolic to authentic to simulated characterised this module's representational system.

Mapping the cycles of context across all the lectures in Figure 5.11 revealed one pattern in the lecturing stages of introductory (Figure 5.7a, page 176) and integration (Figure 5.7b, page 176) lectures and another pattern for the revision lectures (Figure 5.7c, page 176). The introductory lectures started with L1 making his aims and learning outcomes explicit, followed by a review of the previous week or term's topics. In the second cycle, L1 started

Lecturing cycles - New topic

1: Announcements/Aims/content$\longrightarrow$

2: *symbolic* $\longrightarrow$ *parabolic* $\longrightarrow \cdots$

3: *symbolic* $\longrightarrow$ *prototype* $\longrightarrow \cdots$

4: *symbolic* $\longrightarrow$ *prototype* OR *simulated* OR *authentic* $\longrightarrow$

5: Summary/Advice/Resources

Lecturing cycles - Revision

1: Announcements & summary of module **content** $\longrightarrow$ Exam/coursework structure

2: *symbolic* $\longrightarrow$ *prototype* $\longrightarrow \cdots$

3: Summary/Advice/Resources

Figure 5.12: Cycles of context in Module B lectures

new material with *symbolic* followed by *parabolic* contexts. The third cycle in the lecture was characterised by a cycle of *symbolic* representations followed by *authentic* or *simulated* scenarios. Finally, L1 would provide further *symbolic* explanations (in lectures 1-3, 5, 6) followed by multiple choice questions using electronic devices (lectures 1-4 and 6).

Table 5.3 (p. 197) shows the larger volume of prototypical and symbolic cycles in Module B as well as the larger amount of time spent on these two types (approximately 70% of the module). Figure 5.12 shows that Module B's teaching pedagogy involved predominately a pattern of weaving symbolic and parabolic and symbolic and prototypical TEs. Conversely, Module A allocated more time along the non-statistical dimension, with an equivalent proportion of episodes (approximately 70%) spent in the simulated/authentic quadrant).

The pattern in Figure 5.12 was altered in two lectures. In lectures 4 (hypothesis testing) and 9 (ANOVA) L2 started from a parabolic TE, a newspaper article in TE:7, p. 183, and the research questions a t-test or an ANOVA may answer. L2 then moved towards a symbolic TE. In these two cases, the pattern described in Figure 5.12 was temporarily reversed in the second lecturing stage from parabolic to symbolic TEs, although the pair symbolic/ prototypical was maintained. This cycle parabolic-symbolic is similar to the pattern in Module A, TE:1, p. 178, in which L1 scaffolds statistical knowledge using the pattern decontextualised scientific concepts (research methods and key statistical concepts), followed by the history of t-tests and further scientific concepts (a basic formula for t). L2's focus on exam practice meant that the revision lectures were 'atypical' and suggested to me that the message conveyed to students was that statistics, as assessed in the module, focussed on mathematical logic ahead of meaning-in-context. However, example 11, page 306, showing the kind of *advice* L2 offered students at the end of the final lecture 11, indicated that the process of teaching I describe here, although it mirrored some of the elements of a mathematically-based curriculum focused on formal inference, it also appeared to integrate abstract deductive reasoning with some meaning-in-context (Cobb, 2015).

In summary, with this analysis of cycles of “context”, I was able to identify similarities and differences in the teaching of statistics between the two modules. On the one hand, both modules used cycles of “contexts” to help students’ achieve the module learning outcomes. Lecturers’ actions and transitions from one episode to another in Figures 5.11 and 5.12 were broadly similar. However, Module A suggested a ‘fading’ of the concrete context while Module B ‘built up’ from abstract to concrete representations. These differences were not directly apparent in the intended curricula in the module specification and were observable only through this analysis of process of implemented teaching.

5.4.4 Micro model of teaching

The lecturers’ macro-planning involved the preparation of decontextualised knowledge materialised on slides. The micro-teaching was mediated by lecturers’ discourse that contextualised the mediational tools in the learning resources. The teaching activity revealed a contradiction between learning statistics “by doing” and learning “by abstraction”. When compared across setting (the two modules), it becomes apparent that knowledge which I understood to be decontextualised is in fact another form of contextualised knowledge and the teaching observed is contextualised (Lave and Wenger, 1991). The lecturing activity involved telling, narrating and transfering from context to context. The analysis of these two statistical modules (cases) at the micro level gives an indication of the diversity of lecturing practices and processes. Key to lecturing practices are the mediational tools, the ways in which lectures are organised and lecturers enact their pedagogical and statistical discourses (Vygotsky, 2012). The micro-model emerging from this analysis can be conceived from different perspectives as a collection of micro units of analysis (Zinchenko, 1985), in my data the lecture organisation, pedagogic discourse, statistical discourse and lecturing cycles (structure, origin and function).

The analysis of the ways in which lectures were organised showed two patterns. Module A taught the same topic over several lectures using chaining forms of organisation. As a result, I categorised three types of lectures: new topic, integration and revision. L1 used each of these types for different purposes for the students. Module B was organised at macro-level with longer lecturing time and as a result taught a new topic each week, with two revision sessions in the middle and end of the module, to coincide with the assessment schedule on the module. The consequence at the macro-level for the lecturer was that L2 covered a broad curriculum in two types of lectures, new topic and revision in a hierarchical form of organisation.

The pedagogic discourse was organised around a number of routines or strategies (Table 5.2): explanation, motivator and questions. Each of these routines were directly influenced by the object of the activity, the students and could be explained in terms of lecturers' developmental teaching using a SCT lens.

The statistical discourse I categorised in terms of the "contexts" used in the teaching resources (examples, problems or exercises) as authentic, simulated, parabolic, prototypical and symbolic. These were explained in terms of two types of lecturing discourse, statistical and non-statistical. The conceptualisation of these different types of "contexts" involved lecturers' statistical and pedagogical knowledge. Both L1 and L2 used discourse and narrative as a teaching style. Cycles of "contexts" characterised L1's teaching as relying on the non-statistical axis of the model and L2 on the statistical axis (Figure 5.10) which contrast abstract, symbolic formulations at one end and concrete at the other.

5.5 A model of teaching in context

My representation of a model of teaching statistics is alongside two axes, the macro-micro and the external-internal continua, depicted in Figure 5.13. In this model of teaching statistics in context, I considered the distinction between the two types of context, the social situation or environment and the ways in which cognition can relate to that context (Cole, 1996; Cole and Gajdamaschko, 2010). In Section 5.4.3.3, I conceptualised the *objective* statistical axis in Figure 5.9 to differ radically from the *subjective* non-statistical axis. Vygotsky (1978b) promotes the scientific analysis, an internal system of meaning created through social interactions, to reveal internal, individual differences which are "hidden by external similarities" which are more abstract or generalised (p. 63). Lecturing as cultural development is a historically elaborated process which mediates students' learning (Cole, 1985).

Based on data of intended curricula (Figure 2.3, p. 17) from institutional documentation of module planning (quadrant 1) and interviews with lecturers in the Pilot and Main studies (quadrant 2), my conception of macro-teaching is of *movement* with a *spiral* form (Section 5.3.5) composed of social, objective institutional regulations or statistics as subject matter representations (quadrant 1) and individual, subjective and abstract culture, norms and values lecturers held during teaching activities and cycles of planning (quadrant 2).

Cole (1996) suggests that the social context of development is the objective environment. Statistics, as vertical discourse does not consist of culturally specialised segments but of specialised, explicit knowledge (Bernstein, 1999). In Figure 5.13, the macro-objective quadrant 1, contains the objective elements of society to include statistics as a discipline, university regulations and intended curricula, which I discussed in Section 5.3.3. Decisions regarding module planning and intended curricula based on university regulations, expectations in society about what constitutes learning statistics at university and statistics as a domain worthy of enquiry can be characterised within the external-macro world (Section 5.3.3).

From a SCT perspective, Cole (1996) describes tools as both ideal (conceptual) and material and emphasises the cultural-historical nature of the development of subjective, ideal cognition at the macro-micro levels. In statistics education, Gal (2002) considers that statistical literacy involves both knowledge which involves literacy skills, critical questioning,

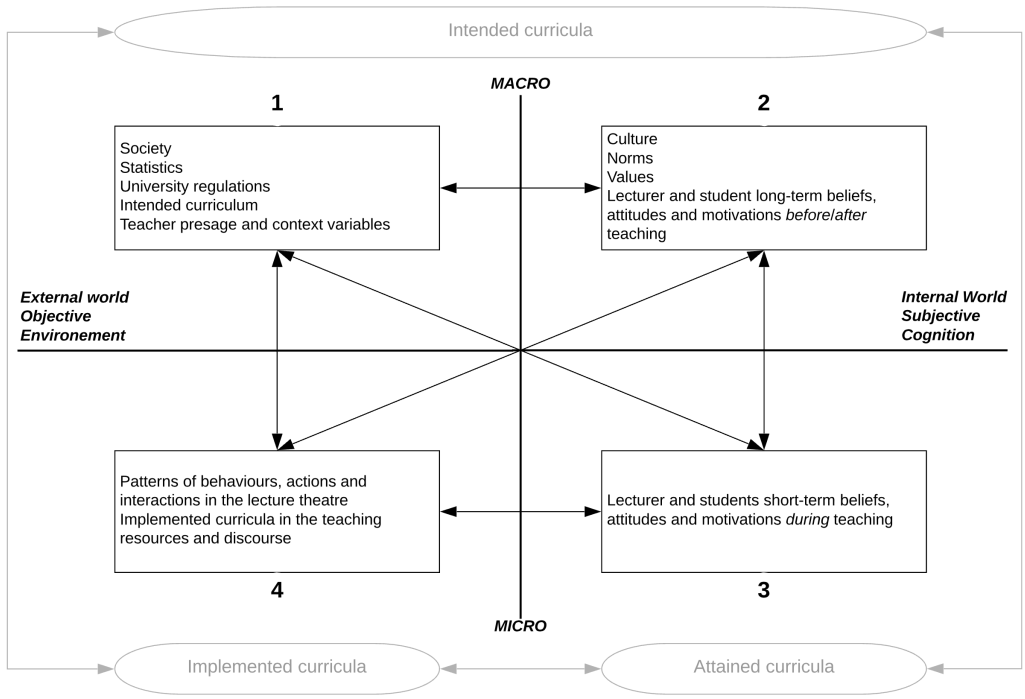


Figure 5.13: A model of teaching statistics in context

statistical, mathematical, and contextual knowledge (in quadrant 1) and also dispositions. Dispositional components are about critical thinking and attitudes towards statistics, people's beliefs and attitudes to enable positive problem-solving. In his model of statistical literacy, Gal (2002) considered that statistical skills include both statistical ability but also the set of cultural practices that people engage in. Similarly, Nisbett et al. (2009) highlighted the importance of cultural prescriptions to reason statistically about events and also recognise that formal training can improve people's statistical reasoning about everyday events. Thus, in quadrant 2 of the macro-subjective, I include culture, norms, dispositions, traditions and values as depicted in lecturers' beliefs about teaching, statistics and the world (Section 5.4.3).

Historic iterations of a module teaching are seen as separate yet related and interconnected activity systems due to which later and current activity systems are realised. Each of the nodes rather than being static, they are dynamic, subject to change and transformation through time. On ModuleA, the same subject mediated students' learning through ever changing tools. On ModuleB, the module team, the subject of the activity, changed yet was planning teaching using the same tools. The object of the activity might be conceptualised differently, depending on the subject. Essential at macro-level activity in quadrant 1 are the historical nature of human experience, the objective social environment and experiences of other students and the existence of subjective prior beliefs and schemas in quandrant 2

(Kozulin, 1998).

The micro-analysis relied on lecturers' discourse, lecture resources and learning objectives during teaching (quadrant 3) and lecture organisation and cycles of context (quadrant 4). The micro-context appears as a movement from one context of activity to another, as interactions between objective and concrete patterns of behaviour which I could observe at the micro-level inside the lecture theatre (lecture organisation, cycles of context, actions and interactions) in quadrant 4 and the subjective beliefs of lecturers about curricula, students or statistics, both during teaching and during planning (before/after teaching) in quadrant 3.

Lecturers and students' beliefs, affective attitudes and motivations were also specific to the situation inside the lecture theatre. The interview data suggested that lecturers' beliefs about statistics, students or teaching, as mental constructs, represented the codification of the lecturers' experiences and understandings about themselves as teachers and as statisticians, the nature of statistics , learning, students and institutions. The subjective world of the individual mental life is represented in quadrant 3. Here, I include the imagined, subjective worlds of the personal capabilities of individuals who actively construct context in action. The Pilot interviews revealed what lecturers perceived to be a challenge in statistics education and statistics as a scientific activity of putting real-world into mathematical frameworks, i.e. connecting and relating quadrants 1 and 2 in the micro-teaching discourse based on lecturers' beliefs about statistics and about teaching. For example, critical beliefs of how to integrate abstract statistical concepts within an external, non-statistical context characterised L1 and L2's views on the teaching of statistics at university.

Bernstein (1999) explains that horizontal, context-specific discourse reinforced through pedagogy facilitates the development of a range of operational 'knowledge'. In quadrant 4 of the micro-objective, I include patterns of behaviours, actions and interactions in the lecture theatre which I could observe directly but were context dependent, embedded in ongoing practices, usually directed towards specific, immediate goals and highly relevant to the people in that activity (Bernstein, 1999). For example, implemented curriculum, lecture organisation pedagogic and statistical routines were positioned within the micro-objective context. Lecturers' intentions and objectives for their students and students motivations were however issues of the micro-subjective context. The lecturers' cycles of "context" (Section 5.4.3.4) were intended to bridge the macro-macro with the objective/external-subjective/internal dimensions.

The ways in which lecturers and students use tools inside the lecture theatre in the relationships between quadrants 3 and 4 *transforms* the model of contexts. The model of contexts which are nested, that surround and those which weave together implies that the 'rings' may be reshaped, transformed and interpreted in a multitude of ways. In my model of context therefore, I had at one end the objective social environment and at the other the subjective cultural cognition which is transformed through 'active construction of context in action' (Bernstein, 1999). On the horizontal axis, I included the relations between objective-subjective, ideal-material that are fundamental within vygotskian frameworks which distinguish between objective, “external” physical objects (sense) and their subjective meanings when used for particular purposes (Section 4.2.1.5). Bakhurst (1997) states that the ideal “involves a resolute defence of the objectivity of ideal phenomena, which are said to exist as aspects of our spiritual culture, embodied in our environment” (p. 34). At the same time, on the vertical axis the model includes my main foci of analysis, the macro-micro dimension (Hammersley, 1993a).

The analysis of external teaching activity can reveal insights of the internal, often hidden activity. Macro-micro relationships could be sought between the ‘average student’ as defined in the module specification (quadrant 1) and the implemented curricula in the lecture organisation and cycles of context (quadrant 4). Further, relationships between the undergraduate students’ intention before the module and lecturers’ intended curricula (quadrant 2) and the realisation of these intentions in implemented and achieved curricula (quadrant 3) could give rise to contradictions in the teaching activity. In this way, the model in Figure 5.13 could be used to characterise the teaching process and curricula.

However, the model presupposes commonality between the contexts analysed. This common ground makes it possible to transfer from one context to another (Greeno, 1997; Derry, 2013). Teaching statistics from concrete, authentic representations to abstract representations is intended to enculturate the students into using different forms of representations. Lecturers’ explanations provide formal rules and procedures which are re-interpreted for students. The lecturers transfer meaning of the statistical models from a real, authentic application (representation) to an adapted or transformed representation. By treating the adapted representations as if they were a real statistical analyses, the lecturers can “act out” a meaning. This means that planning lectures (quadrants 1 and 2) and lecturing (quadrants 3 and 4) requires expertise and a deep knowledge of the subject matter (Davydov, 2010; Hadegaard, 1990).

5.6 Summary of chapter

In this chapter, I presented in some detail the process of my data analysis and interpretations at two levels of analysis: micro and macro. At the macro level, my analysis was structured around the three components of the basic mediational triangle. Using data collected during the pilot and main studies in interviews with lecturers, module documentation and institutional regulations, I gave an account of the curricular planning on two introductory statistics modules and of lecturers’ beliefs regarding their planned curricula. At the micro level, the analysis focused on the interview and observational data from lectures to characterise the teaching strategies in the teaching of statistics using lecturers’ planning/organisation and lecturing discourses on these two modules. The outcome of the analysis was a macro-micro model of statistical teaching practice at university. The model can be used to characterise teaching styles and experiences on statistical modules. In the next Chapter 6, I aim to put my findings into a broader context and discuss the contributions of my work to statistics education research and teaching.

6 Summary and Conclusions

Chapter contents

In my study, I have characterised the teaching of introductory university statistics as a 'service' subject on two modules from a sociocultural perspective on teaching. My aim was to capture specific insights about the teaching of statistics in context using several sources of data. In my pilot study, I interviewed twenty lecturers using the repertory grid technique to identify aspects of how the lecturers perceived intended curricula, their beliefs about specific teaching approaches and students' learning (Chapter 3).

The process of analysing interviewees' beliefs about the intended curricula supported my main study methodology (Chapter 4) and data analysis (Chapter 5) investigating teaching processes. To focus on two levels of analysis of context, macro and micro, in my main study, I used sociocultural theory as an interpretative lens with a grounded analytical approach to analyse observational data and characterise the teaching of statistics at university of two modules. The outcome of my data analysis of the statistics teaching activities on two statistics modules in Chapter 5 is a model of teaching statistics in *context*. In this final chapter, Chapter 6, I draw together the different sources of data and facets of the analyses presented in previous chapters.

First, I consider answers to the first research question *"what characterises the teaching of statistics at the macro-level?"* in Section 6.1. To answer this question, I draw on findings about ***what*** lecturers wish to teach and ***why*** from the Pilot study (Chapter 3), the Main study macro-analysis (Section 5.3 and also Section 5.4). Second, I look at answering my second research question *"what characterises the teaching of statistics at the micro-level"* in Section 6.2. Here, I draw on findings about ***how*** lecturers teach (lecture) inside the lecture theatre and ***why***. Answers to this question involve my analysis of lecturers' teaching in the lecture theatre (Section 5.4) as well as relationships between the intended and implemented curricula, between macro-micro (Section 5.4.4). Next, I reflect on the contributions of my research to the teaching profession in Section 6.4 and the research community in Section 6.5. Following this, I reflect on the quality of my investigation in Section 6.6. Finally, I conclude this chapter with an overview of future research in Section 6.7. In Table 6.1, I summarise the main conceptualisations that emerged in my study and which I discuss in this chapter.

Table 6.1: Components of analysis

	Teaching	**Statistics**
Macro	Module intended curricula	Statistical professional practices
	Lecturers' beliefs about teaching	Lecturers' beliefs about statistics
Micro	Pedagogic discursive routines	Statistical discursive routines
	Lecture organisation	Cycles of "context"

6.1 What characterises the macro-teaching of statistics?

Teachers' belief systems, teaching processes and classroom practices are thought to make a positive impact on student attainment of the learning outcomes and attitudes since, in turn, they are assumed to influence teachers' curriculum decisions-making and teaching approaches (Pajares, 1992; Calderhead, 2004). Teachers' belief systems, composed of 'beliefs' connected to one another and to other cognitive/affective structures', reflect their personal beliefs or attitudes about education, about the subject matter and about the students (Pajares, 1992, p. 316). In my pilot study, I considered Munby's suggestion that it was important to investigate the relationship between beliefs and lecturer behaviours and decisions (Munby, 1982). Thus, I explored lecturers' planning of undergraduate statistics education and in particular what the lecturers in my study perceived to be important for students to learn on statistical modules at university.

In this study, I assumed that lecturers' conceptions of standards such as learning outcomes are tacit knowledge and so difficult to examine and articulate with other methods (Sadler, 1989). My methodology used the repertory grid technique (Kelly, 1955). Repertory grid (RG) interviews were designed to identify and investigate the relationship between lecturers' constructions of different aspects of the intended curricula of university statistics. The resulting data from twenty participants was in the form of statements (words) and ratings. It was therefore possible to analyse this rich dataset using quantitative and qualitative methods.

Following this study, I was able to produce a 'rough' conceptual framework (Figure 3.3) which guided the research design of my Main study (Section 4.4). The research design of the Main study included observations of the teaching of statistics inside the lecture theatre and interviews with the lecturers. In the end, I analysed the teaching of two statistics lecturers, a Psychologist teaching statistics to Psychology students and a Mathematician teaching to Engineering students.

The main source of primary data were macro-interviews and documentation. The data collected *before*, *during* and *after* teaching I could interpret through a Vygotskian theoretical lens using a grounded analytical approach. My analysis used the basic mediational triangle in Figure 4.2 seen as the context for understanding teaching actions carrying out multiple layers of analysis (Section 4.2.1.5).

Recent research in statistics education (Chapter 2), has highlighted the importance of teaching statistical concepts and methods (computations) to non-statisticians using real-world or authentic tasks (Broers, 2006; Seabrook, 2006). The data suggested that the main issue for statistics pedagogy is to engage in and integrate both (advanced) statistics as a scientific activity and statistics as embedded in authentic or real-life situations. The connection between scientific practice contained in objective statistical analysis and cultural practices in subjective interpretations of the world were fundamental issues for a model of teaching (Section 6.3).

Lecturers believed that their role in the lecture theatre was to show students how to use statistical conceptual tools in 'authentic' activity (Brown et al., 1989). The issue of how to integrate "imagined" worlds in the curriculum is reflective of Vygotskian issue with distinguishing between school learning activities and authentic activity. Vygotsky (1978b) saw the role of the teacher to provide students with the external signs or tools and transform these tools into internal signs as means of remembering through speech, since 'the very essence of human memory consists in the fact that human beings actively remember with the help of signs' (Vygotsky, 1978b, p. 51). Vygotsky used the concept of logical memory which is voluntary and made possible by mediational tools. Vygotsky further put forward that 'in this way, external forms of mediated behaviour develop' (p. 45). Thus, by using tools as aids to memory, 'human beings are able to control the conditions of their future remembering' (p. 202) and are able to voluntarily 'remember by constructing narratives which require the recall of past events for their intelligible completion' (Bakhurst, 1998, p. 204). In lectures, lecturers attempt to offer students symbolic devices in conjunction with statistical theory to aid recall and understanding. Thus a scatterplot or realistic TE might be employed by these lecturers to help students remember a statistical model. In his psychological experiments across age ranges, Vygotsky (1978b) concluded that adolescents and adults' mnemonic systems could be disrupted if they were asked to use external aids.

The lecturers in my study had a deep concern for the planning of curricula, including which teaching resources are most beneficial to students' development. There was less agreement for the participants in the Pilot and Main studies on how to integrate abstract, statistical concepts in a non-statistical context. For some of my participants, students required the statistical language while other lecturers believed that students needed concepts in a "context" with interpretations following conventions but importantly without mathematics

In general, for the Pilot study participants, statistical thinking related to knowledge of statistics, ability to communicate with data in reports, reading quantitative studies and interpreting the meaning of data. Statistical thinking was contrasted to objective, routine analysis and techniques. For different lecturers, different statistical models or techniques seemed relevant, depending on the students' programme of study or lecturers' perceptions about which skills are valued in employment. This was in contrast to the curricular reforms reviewed in Section 2.4.2 which aimed to propose a more standardised approach to the teaching of statistics across in-service modules. For the lecturers in my study, the constraints present in their macro-context, such as students' background, hours in the curriculum, availability of resources influenced their beliefs of statistics curricula and shaped their planning.

The concept of ZPD (Section 4.2.1.4) assumes that subjective, mental capabilities function independently of the material within which they operate. However, the data analysis in Chapter 5 shows that the opposition between external and internal worlds can be limited (Zinchenko, 1985). For instance, L2 saw the study of statistics as a useful activity to students' more general academic development and the development of statistical thinking/reasoning as beneficial for the development of other abilities. Students' background, interests and motivations were fundamental to lecturers' planning. Some interviewees in the pilot study believed that students are able to follow *recipes* while other lecturers looked to stretch their students who could think in a mathematical way and know what the statistical *recipes* were made out of.

The majority of interviewees focused on the two later stages of the statistical process, analysing data and interpreting the results. L1 seemed to support this view that students might benefit from an understanding of where the statistical analysis leads before formulating research questions (possibly given to students) or collecting data (too time consuming). However, L2 saw statistics teaching more like a 'conversation' and a 'normal thing to do', implying that statistical reasoning was a key goal in his perception of 'ideal teaching'.

Kozulin (2012) noted that the difference between scientific and everyday concepts is their content, theoretical versus empirical. The comparisons of the two modules' teaching contexts revealed this difference between scientific and everyday concepts, i.e. a contrast between theoretical statistical representations and practical uses of statistical representations. For L1, the purpose of statistics education was to develop in students a practical worldview of statistical representations. By contrast, L2 saw the purpose of statistics education to develop in students a statistical worldview which required some theoretical concepts, including mathematics, to "open up the black box" (16, p. 164). The distinction between scientific concepts and everyday concepts in statistics education is however more difficult to discern distinctions between scientific and everyday concepts in the process of teaching. This is due to the multifaceted nature of statistics which involves statistical and mathematical knowledge and considerations of the contextual knowledge of the target domain. Davydov and Markova (1982) (Section 4.2.1.4) define ZPD as the distance between the cultural knowledge and scientific concepts which is made accessible through formal teaching and the everyday experiences of individuals. The data suggests that statistics lecturers' conceptions of teaching statistics is influenced by their institutional context, cultures, values and resources (Gordon et al., 2007). Overall, the interviewees seemed to focus on statistically-based outcomes, rather than on personal or transferable skills.

The module *planning cycles* identified in this study suggest continuity but also change or interruption in working practices which can disrupt the creation of tools, such as planned curricula (Laufer and Glick, 1998). For example, L2's sudden arrival on the module team meant that he *felt* 'less prepared' than he would have liked to be. This lead to what L2 considered to be minor changes to the way the teaching resources were eventually used in the lecture theatre. Based on my Pilot study observation notes appeared to be less interaction with the students than in the lectures I observed in the Pilot study 3 (Section 3.3).

In summary, in my macro-analysis I have captured three dimensions of module planning: (1) *teaching with "context"*: the tension between planning for theoretical knowledge and understanding (statistical) versus applying statistical methods and techniques in a real-world, authentic context (non-statistical), (2) *contextual constraints*: the tension between various statistical process components, e.g. allocating teaching time for data collection versus analysing or interpreting data. and (3) *student learning*: what students can learn versus what students should learn, curriculum design.

The macro-analysis of lecturers' planning and beliefs demonstrated that they were bringing into the lecture theatre considerable amounts of knowledge about statistics and also about aspects of teaching, crucially the students and also the university environment. The lecturers' history with the students helped them understand the students through past social interactions and in more abstract terms as defined in curricular documentation. Through understanding of the context the lecturers' planned the teaching and the context in which teaching takes place can further help understand what the lecturer does and why (Gordon et al., 2007).

6.2 What characterises the micro-teaching of statistics?

At the micro-level, I used multiple units of analysis to trace the changes and evolutions in the lecturers' teaching practices and tools as they related to each other. During my micro-analysis, I considered what was taking place at any point during the lecture from the lecturers' perspectives. Inside the lecture theatre, there are various rules, routines and moves for implementing the planned teaching. The lecturers in my study, as demonstrated in the macro analysis, were equipped with various beliefs about the purposes of student learning at university, the body of students in front of them, the specific learning goals and plans for the lecture (as evidenced on slides) and the module or programme as a whole.

The micro-transcripts of lecturing brought out different or multiple “voices” of teaching as continuous re-contructions of teaching practices (Bernstein, 1999).

An initial unit of analysis was the **lecturer’s organising the teaching process** using lecturing as a teaching method. By looking at the way lecturers organised the lectures, I could answer the question on *how* the lecturers adopt such a teaching method when teaching introductory statistics. Lectures were characterised into three types: new topic, integration and revision. Each lecture type was organised differently, depending on the purpose of the learning.

Module A used all three types to teach the two main topics planned for the students. The new topic lectures were generally characterised as “hierarchic” in form since they presented statistical theory first, followed by examples. The use of *integration* lectures meant that L1 could use the same “context” or case several times during the module and so chaining of information characterised this style of teaching. In this way, L1 connected the learning across the module. Finally, revision lectures were a combination of hierarchic and chaining for L1 and hierarchical with complex forms of comparing and contrasting for L2. The analysis of lecture organisation revealed the lecturers’ thinking in concepts, the internal connections of things as well as connections and relationships between concepts (Vygotsky, 1998).

My next unit of analysis related to the **lecturing pedagogical routines** or ‘moves’ to further understand how lecturers teach and also why. The variety of pedagogical routines of explanations, motivators and questions used on Modules A and B describes the rich lecturing practices of my participants. The requirement for presenting ‘numbers with a context’ in statistics lead lecturers to use a variety of explanations to provide details about statistical models (with or without the theoretical underpinnings) with a theory-example-theory strategy. Motivators were used to ‘hook’ students into the lecture and deal with possible student anxiety with statistics (Quote 6, p. 156). Finally questions were used on both modules to facilitate verbal and largely non-verbal communication. Rather than passively sitting in the lecture theatre, I assumed that the students were able to give lecturers feedback throughout the lecture using non-verbal cues.

With the unit of analysis of the **teaching of statistics using “context”** in the problems, exercises or narratives during lecturing, “things” such as teaching/learning resources, statistical symbolism or verbal discourse contribute to the creation of meaning inside the lecture theatre. However, the connection between lecturers’ beliefs and actions is not necessarily straightforward (Schoenfeld, 1998). The use of *teaching with “context”* unit of analysis allowed fresh insights into lecturers’ statistical routines inside the lecture theatre.

In this analysis, I considered that **lecturing routines** embraced the messy, improvisational, unexpected qualities of teaching and learning *context* of making meaning of school situations (Preskill, 1998; Connelly and Clandinin, 1990). In my study, discursive routines and narratives are therefore situated, connected to lived experiences and embedded in a *context* (Pulvermacher and Lefstein, 2016). To become members of a particular professional community of practice, students need to learn representations of practice and use them correctly (Goodwin, 1994). Similarly, learning statistics is an opportunity to be exposed to statistical models, problems or authentic situations of practice which can be used in other future contexts, e.g. with "authentic" or "simulated" TEs.

L1's statistical discourse was grounded in a series of 'cases' or problems in a 'context'. The affinities between the two main topics of the module, t-tests and ANOVA, allowed him to present these cases across several lectures and different content. L2's lecturing was also characterised by a large number of problems or exercises which he presented to students. In his case, the focus seemed to be on teaching for theoretical and procedural knowledge, although this varied depending on the topic of the lecture.

Previous studies have focused on the use of metaphoric discourse as teaching devices to enhance learning (Glynn, 1974; English, 2007; DelMas, 2004; Groth and Bergner, 2009). In my study, the focus on types of statistical routines revealed important aspects of lecturers' teaching. For instance with "authentic" and "simulated" routines, the lecturers emphasised statistical reasoning and thinking. With such examples, the lecturers were able to connect one domain of experience to another (e.g. statistics and the real world) and in *thinking with concepts*, the lecturers created meaning for the students from that connection.

In other cases, using "parabolic" examples, before or after representing abstract statistical theory in formulae, lecturers depicted concrete cases to replace the abstract theory (Presmeg, 1992). The lecturers further used "everyday" language to guide the reasoning process of the students visualising the statistical models. L1 for example in TE:1, p. 178 made sense of the statistical concept of sample by relating it to concrete objects such as the production of Guinness (Jacob, 1997). L2 used differences between driving ability of men and women in a sample as a representative part of a whole (Watson and Moritz, 2000). In this example (TE:7), L2 appeals to prescriptive gender and language cultural stereotypes. Such example were however lacking in statistical detail thus allowing informal, non-statistical language to represent a theoretical statistical model.

Two other types of examples, "(proto)typical" and "symbolic" used in lectures focused on statistical theory to the expense of 'context'. Statistical concepts such as averages were conceived first as abstract representations. When data were used to compute the average for example, the numerical values obtained through these calculations represented concrete objects. The lecturers were then able to connect abstract symbolism to the concrete objects.

Lecturing cycles of "context" described different patterns of teaching process on the two modules, thus allowing me to connect process with the content of teaching (Gage, 2009). Transitions from one "context" to another meant that lecturers were able to emphasise key statistical goals of reasoning and thinking. On the surface, the lecturing materials might not reveal the depth of knowledge that lecturers' discursive routines unveiled.

L1 in particular started new topics by highlighting the key statistical concepts required to access the learning. Symbolic content was then immediately followed by authentic or simulated. L1's style of teaching was to weave together statistical concepts and theory with everyday "context". L2 transitioned from symbolic to parabolic in the first lecturing 'move' and then repeated the symbolic-prototypical cycle about three times during the lecture. Only at the end of lectures, L2 discussed authentic problems, once he was confident the students had seen the theoretical underpinnings.

6.3 A model of teaching statistics at university

Pedagogic practice interpreted from a vygotskian perspective on teaching is a fundamental social context through which cultural reproduction-production takes place (Daniels, 2001). In the model of teaching of statistics that I proposed in this study, I attempt to make various teaching practices explicit and open to further analysis and debate. The model of teaching includes descriptors of teaching actions on two dimensions, macro - micro and objective/environment - subjective/cognition. At macro level I describe the lecturers' teaching activity mediated by tools such as statistical knowledge, curriculum, university regulations within the objective environment. Also at the macro level, the lecturers' beliefs about planning teaching focussing on "context" or on statistical theoretical knowledge were influenced by external tensions and constraints, such as time allocated to teaching, existing teaching resources which could not be changed. Lecturers' set of beliefs was also internal, reflecting their own set of expectations and experiences. The macro dimension interacted with the lecturers' behaviours, actions and interactions in the lecture theatre mediated by teaching methods and routines.

The visual representation of my model (Figure 5.13, p. 213) is intended to have a reflexive function and summary of the written argument. The intention for example was to show that teaching can be planned and/or implemented in different ways, reflecting different conceptions of teaching. The model further suggests the possibility of studying relationships amongst all nodes at macro-micro and objective-subjective dimensions and different pathways through the system.

The model is not intended as an absolute "truth" of the teaching I observed in my study and does not claim to capture lecturers' full decision making accurately. As with other models, it aims to simplify the complex phenomena I dealt with in this work. Some of the behaviours and actions might not be represented, such as the complexity of students' actions. However, my intention with this model is to reflect the lecturers' behaviour in the activity of teaching to try understand it a little bit better.

6.4 Contributions to the research of university teaching practices

My work fits within the traditions of sociocultural studies of developmental teaching and learning. Previous research has looked at curricular planning (Section 2.5.1) and student learning and difficulties with statistics (Section 2.6.4) and at lecturers' beliefs. However, previous research in statistics education has not looked at putting together macro-micro analyses of statistics lecturers intended and implemented curricula with a focus on components of teaching in action.

My study has contributed to statistics education in three different ways:

- A methodological approach using the basic vygotskian triangle.
- A characterisation of lecturing practice of university statistics.
- Two accounts of teaching introductory statistics at university.

The application of sociocultural theory in my study was an effective tool for understanding teaching (lecturing). Sociocultural theory explains that subjects and objects are indirectly connected through tools. The study of mediational (indirect) tools in my study in relation to the subject-object relationship was a way of representing the teaching processes and content and helped me understand the phenomena I was observing.

The importance of sociocultural theory in my study is that it brings attention to lecturers' actions, including curricular planning and teaching in the lecture theatre. The lecturers in my study had a complex set of beliefs about teaching, the students and about statistics. So the basic mediational triangle helped me capture what lecturers did and why they did so.

The second important contribution of sociocultural theory was that it brought the context of teaching and learning into view. When lecturers engage in teaching actions, they interpret their teaching within their own belief structures and schemas. The ways lecturers teach statistics reflect their own personal goals and experiences with teaching, students and statistics (Gordon and Fittler, 2004). Lecturing activities are shaped and influenced by the histories and cultures of the university in which they carry out the activity, nested within regions, countries and the world. Social factors within and outside the lecture theatre also influence lecturers' teaching practices. In my study it was possible to make sense of lecturers' actions within a complex interpretation of the context of the teaching.

In my study, I characterised the lecturing of statistics as a complex process realised in lecturing actions. The data in the Pilot study underpins a focus on tensions in developmental teaching. The Main study observations illustrate the agency of lecturers in dealing with these tensions and constraints. As stated in Chapter 2, statistics educators have expressed strong beliefs about the value of lectures and lecturing. However, there are no empirical studies which have investigated the lecturing method from the perspectives of lecturers. The micro-analysis presented Section 5.4 contributes to addressing this lack of studies of university lecturing practice and shows the central role of representations in the teaching and learning processes.

The study of cultural tools used in lecturing presented in this analysis can be used to describe and analyse teaching practice in statistics more generally. In particular, the classification of different types of lecture organisation, although not exhaustive, I have not found in the literature. The classification highlights the importance of transitions and time spent on different episodes of learning through the lecture.

The classifications used to describe pedagogical routines primarily drew on previous research by Bellack et al. (1966) in the context of school-based teaching and Viirman (2015) in the context of university mathematics. While the classifications are similar, my study offers a different perspective on the importance of explanations, motivators and questions specific to statistics lectures.

The classification used to describe statistical routines I have produced through my analytical process. I believe that a focus on the statistical and non-statistical dimensions of the lecturing discourse and teaching resources is applicable more broadly, to other statistical modules at university. Moreover, the analysis of cycles of "context" reveal that statistics lectures do not have to be primarily about the passive transmission of knowledge and that the lecturing activity can be conducive to statistical reasoning and thinking. My interpretations of the data confirm previous research by Rodd (2003) in the context of mathematics education, as discussed in Section 2.6.1, p. 49. In my data, the lecturers encouraged the students to uncover the 'subtext' of lecturers' discourse, which I interpreted it to be similar to uncovering the 'subtext' or thoughts hidden behind an actor's lines in a play (Vygotsky, 2012).

6.5 Contributions to the teaching of university statistics

In my study, lecturers had different views about the teaching of statistics in terms of what to teach and how to teach. This thesis can offer lecturers and teachers of statistics the opportunity to reflect on the range of teaching practices available. Statistics educators could use this study to compare their own teaching practices to the two cases depicted in this study and identify similarities or alternatives they might not have been aware of. This study aimed to represent different practices and viewpoints which means that I was able to offer a view of statistics education which includes a range of contexts with diverse cultures, norms, values and beliefs.

My depiction of *context* offers the statistics education community a description and analysis of teaching inferential statistics to non-statisticians. The multifaceted teaching approaches observed in my data provide an alternative perspective on the debate between teacher and student-lead teaching styles. An implication of the developmental teaching interpreted through the sociocultural lens is to view lecturers' interpretations of statistics captured in their pedagogical and statistical discourses as having been developed through their experiences with Statistics, Psychology, Mathematics or Engineering through work and school experiences over a long period of time. For example, both lecturers commented on their own experiences of learning statistics at school or of using statistics in their own work.

Through teaching statistics over several cycles of learning, lecturers' expertise in teaching and in statistics develops. The findings in my study capture one instance in the ongoing progression of these lecturers' teaching. The macro-micro analysis also highlights the relationships and connections between internal, inner planes of lecturers' intentions, verbal thoughts and speech in the lecture theatre, with its possible interpretations. The challenge for teaching and learning of statistics include identifying the tools which enable lecturers to reinterpret statistics for the students as meaningful knowledge rather than dry procedures and techniques. Through meaningful representations of statistics, lecturers might be able to change symbolic and prototypical exercises to a tool for use in authentic contexts outside of university.

6.6 Reflections on the quality of the study

In this section I turn to the question of the quality of my model of teaching and of my study. In Section 4.1, I outline the criteria for good quality research. In this section, I discuss various aspects of the quality of my study drawing on the points made in the previous chapter. Charmaz (2014) considers that good quality research is interesting, logical and makes sense to the reader. In reflecting on the quality of my study, I considered the justifications for my research questions, my methodology, the research process and the plausibility of my findings.

Before selecting a topic for my main study and deciding on the research questions, I carried out three different pilot studies. These pilot studies were an opportunity for me to review existing literature in statistics education, have conversations with twenty lecturers engaged in the teaching of statistics across disciplines and academic levels and observe lectures and tutorials. The literature review of professional, didactic and research studies revealed a keen interest in researching the teaching practices of inferential statistics at all levels of education and especially at university introductory level. Inferential statistics is a topic which is of interest in a large number of disciplines that use statistics in their programmes of study. However, there are no observational studies that I know of which were concerned with lecturers' teaching practices. My basic unit of analysis in the main study was the teaching of statistics activity and I believe that characterising the teaching of statistics at university can help both teachers of statistics and researchers in statistics education better understand teaching practices in context and the concern for how teaching statistics is interwoven with the "context" of the activity.

My methodological approach was guided by an interpretative paradigm. Given my focus on lecturers' teaching practice, I could carry out detailed analysis of this observational and interview data using the sociocultural theoretical lens and applying a grounded analytical approach. The sociocultural framework has been used in mathematics and statistics education as a useful tool for interpreting teaching and learning activities. At the same time, sociocultural perspectives on teaching and learning are not prescriptive of particular teaching approaches. As I showed in my account of the theoretical perspectives used in this study, sociocultural lens provides useful tools for characterising and analysing teaching activities.

The detailed account of my research design process contributes to the credibility or plausibility of my interpretations. Using an overlap of several data sources (pilot interviews, prolonged observations, interviews with lecturers) lead to the collection of thick descriptive data which lead to various layers of analysis. Throughout the analysis, I made my analytical approach apparent and explicitly connected my research paradigm, theoretical perspectives, research design and data analysis. In my analysis, I provide detailed examples of data and interpretations, using existing research literature for detailed interpretations. This lead to rich descriptions and a detailed account on multiple levels of analysis.

Regarding my categorisations using the grounded analytical approach, I compared and contrasted my findings throughout the study to existing research and I provided a reflective account of my findings. The literature review also helped with defining and using the technical terminology. Throughout this analytical process, I asked myself critical questions regarding the value or worth of my study to the educational research and teaching communities. Presentations and publications of preliminary findings meant that I had the opportunity to discuss and receive constructive criticisms which lead to further improvements of my research study (Harth, 2016, 2017).

Finally, the output of my analysis, a model of teaching statistics in context, can be evaluated using quality criteria. For example, Schoenfeld (1998) proposed quality criteria such as how well a model describes or explains what the lecturers were doing, whether it can predict other lecturers' teaching practices in a closely related context and whether it could capture the range of teaching styles and experiences. The model emerged as part of a lengthy research process in which different sources of data were collected and analysed. Although the analysis might not represent the full complexity of teaching actions accurately, the multiple pilot studies carried out before the main study and the prolonged observations in the main study were intended to give confidence in the analytic process. It could be argued that I included only two lecturers in studying teaching styles. Investigations of teaching practice, not completely specific to the context in which they are applied, opt for depth of observation and analysis (Weber, 2004).

6.7 Conclusions and future research

In the process of conducting this research study and writing my findings, several new areas of enquiry deserve further attention in the teaching of statistics at university. For example, I considered that both levels of analysis, macro and micro and the interlinks between them could be enhanced with carrying out further research.

For example, Vygotsky did not perceive the teacher and the child as a separate reality and was interested in the social environment of the child (Davydov, 2010). In order to provide more depth to this analysis of teaching activity, the expanded triangular representation of activity systems (Figure 5.4) could be used in future research to enable an examination of multiple systems of activity at the macro level of the community and micro levels concentrating on lecturers and students operating with tools such as curricula and teaching resources (Barab et al., 2003).

Such a longitudinal study would involve additional datasets to include student activity inside and outside of the lecture theatre. The aim of such as study would be to represent the social elements of the teaching and learning activity systems with the addition of community of students, rules (within and outside the lecture theatre) and division of labour (Engeström, 1999). The expanded mediational triangle also encourages the analysis of interactions between factors and emphasises the relationships between the subject and the community, an aspect which I did not capture here as it was beyond the scope of my present study. For instance, the use of teaching assistants by module teams is fundamental to students' learning. In this analysis, I did not include the teaching assistants since the focus of my enquiry was the lecturers (Green, 2010; Justice et al., 2017).

Longitudinal data could also be used to capture relationships and connections between the study's model components (Figure 5.13). A future study could for instance investigate the impact on learning of different lecture timings and transitions, different lecture types, curricular delivery and cycles of context as modelled in this study. The lecturers in my study were concerned with the relevance of curricula planned and implemented in their modules. A future study could for example investigate how students are able to transfer statistical skills and knowledge learnt in one context (e.g. introductory statistics module) to another (e.g. a third year project or industrial placement).

The micro-teaching routines are a rich area which in my view require further investigation of both verbal and non-verbal routines. In my study I did not analyse in detail the non-verbal routines such as the arrival of the students, seating or students' reactions. Although I made notes about some of these exchanges, they were not a focus of investigation. However, school-based research has shown how non-verbal routines affect the teaching process since they give purpose and meaning to the teaching activity (Pollard, 2014). A future study could look to use video data of lectures to examine non-verbal routine processes in more detail at the university level.

Amongst verbal routines, discourse markers have been considered to be fundamental for understanding however less is known about their role in the spoken discourse of technical subjects such as statistics (Christodoulidou, 2011). Another area which I think requires further attention is the use of questioning techniques to include formative assessment in statistics lectures (Brophy and Hahn, 2014; Larsen, 2017).

Of particular interest in my study were the links lecturers were making in explaining statistical concepts or non-specific information. Previous studies of school- or university-based teaching agree that teachers' skill in explaining and reasoning in lectures is critical to students' learning (Bligh, 1972). However, there have been very few experimental studies of this aspect of micro-teaching of statistics. Further investigations are required to uncover the relationships between teaching and the development of statistical reasoning at university.

I believe and hope that my study has contributed to some extent towards a (re)conceptualisation of developmental teaching of statistics at university for the 21st century and has raised issues with how to tailor, differentiate and scaffold the teaching to meet the needs of very diverse students, curricula and future work contexts.

References

Abbott, R. D. and Falstrom, P. (1977). Frequent testing and personalized systems of instruction. *Contemporary Educational Psychology*, 2(3):251–257.

Abelson, R. P. (1995). *Statistics as Principled Argument*. Psychology Press, New York.

Aberson, C. L., Berger, D. E., Healy, M. R., and Romero, V. L. (2003). Evaluation of an Interactive Tutorial for Teaching Hypothesis Testing Concepts. *Computers in Teaching*, 30(1):75–78.

Abrahamson, D. (2009). Coordinating phenomenologically immediate and semiotically mediated constructions of statistical distribution. In Gould, R., editor, *The role of context and evidence in informal inferential reasoning. Proceedings of the Sixth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-6). Brisbane, Australia: The University of Queensland.*

ACME (2011). Mathematical Needs: Mathematics in the Workplace and Higher Education. London, UK. Advisory Committee on Mathematics Education.

Afifi, A. A. and Clark, V. A. (1996). *Computer-Aided Multivariate Analysis*. Texts in Statistical Science. Chapman & Hall, London, third edition edition.

Akkerman, S. F. and Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2):132–169.

Allan, J. (1996). Learning outcomes in higher education. *Studies in Higher Education*, 21(1):93–108.

Allen, R. A., Folkhard, A., and Abram, B. (2010). Statistics for the biological and environmental sciences: improving service teaching for postgraduates. In Reading, C. E., editor, *Data and context in statistics education: towards an evidence-based society*. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8), Ljubljana, Slovenia.

Au, K. H. (1990). Changes in a teacher's views of interactive comprehension instruction. In Moll, L. C., editor, *Vygotsky and Education Instructional Implications and Applications of Sociohistorical Psychology*, pages 271–286. Cambridge University Press, Cambridge.

Bakhtin, M. M. (1986). The p roblem of speech genres. In Emerson, C. and Holquist, M., editors, *Speech Genres and Other Late Essays*. University of Texas Press Slavic series, Austin, Texas, first edition.

Bakhtin, M. M. (2007). Vygotsky's demons. In Daniels, H., Cole, M., and Wertsch, J. V., editors, *The Cambridge companion to Vygotsky*, pages 49–75. Cambridge University Press, New York.

Bakhurst, D. (1997). Meaning, normativity and the life of the mind. *Language & Communication*, 17(1):33–51.

Bakhurst, D. (1998). Introduction to Vygotsky. In Daniels, H., editor, *An Introduction to Vygotsky*, pages 196–218. Routledge.

Bakker, A. (2004). *Design Research in Statistics Education on Symbolizing and Computer Tools*. PhD thesis, Center for Science and Mathematics Education, University of Utrecht.

Bakker, A. and Akkerman, S. F. (2013). A boundary-crossing approach to support students' integration of statistical and work-related knowledge. *Educational Studies in Mathematics*, 86(2):223–237.

Bakker, A. and Derry, J. (2011). Lessons from Inferentialism for Statistics Education. *Mathematical Thinking and Learning*, 13(1-2):5–26.

Bakker, A., Kent, P., Derry, J., Noss, R., and Hoyles, C. (2008). Statistical inference at work: Statistical process control as an example. *Statistics Education Research Journal*, 7(2):130–145.

Ball, D. L., Thames, M. H., and Phelps, G. (2008). Content Knowledge for Teaching What Makes It Special? *Journal of Teacher Education*, 59(5):389–407.

Barab, S. A., Barnett, M., Yamagata-Lynch, L., Squire, K., and Keating, T. (2002). Using Activity Theory to Understand the Systemic Tensions Characterizing a Technology-Rich Introductory Astronomy Course. *Mind*, 9(2):76–107.

Barab, S. A., Evans, M. A., and Baek, E.-O. (2003). Activity Theory As a Lens for Characterizing the Participatory Unit. In Jonassen, D. and Driscoll, M., editors, *Handbook of Research on Educational Communications and Technology*, pages 199–214. Routledge, New York.

Barron, K. E. and Apple, K. J. (2014). Debating Curricular Strategies for Teaching Statistics and Research Methods What Does the Current Evidence Suggest? *Teaching of Psychology*, 41(3):187–194.

Batanero, C., Garfield, J. B., Ottaviani, M. G., and Truran, J. (2000). Research in statistical education: Some priority questions. *IASE Newsletter*, 1(2):2–6.

Batanero, C., Tauber, L. M., and Sanchez, V. (2004). Students' reasoning about the normal distribution. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 257–276. Springer Netherlands, Dordrecht.

Baumer, B. (2015). A Data Science Course for Undergraduates: Thinking With Data. *The American Statistician*, 69(4):334–342.

Begg, A. and Edwards, R. (1999). Teachers' Ideas about Teaching Statistics. In *Paper presented at the Australian Association for Research in Education (AARE) Annual Conference*, Melbourne.

Beins, B. C. (1993). Writing Assignments in Statistics Classes Encourage Students to Learn Interpretation. *Teaching of Psychology*, 20(3):161–164.

Bell, G. (2004). The repertory grid technique. In Fransella, F., editor, *International Handbook of Personal Construct Psychology*, pages 95–120. John Wiley & Sons Ltd, Hoboken, NJ.

Bellack, A. A., Kliebard, H. M., Hyman, R. T., and Smith, F. L. (1966). *The Language of the Classroom*. Teachers College Press, New York.

Ben-Zvi, D. (2000). Toward Understanding the Role of Technological Tools in Statistical Learning. *Mathematical Thinking and Learning*, 2(1-2):127–155.

Ben-Zvi, D. (2004). Reasoning about data analysis. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 121–146. Springer Netherlands, New York.

Ben-Zvi, D. and Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, 45(1/3):35–65.

Ben-Zvi, D., Aridor, K., Makar, K., and Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM*, 44(7):913–925.

Ben-Zvi, D. and Aridor-Berger, K. (2016). Children's Wonder How to Wander Between Data and Context. In Ben-Zvi, D. and Makar, K., editors, *The Teaching and Learning of Statistics*, pages 25–36. Springer International Publishing Switzerland.

Ben-Zvi, D. and Garfield, J. B. (2004). Statistical Literacy, Reasoning, and Thinking: Goals, Definitions, and Challenges. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Springer Netherlands, Dordrecht.

Ben-Zvi, D. and Garfield, J. B. (2008). Introducing the Emerging Discipline of Statistics Education. *School Science and Mathematics*, 108(8):355–361.

Ben-Zvi, D. and Makar, K. (2012). The Teaching and Learning of Statistics. In Cho, S. J., editor, *The Proceedings of the th International Congress on Mathematical Education*, page 334, Seoul, Korea. Springer.

Bernstein, B. (1999). Vertical and Horizontal Discourse: An essay. *British Journal of Sociology of Education*, 20(2):157–173.

Bernstein, B. (2000). *Pedagogy, Symbolic Control, and Identity.* Theory, Research, Critique. Rowan & Littlefield Publishers Inc, Lanham, revised edition.

Bessant, K. C. (1992). Instructional Design and the Development of Statistical Literacy. *Teaching Sociology*, 20(2):143.

Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory & Cognition*, 10(6):511–519.

Beyth-Marom, R. and Dekel, S. (1983). A curriculum to improve thinking under uncertainty. *Instr Sci*, 12(1):67–82.

Beyth-Marom, R. and Fidler, F. (2008). Statistical cognition: Towards evidence-based practice in statistics and statistics education. *Statistics Education Research Journal*, 7(2):20–39.

Biggs, J. B. (2003). *Teaching for Quality Learning at University.* Open University Press, Buckingham, UK, second edition.

Biggs, J. B. and Collis, K. F. (1982). *Evaluating the Quality of Learning.* The SOLO Taxonomy (Structure of the Observed Learning Outcome). Academic Press, New York.

Biggs, J. B. and Tang, C. (2011). *Teaching for Quality Learning at University: What the student does.* Society for Research into Higher Education & Open University Press, fourth edition edition.

Bjorklund, L. (2008). The Repertory Grid Technique: Making Tacit Knowledge Explicit: Assessing Creative work and Problem solving skills . In Middleton, H., editor, *Researching Technology Education: Methods and techniques.* Sense Publishers, Rotterdam.

Blades, N. J., Schaalje, G. B., and Christensen, W. F. (2015). The Second Course in Statistics: Design and Analysis of Experiments? *The American Statistician*, 69(4):326–333.

Blanton, M. L., Westbrook, S., and Carter, G. (2005). Using Valsiner's zone theory to interpret teaching practices in mathematics and science classrooms. *Journal of Mathematics Teacher Education*, 8(1):5–33.

Blanton, W. E., Simmons, E., and Warner, M. (2016). The Fifth Dimension: Application of Cultural-Historical Activity Theory, Inquiry-Based Learning, Computers, and Telecommunications to Change Prospective Teachers' Preconceptions:. *Journal of Educational Computing Research*, 24(4):435–463.

Bligh, D. A. (1972). *What's the use of lectures?* Harmondsworth, Penguin.

Bloom, B. S. (1953). Thought-processes in lectures and discussions. *The Journal of General Education*, 7(3):160–169.

Bloom, B. S. (1956). *Taxonomy of educational objectives: the classification of educational goals*. Addison-Wesley Longman Ltd., Longman, New York, 2nd edition.

Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more "real". *For the learning of mathematics*, 13(2):12–17.

Box, G. E. P. and Draper, N. R. (1986). *Empirical Model Building and Response Surfaces*. Wiley-Blackwell, Hoboken, NJ, USA.

Bradstreet, T. E. (1996). Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning. *The American Statistician*, 50(1):69–78.

Broers, N. J. (2006). Learning goals: The primacy of statistical knowledge. In Rossman, A. J. and Chance, B. L., editors, *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)*, Salvador, Bahia, Brazil.

Brophy, C. and Hahn, L. (2014). Engaging Students in a Large Lecture: An Experiment using Sudoku Puzzles. *Journal of Statistics Education*, 22(1).

Brown, G. A., Bakhtar, M., and Youngman, M. B. (1984). Toward a typology of lecturing styles. *British Journal of Educational Psychology*, 54(1):93–100.

Brown, J. S., Collins, A., and Duguid, P. (1989). Situated Cognition and the Culture of Learning:. *Educational Researcher*, 18(1):32–42.

Bryant, A. and Charmaz, K. (2007). *The SAGE Handbook of Grounded Theory*. SAGE Publications, London.

Bryce, G. B. (2002). Undergraduate statistics education: An introduction and review of selected literature. *Journal of Statistics Education*, 10(2).

Budé, L., Van De Wiel, M., Imbos, T., and Candel, M. (2007). Students' achievements in a statistics course in relation to motivational aspects and study behaviour. *Statistics Education Research Journal*, 6(1):5–21.

Burgess, R. G. (1985). The whole truth? Some ethical problems of research in a comprehensive school. In Burgess, R. G., editor, *Field methods in the study of education*, pages 139–162. Falmer Press, Lewes.

Burgess, T. (2008). Teacher knowledge for teaching statistics through investigations. In Batanero, C., Burrill, G., Reading, C. E., and Rossman, A. J., editors, *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.*

Burton, L. (2002). Methodology and Methods in Mathematics Education Research: Where is the Why. In Goodchild, S. and English, L. D., editors, *Researching mathematics classrooms A critical examination of methodology*, pages 1–10. Praeger, London.

Calderhead, J. (2004). Teachers: beliefs and knowledge. In Berliner, D. C. and Calfee, R. C., editors, *Handbook of Educational Psychology*, pages 709–725. Routledge, New York and London.

Callingham, R., Carmichael, C., and Watson, J. M. (2015). Explaining Student Achievement: the Influence of Teachers' Pedagogical Content Knowledge in Statistics. *International Journal of Science and Mathematics Education*, pages 1–19.

Callingham, R. and Watson, J. M. (2011). Measuring Levels of Statistical Pedagogical Content Knowledge. In *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, pages 283–293. Springer Netherlands, Dordrecht.

Cardona, L. Z. (2008). How Do Teachers Deal with the Heuristic of Representativeness? In *Paper presented at the International Congress on Mathematics Education (ICME) 11*, Monterrey, Mexico.

Champkin, J. (2014). Timeline of statistics. *Significance*, 10:23–26.

Chance, B. L. (1997). Experiences with Authentic Assessment Techniques in an Introductory Statistics Course. *Journal of Statistics Education*, 5(3):1–15.

Chance, B. L. (2002). Components of Statistical Thinking and Implications for Instruction and Assessment. *Journal of Statistics Education*, 10(3):1–15.

Chance, B. L., Ben-Zvi, D., Garfield, J. B., and Medina, E. (2007). The Role of Technology in Improving Student Learning of Statistics. *Technology Innovations in Statistics Education*, 1(1).

Charmaz, K. (1990). Discovering chronic illness: Using grounded theory. *Social Science and Medicine*, 30(11):1161–1172.

Charmaz, K. (2006). *Constructing Grounded Theory.* A practical guide through qualitative analysis. Sage Publications Ltd.

Charmaz, K. (2014). *Constructing Grounded Theory (Introducing Qualitative Methods series).* Sage Publications Ltd, second edition.

Chatfield, C. (1988). *Problem Solving: A Statistician's Guide*. Texts in Statistical Science. Chapman & Hall, London.

Chervany, N. L., Collier, R. O., and Fienberg, S. E. (1977). A framework for the development of measurement instruments for evaluating the introductory statistics course. *The American Statistician*, 31(1):17–23.

Chick, H. L. and Pierce, R. U. (2008). Teaching statistics at the primary school level: Beliefs, affordances, and pedagogical content knowledge. In Batanero, C., Burrill, G., Reading, C. E., and Rossman, A. J., editors, *Joint ICMI/IASE Study Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18*.

Chilisa, B. and Kawulich, B. (2012). Selecting a research approach: Paradigm, methodology and methods. In Wagner, C., Kawulich, B., and Garner, M., editors, *Doing Social Research: A global context*. McGraw-Hill Higher Education, London.

Chong, H. Y., Sandra, A., David, W., Angel, J.-P., and Samuel, A. D. (2012). Teaching Factor Analysis in Terms of Variable Space and Subject Space Using Multimedia Visualization. *Journal of Statistics Education*, 10(1):620.

Christodoulidou, M. (2011). Lexical Markers Within the University Lecture. *Novitas-Royal (Research on Youth and Language)*, 5(1):143–160.

Clark, C. M. and Penelope, P. L. (1984). Teachers' Thought Processes. Occasional Paper No. 72. Technical report, Institute for Research on Teaching, College Education, Michigan State University, East Lansing, MI.

Cobb, G. W. (1992). Teaching statistics. In Steen, L. A., editor, *Heeding the call for change: Suggestions for Curricular Action*, pages 3–43. The Mathematical Association of America.

Cobb, G. W. (1993). Reconsidering Statistics Education: A National Science Foundation Conference. *Journal of Statistics Education*, 1(1).

Cobb, G. W. (2000). Teaching statistics: more data, less lecturing. In Moore, T. L., editor, *Teaching statistics Resources for undergraduate instructors*. Mathematics Association of America.

Cobb, G. W. (2005). Introductory Statistics: A Saber Tooth Curriculum? In *After dinner talk given at the United States Conference on Teaching Statistics, USCOTS*.

Cobb, G. W. (2007). The introductory statistics course: a Ptolemaic curriculum? *Technology Innovations in Statistics Education*, 1(1).

Cobb, G. W. (2015). Mere Renovation is Too Little Too Late: We Need to Rethink our Undergraduate Curriculum from the Ground Up. *The American Statistician*, 69(4):266–282.

Cobb, G. W. and Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, 104(9):801–823.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning*, 1(1):5–43.

Cobb, P., Perlwitz, M., and Underwood, D. (1996). An approach to arithmetic computation compatible with constructivism. In Mansfield, H., Pateman, N. A., and Bednarz, N., editors, *Mathematics for Tomorrows Young Children: International perspectives on curriculum*, pages 31–46. Kluwer Academic Publishers, Montréal, Québec.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatley, G., Trigatti, B., and Perlwitz, M. (1991). Assessment of a Problem-Centered Second-Grade Mathematics Project. *Journal for Research in Mathematics Education*, 22(1):3–39.

Cole, M. (1985). The zone of proximal development: where culture and cognition create each other. In Wertsch, J. V., editor, *Culture, Communication and Cognition: Vygotskian Perspectives*, pages 146–161. Cambridge University Press, Cambridge.

Cole, M. (1996). *Cultural Psychology: A Once and Future Discipline*. Harvard University Press, Cambridge.

Cole, M. and Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In Salomon, G., editor, *Distributed cognitions Psychological and educational considerations*, pages 1–46. Cambridge University Press, Cambridge.

Cole, M. and Gajdamaschko, N. (2007). Vygotsky and Culture. In Daniels, H., Cole, M., and Wertsch, J. V., editors, *The Cambridge companion to Vygotsky*, pages 193–210. Cambridge University Press, New York.

Cole, M. and Gajdamaschko, N. (2010). Vygotsky and context: Toward a resolution of theoretical disputes. In Kirschner, S. R. and Martin, J., editors, *The sociocultural turn in psychology The contextual emergence of mind and self*, pages 253–280. Columbia University Press, New York, NY, US.

Confrey, J. (1995). A Theory of Intellectual Development. *For the Learning of Mathematics*, 15(1):38–48.

Connelly, F. M. and Clandinin, D. J. (1990). Stories of Experience and Narrative Inquiry. *Educational Researcher*, 19(5):2–14.

Corbin, J. and Morse, J. M. (2016). The Unstructured Interactive Interview: Issues of Reciprocity and Risks when Dealing with Sensitive Topics:. *Qualitative Inquiry*, 9(3):335–354.

Corbin, J. and Strauss, A. L. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publishing, Inc, Thousand Oaks, 3rd edition.

Corbin, J. M. and Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1):3–21.

Corbin, J. M. and Strauss, A. L. (1998). The Conditional/Consequential Matrix. In *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, pages 181–199. Sage Publishing, Inc, Thousand Oaks.

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Sage Publications, Inc., Thousand Oaks, 2nd edition.

Cuban, L. (1984). *How Teachers Taught: Constancy and Change in American Classrooms, 1890-1980*. Longman, New York, 1st edition.

Curtis, A. M., Wells, T. M., Higbee, T., and Lowry, P. B. (2008). An overview and tutorial of the repertory grid technique in information systems research. *Communications of the Association for Information Systems (CAIS)*, 23(3):37–62.

Daniels, H. (2001). *Vygotsky and Pedagogy*. Routledge Falmer, London.

Daniels, H. (2015). Boundaries Within and Between Contexts. In Marsico, G., Dazzani, V., Ristum, M., and de Souza Bastos, A. C., editors, *Educational Contexts and Borders through a Cultural Lens*, pages 11–27. Springer International Publishing, Cham.

Davies, N., Marriott, J., and Martignetti, D. (2012). A Statistical Awareness Curriculum for STEM Employees. *MSOR Connections*, 12(2):12–16.

Davydov, V. V. (2010). The Concept of Developmental Teaching. *Journal of Russian and East European Psychology*, 36(4):11–36.

Davydov, V. V. and Markova, A. K. (1982). A Concept of Educational Activity for Schoolchildren. *Soviet Psychology*, 21(2):50–76.

Davydov, V. V. and Radzikhovskii, L. A. (1985). Vygotsky's theory and the activity-oriented approach in psychology. In Wertsch, J. V., editor, *Culture, Communication and Cognition: Vygotskian Perspectives*, pages 35–65. Cambridge University Press, Cambridge.

Delamont, S. and Hamilton, D. (1993). Revisiting classroom research: a continuing cautionary tale. In Hammersley, M., editor, *Controversies in classroom teaching*, pages 25–43. Open University Press, Buckingham.

DelMas, R. C. (2004). A Comparison of Mathematical and Statistical Reasoning . In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 79–96. Springer Netherlands, Dordrecht.

DelMas, R. C., Garfield, J. B., and Chance, B. L. (1999). A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning. *Journal of Statistics Education*, 7(3).

DelMas, R. C., Garfield, J. B., Ooms, A., and Chance, B. L. (2007). Assessing students' conceptual understanding of statistical relationships. *Statistics Education Research Journal*, 6(2):28–58.

Derry, J. (2007). Abstract rationality in education: from Vygotsky to Brandom. *Studies in Philosophy and Education*, 27(1):49–62.

Derry, J. (2013). *Vygotsky, Philosophy and Education*. Wiley Blackwell, Chichester.

Derry, S., Levin, J. R., and Schauble, L. (1995). Stimulating Statistical Thinking through Situated Simulations. *Teaching of Psychology*, 22(1):51–57.

Derry, S. J., Levin, J. R., Osana, H. P., Jones, M. S., and Peterson, M. (2000). Fostering Students' Statistical and Scientific Thinking: Lessons Learned From an Innovative College Course. *American Educational Research Journal*, 37(3):747–773.

Dewey, J. (2011). *Democracy and Education: An Introduction to the Philosophy of Education by Dewey*. Simon and Brown, Milton Keynes.

Dierdorp, A., Bakker, A., Eijkelhof, H., and van Maanen, J. (2011). Authentic Practices as Contexts for Learning to Draw Inferences Beyond Correlated Data. *Mathematical Thinking and Learning*, 13(1-2):132–151.

Dixon, P. N. and Judd, W. A. (1977). A Comparison of Computer-Managed Instruction and Lecture Mode for Teaching Basic Statistics. *Journal of Computer-Based Instruction*, 4(1):22–25.

Dwyer, C. A., Gallagher, A., Levin, J., and Morley, M. E. (2003). What is Quantitative Reasoning? *ETS Research Report Series*, (RR-03-30).

Eckstein, M. A., Travers, K. J., and Shafer, S. M. (1982). A Comparative Review of Curriculum: Mathematics and International Studies in the Secondary Schools of Five Countries. Technical report, National Commission on Excellence in Education, Washington, DC.

Efron, B. and Hastie, T. (2016). *Computer Age Statistical Inference*. Cambridge University Press, Cambridge.

Eichler, A. (2006). Individual curricula-beliefs behind teachers' beliefs. In Rossman, A. J. and Chance, B. L., editors, *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)*, Salvador, Bahia, Brazil.

Eichler, A. (2007). Individual curricula: teachers' beliefs concerning stochastics instruction. *International Electronic Journal of Mathematics Education (IEJME)*, 2(3):208–226.

Eichler, A. (2008a). Statistics teaching in German secondary high schools. In Batanero, C., Burrill, G., Reading, C. E., and Rossman, A. J., editors, *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, Monterrey, Mexico.

Eichler, A. (2008b). Teachers' classroom practice and students' learning. In Batanero, C., Burrill, G., Reading, C. E., and Rossman, A. J., editors, *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, Monterrey, Mexico.

Eichler, A. (2009). The role of context in stochastics instruction. In Durand-Guerrier, V., Soury-Lavergne, S., and Arzarello, F., editors, *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education*, Lyon, France.

Eichler, A. (2011). Statistics Teachers and Classroom Practices. In Batanero, C., Burrill, G., and Reading, C. E., editors, *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, pages 175–186. Springer, Dordrecht.

Elliott, G. and Greatorex, J. (2002). A fair comparison? The evolution of methods of comparability in national assessment. *Educational Studies*, 28(3):253–264.

Elton, L. and Johnston, B. (2002). Assessment in Universities: a critical review of research. *Learning and Teaching Support Network Generic Centre.*

Engeström, Y. (1987). *Learning by Expanding: An Activity-theoretical Approach to Developmental Research.* Orienta-Konsultit, Helsinki.

Engeström, Y. (1995). Objects, contradictions and collaboration in medical cognition: an activity-theoretical perspective. *Artificial intelligence in medicine*, 7(5):395–412.

Engeström, Y. (1996). Developmental studies of work as a testbench of activity theory: the case of primary care medical practice. In Chaiklin, S. and Lave, J., editors, *Understanding practice perspectives on activity and context*, pages 64–103. Cambridge University Press, Cambridge.

Engeström, Y. (1999). Activity theory and individual and social tranformation. In Engeström, Y., Miettinen, R., and Punamaki, R.-L., editors, *Perspectives on Activity Theory*, page 480. Cambridge University Press, Cambridge, UK.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, 43(7):960–974.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1):133–156.

Engeström, Y. and Middleton, D. (1998). Introduction: Studying work as mindful practice. In Engeström, Y. and Middleton, D., editors, *Cognition and Communication at Work*, pages 1–14. Cambridge University Press, Cambridge.

English, L. D. (2007). Cognitive Psychology and Mathematics Education: Reflections on the Past and the Future. In Sriraman, B., editor, *Zoltan Paul Dienes and the Dynamics of Mathematical Learning*, pages 119–126. Lawrence Erlbaum Associates, Mahwah, Charlotte, NC, USA.

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, 70(1):113–136.

Ericsson, K. A. and Simon, H. A. (1998). How to Study Thinking in Everyday Life: Contrasting Think-Aloud Protocols With Descriptions and Explanations of Thinking. *Mind*, 5(3):178–186.

Foley, J. (2009). Vygotsky, bernstein and halliday: Towards a unified theory of L1 and L2 learning. *Language, Culture and Curriculum*, 4(1):17–42.

Fong, G. T., Krantz, D. H., and Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18(3):253–292.

Fong, G. T. and Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, 120(1):34–45.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., and Scheaffer, R. (2007). Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: a Pre-K-12 Curriculum Framework. Technical report.

Fransella, F., Bell, R. C., and Bannister, D. (2004). *A Manual for Repertory Grid Technique*. John Wiley & Sons, Ltd, Chichester, West Sussex, England, 2nd edition.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415.

Friedrich, J., Buday, E., and Kerr, D. (2000). Statistical Training in Psychology: A National Survey and Commentary on Undergraduate Programs. *Teaching of Psychology*, 27(4):248–257.

Gage, N. (2009). *A Conception of Teaching*. Springer Science + Business Media, LLC, New York.

Gage, N. L. (1978). *The Scientific Basis of the Art of Teaching*. Teachers' College Press, Columbia University, New York.

GAISE (2005). Guidelines for assessment and instruction in statistics education (GAISE) College Report. Technical report, Alexandria, VA.

GAISE (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report. Technical report, Alexandria, VA.

Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review*, 70(1):1–51.

Gal, I. (2004). Statistical Literacy. In Ben-Zvi, D. and Garfield, J. B., editors, *The challenge of developing statistical literacy, reasoning and thinking*, pages 79–95. Springer International Publishing, Dordrecht.

Gal, I. and Garfield, J. B. (1997). Curricular Goals and Assessment Challenges in Statistics Education. In Gal, I. and Garfield, J. B., editors, *The Assessment Challenge in Statistics Education*, pages 1–13. IOS Press, Amsterdam.

Gal, I. and Ginsburg, L. (1994). The Role of Beliefs and Attitudes in Learning Statistics: Towards an Assessment Framework. *Journal of Statistics Education*, 2(2).

Gal, I., Ginsburg, L., and Schau, C. (1997). Monitoring Attitudes and Beliefs in Statistics Education. In Gal, I. and Garfield, J. B., editors, *The Assessment Challenge in Statistics Education*. IOS Press, Amsterdam.

Gal, l. (2000). Statistical Literacy: Conceptual and Instructional issues. In *Perspectives on Adults Learning Mathematics*, pages 135–150. Springer, Dordrecht, Dordrecht.

Gardner, P. L. and Hudson, I. (1999). University Students' Ability to Apply Statistical Procedures. *Journal of Statistics Education*, 7(1).

Garfield, J. B. (1994). Beyond Testing and Grading: Using Assessment To Improve Student Learning. *Journal of Statistics Education*, 2(1):1–10.

Garfield, J. B. (1995). How Students Learn Statistics. *International Statistical Review*, 63(1):25–34.

Garfield, J. B. (1996). Assessing student learning in the context of evaluating a chance course. *Communications in Statistics–Theory and Methods*, 25(11):2863–2873.

Garfield, J. B. (1998). The statistical reasoning assessment: Development and validation of a research tool. In Pereira-Mendoza, L., Kea, L. S., Kee, T. W., and Wong, W.-K., editors, *The Proceedings of the Fifth International Conference on Teaching Statistics (ICOTS5)*, Singapore.

Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2(1):22–38.

Garfield, J. B. (2017). Teaching Statistics Using Small-Group Cooperative Learning. *Journal of Statistics Education*, 1(1).

Garfield, J. B. and Ahlgren, A. (1988). Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal for Research in Mathematics Education*, 19(1):44.

Garfield, J. B. and Ben-Zvi, D. (2004). Research on Statistical Literacy, Reasoning, and Thinking: Issues, Challenges, and Implications. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 397–410. Springer Netherlands, Dordrecht.

Garfield, J. B. and Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3):372–396.

Garfield, J. B. and Ben-Zvi, D. (2008). Developing Students' Statistical Reasoning. In *Developing students' statistical reasoning: Connecting research and teaching practice*, pages 25–55. Springfield.

Garfield, J. B. and Chance, B. L. (2000). Assessment in statistics education: Issues and challenges. *Mathematical Thinking and Learning*, 2(1-2):99–125.

Garfield, J. B., Chance, B. L., and Snell, L. S. (2002). Technology in College Statistics Courses. In Holton, D., Artigue, M., Kirchgraber, U., Hillel, J., Niss, M., and Schoenfeld, A., editors, *The Teaching and Learning of Mathematics at University Level*, pages 357–370. Kluwer Academic Publishers, New York.

Garfield, J. B., delMas, R. C., and Chance, B. L. (2012a). Using Students' Informal Notions of Variability to Develop an Understanding of Formal Measures of Variability. In *Thinking With Data*, pages 132–162. Psychology Press.

Garfield, J. B., DelMas, R. C., and Zieffler, A. (2012b). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM*, 44(7):883–898.

Garfield, J. B. and Gal, I. (1999). Teaching and assessing statistical reasoning. In Stiff, L., editor, *Developing Mathematical Reasoning in Grades K-12*, pages 207–219. National Council Teachers of Mathematics 1999 Year.

Garfield, J. B., Le, L., Zieffler, A., and Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics*, 88:327–342.

Garfield, J. B., Zieffler, A., Kaplan, D., Cobb, G. W., Chance, B. L., and Holcomb, J. P. (2011). Rethinking assessment of student learning in statistics courses. *The American Statistician*, 65(1):1–10.

Gentner, D. and Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist*, 52(1):32–34.

Gentner, D. and Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65(2-3):263–297.

Gibbs, G. and Dunbar-Goddet, H. (2007). The effects of programme assessment environments on student learning. *Higher Education Academy.*

Gil, E. and Ben-Zvi, D. (2011). Explanations and Context in the Emergence of Students' Informal Inferential Reasoning. *Mathematical Thinking and Learning*, 13(1-2):87–108.

Glaser, B. G. and Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Aldine Publishing Company, Chicago.

Glynn, S. M. (1974). The teaching with analogies model. In Muth, D. K., editor, *Childrens comprehension of text: Research into Practice*, pages 185–204. International Reading Association, Newark, Delaware.

Gold, R. L. (1958). Roles in Sociological Field Observations. *Social Forces*, 36:217–223.

González, O. (2011). Secondary School Teachers' Statistical Knowledge for Teaching and Espoused Beliefs on Teaching and Learning of Variability-Related Concepts. In Ubuz, B., editor, *Proceedings of the th Conference of the 35th International Group for the Psychology of Mathematics Education*, Ankara, Turkey.

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3):606–633.

Goos, M. (2005). A sociocultural analysis of the development of pre-service and beginning teachers' pedagogical identities as users of technology. *Journal of Mathematics Teacher Education*, 8(1):35–59.

Gordon, F. S. and Gordon, S. P. (1989). Computer graphics simulations of sampling distributions. *Collegiate Microcomputing*, 7(2):185–189.

Gordon, S. (1993). Mature students learning statistics: The activity theory perspective. *Mathematics Education Research Journal*, 5(1):34–49.

Gordon, S. (1995). A theoretical approach to understanding learners of statistics. *Journal of Statistics Education*, 3(3):1–21.

Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, 3(1):40–59.

Gordon, S. and Fittler, K. (2004). Learning By Teaching: A Cultural Historical Perspective On A Teacher's Development. *Outlines. Critical Practice Studies*, 6(2):35–46.

Gordon, S., Petocz, P., and Reid, A. (2007). Teachers' Conceptions of Teaching Service Statistics Courses. *International Journal for the Scholarship of Teaching and Learning*, 1(1).

Gould, R. (2010). Statistics and the Modern Student . *International Statistical Review*, 78(2):297–315.

Greatorex, J. (2002). Making accounting examiners' tacit knowledge more explicit: developing grade descriptors for an Accounting A-level. *Research Papers in Education*, 17(2):211–226.

Green, J. L. (2010). Teaching highs and lows: Exploring university teaching assistants' experiences. *Statistics education Research Journal*, 9(2):108–122.

Greenfield, P. (1984). Theory of the teacher in learning activities. In Rogoff, B. and Lave, J., editors, *Everyday cognition its development in social context*, pages 117–138. Harvard University Press, Cambridge, Massachusetts.

Greeno, J. G. (1991). Number Sense as Situated Knowing in a Conceptual Domain. *Journal for Research in Mathematics Education*, 22(3):170.

Greeno, J. G. (1997). Response: On Claims That Answer the Wrong Questions. *Educational Researcher*, 26(1):5–17.

Grice, J. W. (2002). Idiogrid: Software for the management and analysis of repertory grids. *Behavior Research Methods*, 34(3):338–341.

Grice, J. W. (2007). Generalized Procrustes Analysis Example with Annotation. Technical report.

Grice, J. W. and Assad, K. K. (2009). Generalized procrustes analysis: a tool for exploring aggregates and persons. *Applied Multivariate Research*, 13(1):93–112.

Grice, J. W., Burkley, E., Burkley, M., Wright, S., and Slaby, J. (2004). A sentence completion task for eliciting personal constructs in specific domains. *Personal Construct Theory and Practice*, 1(2):60–75.

Groth, R. and Meletiou-Mavrotheris, M. (2017). Research on Statistics Teachers' Cognitive and Affective Characteristics. In Ben-Zvi, D., Makar, K., and Garfield, J. B., editors, *International Handbook of Research in Statistics Education*, pages 327–355. Springer International Publishing, Cham.

Groth, R. E. (2017). Developing Statistical Knowledge for Teaching During Design-Based Research. *Statistics Education Research Journal*, 16(2).

Groth, R. E. and Bergner, J. A. (2005). Pre-service elementary school teachers' metaphors for the concept of statistical sample. *Statistics Education Research Journal*, 4(2):27–42.

Groth, R. E. and Bergner, J. A. (2009). Preservice Elementary Teachers' Conceptual and Procedural Knowledge of Mean, Median, and Mode. *Mathematical Thinking and Learning*, 8(1):37–63.

Guba, E. and Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In Denzin, N. K. and Lincoln, Y. S., editors, *Hanbook of Qualitative Research*, pages 105–117. SAGE Publications, Inc, Thousand Oaks, CA.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ*, 29(2):75–91.

Hadegaard, M. (1990). The zone of proximal development as basis for instruction. In Moll, L. C., editor, *Vygotsky and Education Instructional Implications and Applications of Sociohistorical Psychology*, pages 349–371. Cambridge University Press, Cambridge.

Haenen, J., Schrijnemakers, H., and Stufkens, J. (2003). Sociocultural theory and the practice of teaching historical concepts. In Kozulin, A., Gindis, B., Ageyev, V. S., and Miller, S. M., editors, *Vygotsky's Educational Theory in Cultural Context*, pages 246–266. Cambridge University Press, Cambridge.

Hall, S. and Vance, E. A. (2010). Improving self-efficacy in statistics: Role of self-explanation and feedback. *Journal of Statistics Education*, 18(3).

Hammersley, M. (1993a). Introduction. In Hammersley, M., editor, *Controversies in classroom teaching*, pages x–xxii. Open University Press, Buckingham.

Hammersley, M. (1993b). Some reflections upon the macro-micro problem. In Hammersley, M., editor, *Controversies in classroom teaching*, pages 153–163. Open University Press, Buckingham.

Handal, B. and Herrington, A. (2003). Mathematics teachers' beliefs and curriculum reform. *Mathematics Education Research Journal*, 15(1):59–69.

Harry, B., Sturges, K. M., and Klingner, J. K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational Researcher*, 34(2):3–13.

Harth, H. (2016). Curricular goals: lecturers' beliefs concerning in-service undergraduate statistics education. In *Proceedings of the British Society for Research into Learning Mathematics*, volume 36(2).

Harth, H. (2017). Context as action in the teaching of statistical concepts. In Ramirez-Montoya, M.-S., editor, *Handbook of Research on Driving STEM Learning With Educational Technologies*, pages 451–470. IGI Global, Hershey, PA.

Hassad, R. (2011). Constructivist and Behaviorist Approaches: Development and Initial Evaluation of a Teaching Practice Scale for Introductory Statistics at the College Level. *Numeracy*, 4(2).

Hawkins, A. (1997). Discussion: Forward to Basics! A Personal View of Developments in Statistical Education. *International Statistical Review*, 65(3):280–287.

Hiebert, J. and Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In *Second Handbook of Research on Mathematics Teaching and Learning*. Iap.

Higgins, J. J. (1999). Nonmathematical Statistics: A New Direction for the Undergraduate Discipline. *The American Statistician*, 53(1):1–6.

Himmelberger, K. S. and Schwartz, D. L. (2007). It's a Home Run! Using Mathematical Discourse to Support the Learning of Statistics. *The Mathematics Teacher*, 101(4):250–256.

Hirsch, L. S. and O'Donnell, A. M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education*, 9(2).

Hoerl, R. W. (1997). Introductory statistical education: radical redesign is needed, or is it. *Statistical Education Section*, 3(1).

Hogg, R. V. (1990). Statisticians gather to discuss statistical education. *Amstat News*, pages 19–20.

Hogg, R. V. (1991). Statistical Education: Improvements are Badly Needed. *The American Statistician*, 45(4):342–343.

Holcomb, J., Chance, B. L., Rossman, A. J., and Cobb, G. (2010a). Assessing student learning about statistical inference. In Reading, C., editor, *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)*.

Holcomb, J., Chance, B. L., Rossman, A. J., and Tietjen, E. (2010b). Introducing concepts of statistical inference via randomization tests. In Reading, C., editor, *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)*.

Holmes, P. (2003). 50 years of statistics teaching in English schools: some milestones. *The Statistician*, 52(4):439–463.

Holmqvist Olander, M. (2016). Development of students' knowledge about didactics during their first year of the master's programme in didactics. *Reflective Practice*, 17(4):507–521.

Honey, P. (1992). The first step on the ladder - a practical case in identifying competencies. In Boam, R. and Sparrow, P., editors, *Designing and achieving competency*. McGraw-Hill Publishing Co., McGraw-Hill Training Series, London.

Hoogveld, A. W. M., Paas, F., Jochems, W. M. G., and Van Merriënboer, J. J. G. (2002). Exploring teachers' instructional design practices from a systems design perspective - Springer. *Instructional Science*, 30(4):291–305.

Hussey, T. and Smith, P. (2002). The trouble with learning outcomes. *Active Learning in Higher Education*, 3(3):220–233.

Hussey, T. and Smith, P. (2003). The Uses of Learning Outcomes. *Teaching in Higher Education*, 8(3):357–368.

Huxham, M. (2005). Learning in lectures Do 'interactive windows' help? *Active Learning in Higher Education*, 6(1):17–31.

IAEEA (1979). Second Study of Mathematics. Technical Report Bulletin No. 4, Wellington, New Zealand.

Iannone, P. and Nardi, E. (2005). On the pedagogical insight of mathematicians: 'Interaction' and 'transition from the concrete to the abstract'. *The Journal of Mathematical Behavior*, 24(2):191–215.

Inhelder, B. and Piaget, J. (1958). *The Growth of Logical Thinking: From Childhood to Adolescence*. Routledge and Kegan Paul Ltd, London.

Jacob, E. (1997). Context and Cognition: Implications for Educational Innovators and Anthropologists. *Anthropology & Education Quarterly*, 28(1):3–21.

Jaworski, B. (1994). *Investigating mathematics teaching: A constructivist enquiry*. Taylor and Francis, London.

Jaworski, B. (2002). Sensitivity and Challenge in University Mathematics Tutorial Teaching. *Educational Studies in Mathematics*, 51(1/2):71–94.

Jaworski, B. (2008). Building and sustaining inquiry communities in mathematics teaching development: teachers and didacticians in collaboration . In Krainer, K. and Wood, T., editors, *The International Handbook of Mathematics Teacher Education: Participants in Mathematics Teacher Education Volume 3: Individuals, Teams, Communities and Networks*. Sense Publishers, Rotterdam.

Jaworski, B. and Potari, D. (2009). Bridging the macro-and micro-divide: Using an activity theory model to capture sociocultural complexity in mathematics teaching and its development. *Educational Studies in Mathematics*, 72(2):219–236.

Jaworski, B., Treffert-Thomas, S., and Bartsch, T. (2009). *Characterising the teaching of university mathematics: a case of linear algebra*. International Group for the Psychology of Mathematics Education.

Jolliffe, F. R. (1976). A continuous assessment scheme for statistics courses for social scientists. *International Journal of Mathematical Education in Science and Technology*, 7(1):97–103.

Jones, G. A., Lagrall, C. W., Mooney, E. S., and Thornton, C. A. (2004). Models of Development in statistical reasoning. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 79–95. Springer Netherlands, Dordrecht.

Justice, N., Zieffler, A., and Garfield, J. B. (2017). Statistics Graduate Teaching Assistants' Beliefs, Practices and Preparation for Teaching Introductory Statistics. *Statistics Education Research Journal*, 16(1):294–319.

Kahneman, D. (2012). *Thinking, Fast and Slow.* Penguin, London.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454.

Kalton, G. (1973). Problems and possibilities with teaching introductory statistics to social scientists. *International Journal of Mathematical Education in Science and Technology*, 4(1):6–16.

Kandola, R. and Pearn, M. (1992). Identifying competencies. In Boam, R. and Sparrow, P., editors, *Designing and Achieving Competency: A Competency-based Approach to Developing People and Organizations*, pages 31–50. McGraw-Hill Publishing Co., McGraw-Hill Training Series, London.

Kane, M. T. (2006). Validation. In Brennan, R. L., editor, *Educational Measurement*, pages 17–64. NCME.

Kapadia, R. (1980). Developments in statistical education. *Educational Studies in Mathematics*, 11(4):443–461.

Kay, R. H. and LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3):819–827.

Kelly, G. A. (1955). *The Psychology of Personal Constructs*, volume 1. W.W. Norton and Company Inc., New York.

Kempthorne, O. (1980). The Teaching of Statistics: Content versus Form. *The American Statistician*, 34(1):17–21.

Khachatryan, D. and Karst, N. (2017). V for Voice: Strategies for Bolstering Communication Skills in Statistics. *Journal of Statistics Education*, 25(2):68–78.

King, K., Hillel, J., and Artigue, M. (2002). Technology. In Holton, D., Artigue, M., Kirchgraber, U., Hillel, J., Niss, M., and Schoenfeld, A., editors, *The Teaching and Learning of Mathematics at University Level*, pages 349–356. Kluwer Academic Publishers, New York.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Prentice Hall, Englewood Cliffs, NJ.

Konold, C. (1995). Issues in Assessing Conceptual Understanding in Probability and Statistics. *Journal of Statistical Education*, 3(1).

Konold, C., Pollatsek, A., Well, A. D., Lohmeier, J., and Lipson, A. (1993a). Inconsistencies in Students' Reasoning about Probability. *Journal for Research in Mathematics Education*, 24(5):392.

Konold, C., Well, A. D., Lohmeier, J., and Pollatsek, A. (1993b). Understanding the law of large numbers. Fifteenth Annual Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Pacific Grove, California.

Kovecses, Z. (2004). *Metaphor and emotion*. Language, culture, and body in human feeling. Maison des Sciences de l'Homme and Cambridge University Press, Cambridge.

Kozulin, A. (1986). The concept of activity in Soviet psychology: Vygotsky, his disciples and critics. *American Psychologist*, 41(3):264.

Kozulin, A. (1998). *Psychological Tools: A Sociocultural Approach to Education*. Harvard University Press, Cambridge, MA.

Kozulin, A. (2012). Vygotsky in Context. In Kozulin, A., editor, *Thought and Language*, pages XXV–XXIII. The MIT Press.

Krathwohl, D. R. and Payne, D. A. (1971). Defining and assessing educational objectives. In Thorndike, R. L., editor, *Educational measurement*, pages 17–129. American Council on Education, Washington DC.

Kuutti, K. (1996). Activity Theory as a Potential Framework for Human-Computer Interaction Research. In Nardi, B. A., editor, *Studying context: a comparison of activity theory, situated action models, and distributed cognition*. MIT Press.

Lajoie, S. P. (1998). Identifying an Agenda for Statistics Instruction and Assessment. In Lajoie, S. P., editor, *Reflections on Statistics Learning, Teaching, and Assessment in Grades K-12*, pages ix–xx. Routledge, New York.

Lane, D. M. and Peres, S. C. (2006). Interactive Simulations in the Teaching of Statistics: Promise and Pitfalls. In Rossman, A. J. and Chance, B. L., editors, *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)*, Salvador, Bahia, Brazil. International Statistical Institute.

Langrall, C., Nisbet, S., and Mooney, E. (2006). The interplay between students' statistical knowledge and context knowledge in analyzing data. In Rossman, A. J. and Chance, B. L., editors, *International Conference on the Teaching of Mathematics*, Bahia, Brazil.

Langrall, C. W., Makar, K., Nilsson, P., and Shaughnessy, J. M. (2017). Teaching and Learning Probability and Statistics: An Integrated Perspective. In Cai, J., editor, *Compendium* *for Research in Mathematics Education*, chapter 18. National Council for Teachers of Mathematics, Reston, VA.

Lantolf, J. P. and Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development.* Oxford University Press, Oxford.

Lantolf, J. P. and Thorne, S. L. (2007). Sociocultural theory and second language learning. In van Patten, B. and Williams, J., editors, *Theories in second language acquisition: An introduction*, pages 201–224. Lawrence Erlbaum, Mahwah, NJ.

Larsen, M. D. (2017). Advice for New and Student Lecturers on Probability and Statistics. *Journal of Statistics Education*, 14(1):77.

Laufer, E. A. and Glick, J. (1998). Expert and novice differences in cognition and activity: a practical work activity. In Engeström, Y. and Middleton, D., editors, *Cognition and communication work*, pages 177–198. Cambridge University Press, Cambridge.

Lave, J. (1988). *The Culture of Acquisition and the Practice of Understanding.* Institute for Research on Learning.

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation.* Cambridge University Press, Cambridge.

Lee, B. (1985). Intellectual origins of Vygotsky's semiotic analysis. In Wertsch, J. V., editor, *Culture, Communication and Cognition: Vygotskian Perspectives*, pages 35–65. Cambridge University Press, New York.

Lemke, J. L. (1990). *Talking Science: Language, Learning, and Values.* Ablex Publishing Corporation, Norwood, New Jersey.

Lemke, J. L. (1997). Cognition, Context and Learning: A social Semiotic Perspective. In Kirshner, D. and Whitson, J. A., editors, *Situated Cognition Social, Semiotic, and Psychological Perspectives*, pages 37–56. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Leontiev, A. N. (1972). Soviet Psychology: Activity and Consciousness. *Voprosy filosofii*, pages 180–202.

Leontiev, A. N. (1978). *Activity, Consciousness, and Personality.* Prentice Hall.

Lerman, S. (1996). Intersubjectivity in mathematics learning: A challenge to the radical constructivist paradigm? *Journal for Research in Mathematics Education*, 27(2):133–150.

Lerman, S. (2001). Cultural, Discursive Psychology: A Sociocultural Approach to Studying the Teaching and Learning of Mathematics. *Educational Studies in Mathematics*, 46(1/3):87–113.

Luria, A. R. (1928). The problem of the cultural behavior of the child. *The Pedagogical Seminary and Journal of Genetic Psychology*, 35(4):493–506.

Luria, A. R. (1976). *Cognitive Development: Its Cultural and Social Foundations*. Harvard University Press, Cambridge, MA.

Lutz, F. W. (1993). Ethnography: the wholistic approach to understanding schooling. In Hammersley, M., editor, *Controversies in classroom teaching*, pages 107–119. Open University Press, Buckingham.

Magel, R. C. (1996). Increasing student participation in large introductory statistics classes. *The American Statistician*, 50(1):51–56.

Maher, A. (2004). Learning outcomes in higher education: implications for curriculum design and student learning. *Journal of Hospitality, Leisure, Sport and Tourism Education*, 3(2):46–54.

Makar, K. (2018). Rethinking the Statistics Curriculum: Holistic, Purposeful and Layered. In Sorto, M. A., White, A., and Guyot, L., editors, *Proceedings of the Tenth International Conference on Teaching Statistics*, Kyoto, Japan. International Statistical Institute.

Makar, K. and Ben-Zvi, D. (2011). The role of context in developing reasoning about informal statistical inference. *Mathematical Thinking and Learning*, 13(1-2):1–4.

Makar, K. and Confrey, J. (2004). Secondary Teachers' Statistical Reasoning in Comparing Two Groups. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 353–373. Springer, Dordrecht, Dordrecht.

Makar, K. and Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics education Research Journal*, 8(1):82–105.

Mali, A. (2016). *Lecturers' tools and strategies in university mathematics teaching: an ethnographic study*. PhD thesis, Loughborough University, UK.

Malone, C. J., Gabrosek, J., Curtiss, P., and Race, M. (2012). Resequencing Topics in an Introductory Applied Statistics Course. *The American Statistician*, 64(1):52–58.

Manly, B. F. J. (2005). *Multivariate Statistical Methods: A primer*. Chapman & Hall/CRC, Boca Raton, third edition.

Mapolelo, D. C. (2009). Students' experiences with mathematics teaching and learning: listening to unheard voices. *International Journal of Mathematical Education in Science and Technology*, 40(3):309–322.

Marr, B. (2017). *Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things*. Kogan Page, London, 1 edition.

Marriott, N. (2014). The future of statistical thinking. *Significance*, 11(5):78–80.

Martin, M. A. (2003). "It's Like... You Know": The Use of Analogies and Heuristics in Teaching Introductory Statistical Methods. *Journal of Statistics Education*, 11(2):21.

Martínez-Mesa, J., González-Chica, D. A., Duquia, R. P., Bonamigo, R. R., and Bastos, J. L. (2016). Sampling: how to select participants in my research study? *Anais Brasileiros de Dermatologia*, 91(3):326–330.

Mathews, D. and Clark, J. (2007). Successful students' conceptions of mean, standard deviation, and the Central Limit Theorem. In *Paper presented at the Midwest Conference on Teaching Statistics*, Oshkosh, WI.

Maxwell, J. A. (2012). *Qualitative Research Design: An Interactive Approach.* SAGE Publications, London, 3rd edition.

McNamara, A. A. (2015). *Bridging the Gap Between Tools for Learning and for Doing Statistics.* PhD thesis, University of California, Los Angeles.

Miles, M. B., Huberman, M. A., and Saldana, J. M. (2013). *Qualitative Data Analysis: An Expanded Sourcebook.* Sage Publications Ltd, Thousand Oaks, California, 3rd edition.

Mills, D. J. (2002). Using Computer Simulation Methods to Teach Statistics: A Review of the Literature. *Journal of Statistics Education*, 10(1):182.

Moll, L. C. (1990). Vygotsky and Education. In Moll, L. C., editor, *Vygotsky and Education*, pages 1–30. Cambridge University Press, Cambridge.

Moore, D. S. (1988). Should Mathematicians Teach Statistics? *The College Mathematics Journal*, 19(1):3.

Moore, D. S. (1990). *On the Shoulders of Giants.* New Approaches to Numeracy. National Academies Press, Washington, D.C.

Moore, D. S. (1997). New Pedagogy and New Content: The Case of Statistics. *International Statistical Review*, 65(2):123–137.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, 93(444):1253–1259.

Moore, D. S. (2001). Undergraduate programs and the future of academic statistics. *The American Statistician*, 55(1):1–6.

Moore, T., College, G., and Legler, J. (2003). Survey on Statistics within the Liberal Arts College. Technical report.

Munby, H. (1982). The place of teachers' beliefs in research on teacher thinking and decision making, and an alternative methodology. *Instructional Science*, 11(3):201–225.

Nardi, E. (1998). Tutors' reflections upon difficulties of learning and taching mathematics at university level. In *Proceedings of the Conference of The British Society for Research into Learning Mathematics*, volume 18(1–2).

Nardi, E., Jaworski, B., and Hegedus, S. (2005). A Spectrum of Pedagogical Awareness for Undergraduate Mathematics: From tricks to techniques. *Journal of Research in Mathematics Education*, 36:284–316.

Newman, D., Griffin, P., and Cole, M. (1995). *The Construction Zone: working for cognitive change in school*. Cambridge University Press, Cambridge, 3rd edition.

Newton, J. A., Dietiker, L., and Horvath, A. (2011). Statistics education in the United States: Statistical reasoning and the statistical process. In Batanero, C., Burrill, G., and Reading, C. E., editors, *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, pages 9–13. Springer Science + Business Media B.V.

Nicholl, D. F. (2001). Future Directions for the Teaching and Learning of Statistics at the Tertiary Level. *International Statistical Review*, 69(1):11–15.

Nisbett, R. E., Krantz, D. H., Jepson, C., and Kunda, Z. (2009). The Use of Statistical Heuristics in Everyday Inductive Reasoning. Cambridge University Press, Cambridge.

O'Donovan, B., Price, M., and Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9(3):325–335.

Ofqual (2012). Qualification levels. Technical report, The Office of Qualifications and Examinations Regulation (Ofqual).

Ottaviani, M. G. (1989). A history of the teaching of statistics in higher education in Europe and the United States, 1660 to 1915. In Morris, R., editor, *Studies in mathematics education*, pages 243–252. UNESCO.

Otter, S. (1994). *Learning Outcomes in Higher Education: A Development Project Report*. National Institute of Adult Continuing Education.

Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62(3):307–332.

Park, J., DelMas, R. C., and Zieffler, A. (2011). A research-based statistics course for tertiary students. In *Proceedings of the 58th World Statistical Congress*, Dublin.

Pedrosa de Jesus, M. H. and da Silva Lopes, B. (2011). The relationship between teaching and learning conceptions, preferred teaching approaches and questioning practices. *Research Papers in Education*, 26(2):223–243.

Petocz, P. and Reid, A. (2003). Relationships between students' experience of learning statistics and teaching statistics. *Statistics education Research Journal*, 2(1):39–53.

Petocz, P., Reid, A., and Gal, I. (2017). Statistics Education Research. In Ben-Zvi, D., Makar, K., and Garfield, J. B., editors, *International Handbook on Research in Statistics Education*, pages 71–104. Springer.

Pfannkuch, M. (2006). Comparing Box Plot Distributions: A Teacher's Reasoning . *Statistics education Research Journal*, 5(2):27–45.

Pfannkuch, M. (2008). Training teachers to develop statistical thinking. In Batanero, C., Burrill, G., and Rossman, A. J., editors, *Joint ICMIIASE Study Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education.* Proceedings of the ICMI Study and IASE Round Table Conference.

Pfannkuch, M. (2017). Reimagining Curriculum Approaches. In *International Handbook of Research in Statistics Education*, pages 387–413. Springer International Publishing, Cham.

Pfannkuch, M. and Wild, C. J. (2000). Statistical Thinking an Statistical Practice: Themes Gleaned from Professional Statisticians. *Statistical Science*, 15(2):132–152.

Pfannkuch, M. and Wild, C. J. (2004). Towards an understanding of statistical thinking. In Ben-Zvi, D. and Garfield, J. B., editors, *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, pages 17–46. Kluwer Academic Publishers.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In Lester, F. K., editor, *Second handbook of research on mathematics teaching and learning*, chapter 7, pages 257–315. Information Age, Charlotte, NC.

Piaget, J. and Inhelder, B. (1975). *The Origin of the Idea of Chance in Children.* Psychology Press.

Pierce, R. and Chick, H. (2011). Teachers' Beliefs and Views About Statistics Education. In Batanero, C., Burrill, G., and Reading, C. E., editors, *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education*, pages 151–162. Springer Netherlands, New York.

Pimenta, R. (2006). Assessing statistical reasoning through project work. In Rossman, A. J. and Chance, B. L., editors, *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)*, Salvador, Bahia, Brazil.

Polanyi, M. (1974). Scientific Thought and Social Reality: Essays by Michael Polanyi. *Psychological Issues*, 8(4, Mono 32).

Polanyi, M. (2009). *The Tacit Dimension.* University of Chicago Press, Chicago.

Pollard, A. (2014). *Reflective Teaching in Schools.* Bloomsbury USA Academic.

Pollatsek, A., Konold, C. E., Well, A. D., and Lima, S. D. (1984). Beliefs underlying random sampling. *Memory & Cognition*, 12(4):395–401.

Pollatsek, A., Lima, S., and Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12(2):191–204.

Popkewitz, T. S. (1998). Dewey, Vygotsky, and the Social Administration of the Individual: Constructivist Pedagogy as Systems of Ideas in Historical Spaces. *American Educational Research Journal*, 35(4):535–570.

Porter, T. M. (1986). *The Rise of Statistical Thinking, 1820-1900.* Princeton University Press.

Preskill, S. (1998). Narratives of Teaching and the Quest for the Second Self. *Journal of Teacher Education*, 49(5):344–357.

Presmeg, N. C. (1992). Prototypes, metaphors, metonymies and imaginative rationality in high school mathematics. *Educational Studies in Mathematics*, 23(6):595–610.

Price, D. W. W. (1990). A Model for Reading and Writing about Primary Sources: The Case of Introductory Psychology. *Teaching of Psychology*, 17(1):48–53.

Price, M. (2005). Assessment standards: the role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education*, 30(3):215–230.

Price, M. and Rust, C. (1999). The Experience of Introducing a Common Criteria Assessment Grid Across an Academic Department. *Quality in Higher Education*, 5(2):133–144.

Pulvermacher, Y. and Lefstein, A. (2016). Narrative representations of practice: What and how can student teachers learn from them? *Teaching and Teacher Education*, 55:255–266.

QAA (2007a). Outcomes from institutional audit The adoption and use of learning outcomes. Technical report.

QAA (2007b). Subject Benchmark Statement: Economics. The Quality Assurance Agency for Higher Education, UK.

QAA (2010). Subject Benchmark Statement: Psychology. The Quality Assurance Agency for Higher Education, UK.

QAA (2011). UK Quality Code for Higher Education, Part B: Assuring and enhancing academic quality, Chapter B6: Assessment of students and accreditation of prior learning. *The Quality Assurance Agency for Higher Education, UK.*

QAA (2013). UK Quality Code for Higher Education, Part A: Setting and maintaining academic standards. *QAA.*

Quilici, J. L. and Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, 88(1):144–161.

Quilici, J. L. and Mayer, R. E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Applied Cognitive Psychology*, 16(3):325–342.

Ragasa, C. Y. (2008). A Comparison of Computer-Assisted Instruction and the Traditional Method of Teaching Basic Statistics. *Journal of Statistics Education*, 16(1):249.

Rayment, T. (2000). Art teachers' views of national curriculum art: a repertory grid analysis. *Educational Studies*, 26(2):165–176.

Raymond, M. R. and Neustel, S. (2006). Determining the Content of Credentialing Examinations. In Downing, S. M. and Haladyna, T. M., editors, *Handbook of test development*, pages 181–223. Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Reading, C. E. (2011). Fundamentals for teaching statistics. In Batanero, C., Burrill, G., and Reading, C. E., editors, *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study New ICMI Study Series*. Springer Dordrecht, Heidelberg.

Renshaw, P. (1996). Mathematics for Tomorrow's Young Children. In Mansfield, C. S., Pateman, N. A., and Bednarz, N., editors, *Mathematics for Tomorrows Young Children*. Kluwer Academic Publishers, Quebec.

Resnick, L. B. (1987a). *Education and Learning to Think - Committee on Research in Mathematics, Science, and Technology Education, Commission on Behavioral and Social Sciences and Education, Division of Behavioral and Social Sciences and Education - Google Books*. National Academy Press, Washington, D.C.

Resnick, L. B. (1987b). The 1987 Presidential Address Learning In School and Out. *Educational Researcher*, 16(9):13–54.

Rezat, S. and Sträßer, R. (2012). From the didactical triangle to the socio-didactical tetrahedron: artifacts as fundamental constituents of the didactical situation. *ZDM*, 44(5):641–651.

Rodd, M. (2003). Witness as participation: the lecture theatre as site for mathematical awe and wonder. *For the Learning of Mathematics*, 23(1):15–21.

Rogoff, B. (1982). Integrating context and cognitive development. In Lamb, M. E. and Brown, A. L., editors, *Advances in developmental psychology*, pages 125–170. Psychology Press, Hillsdale, NJ.

Rosenshine, B. (1995). Advances in Research on Instruction. *Journal of Educational Research*, 88(5):262–268.

Rosenshine, B. (2010). Principles of instruction series 21. Technical Report IBE/2010/ST/EP21.

Rosenshine, B. (2012). Principles of Instruction: Research-Based Strategies That All Teachers Should Know. *American Educator*, 36(1).

Rossman, A. J. and Chance, B. L. (1999). Teaching the Reasoning of Statistical Inference: A "Top Ten" List. *The College Mathematics Journal*, 30(4):297.

Rossman, A. J. and Chance, B. L. (2014). Using simulation-based inference for learning introductory statistics. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(4):211–221.

Ryan, J. and Williams, J. (2007). *Children's mathematics 4 –15.* Learning from errors and misconceptions. Open University Press, Berkshire.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2):119–144.

Sahai, H. (1990). Some Problems of Teaching an Introductory Biostatistics Course to Graduate Students in Health Sciences: Coping with the Diversity of Student Aptitudes, Interests and Objectives. *The Statistician*, 39(4):341.

Saroyan, A. and Snell, L. S. (1997). Variations in lecturing styles. *Higher Education*, 33(1):85–104.

Schau, C. and Mattern, N. (1997). Use of Map Techniques in Teaching Applied Statistics Courses. *The American Statistician*, 51(2):171–175.

Schmidt, H. G., Cohen-Schotanus, J., van der Molen, H. T., Splinter, T. A. W., Bulte, J., Holdrinet, R., and van Rossum, H. J. M. (2009). Learning more by being taught less: a "time-for-self-study" theory explaining curricular effects on graduation rate and study duration. *Higher Education*, 60(3):287–300.

Schmittau, J. (2004). Vygotskian theory and mathematics education: Resolving the conceptual-procedural dichotomy. *European Journal of Psychology of Education*, 19(1):19–43.

Schoenfeld, A. (1998). Toward a theory of teaching-in-context. *Issues in Education*, 4(1):1–94.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In Grouws, D., editor, *Handbook for Research on Mathematics Teaching and Learning*. MacMillan, New York.

Schwandt, T. A. (2001). *Dictionary of Qualitative Inquiry.* Sage Publications, Inc, London, 2nd edition.

Seabrook, R. (2006). Is the Teaching of Statistical Calculations Helpful to Students' Statistical Thinking? *Psychology Learning & Teaching*, 5(2):153–161.

Seaman, J. (2008). Adopting a Grounded Theory Approach to Cultural-Historical Research: Conflicting Methodologies or Complementary Methods?:. *International Journal of Qualitative Methods*, 7(1):1–17.

Sedlmeier, P. and Wassner, C. (2008). German Mathematics Teachers' Views on Statistics Education. In Batanero, C., Burrill, G., and Rossman, A. J., editors, *Joint ICMI/IASE Study Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education*. Proceedings of the ICMI Study and IASE Round Table Conference.

Seeger, F. (1998). Representations in the mathematics classroom: reflections and constructions. In Seeger, F., Voigt, J., and Waschescio, U., editors, *The Culture of the Mathematics Classroom*, pages 308–343. Cambridge University Press, Cambridge.

Senior, B. and Swailes, S. (2004). The dimensions of management team performance: a repertory grid study. *International Journal of Productivity and Performance Management*, 53(4):317–333.

Sfard, A. (2015). Learning, commognition and mathematics. In Scott, D. and Hargreaves, E., editors, *The Sage Handbook of Learning*, pages 129–138. Sage, London.

Shadbolt, N. and Milton, N. (1999). From knowledge engineering to knowledge management. *British Journal of Management*, 10(4):309–322.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In Grouws, D. A., editor, *Handbook of Research on Mathematics Teaching and Learning*, pages 465–494. National College of Teachers of Mathematics, Reston, Virginia.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In Biddulph, F. and Carr, K., editors, *Proceedings of the 20th annual conference of the Mathematics Education Research Group of Australasia*, pages 6–22, Waikato, NZ. MERGA, People in mathematics education.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In Lester, F. K. J., editor, *Second Handbook of Research on Mathematics Teaching and Learning*. Iap.

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2):4–14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57(1):1–23.

Sirotnik, K. (1983). What You See Is What You Get-Consistency, Persistency, and Mediocrity in Classrooms. *Harvard educational review*, 53(1):16–31.

Snee, R. D. (1993). What's missing in statistical education? *American Statistician*, 47(2):149–154.

Sowey, E. R. (1995). Teaching Statistics: Making It Memorable. *Journal of Statistics Education*, 3(2):1–9.

Speer, N. M., Smith III, J. P., and Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *The Journal of Mathematical Behavior*, 29(2):99–114.

Staats, S. (2007). Dynamic contexts and imagined worlds: An interdisciplinary approach to mathematics applications. *For the learning of mathematics*, 27(1):4–9.

Steffe, L. P. and Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In Lesh, R. A. and Kelly, A. E., editors, *Research design in mathematics and science education*, pages 267–307. Erlbaum, Hillside, NJ.

Stein, M. K., Remillard, J., and Smith, M. S. (2007). Second Handbook of Research on Mathematics Teaching and Learning. In Lester, F. K. J., editor, *How curriculum influences student learning*. Information Age Publishing.

Stetskenko, A. (2011). Darwin and Vygotsky on Development: An Exegesis on Human Nature. In Kontopodis, M., Wulf, C., and Fichtner, B., editors, *Children, Development and Education Cultural, Historical, Anthropological Perspectives*, pages 25–40. Springer Netherlands, Dordrecht.

Stoloff, M. L., Curtis, N. A., Rodgers, M., Brewster, J., and McCarthy, M. A. (2012). Characteristics of Successful Undergraduate Psychology Programs. *Teaching of Psychology*, 39(2):91–99.

Suchman, L. A. (1985). *Plans and situated actions*. PhD thesis, Palo Alto Research Centre, Palo Alto.

Summers, J. J. and Davis, H. A. (2006). Introduction: The Interpersonal Contexts of Teaching, Learning, and Motivation. *The Elementary School Journal*, 106(3):189–191.

Swanson, D. B. and McKibben, J. N. (1998). On Teaching Statistics to Non-Specialists: a Course Aimed at Increasing Both Learning and Retention. In Pereira-Mendoza, L., Kea, L. S., Kee, T. W., and Wong, W.-K., editors, *The Proceedings of the Fifth International Conference on Teaching Statistics (ICOTS5)*, Singapore.

Swanson, T., VanderStoep, J., and Tintle, N. L. (2014). Student attitudes toward statistics from a randomization-based curriculum. In Makar, K., de Sousa, B., and Gould, R., editors, *The Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)*, pages 1–5, Flagstaff, Arizona, United States.

Switzer, S. S. and Horton, N. J. (2007). What your doctor should know about statistics (but perhaps doesn't...). *Chance*, 20(1):17–21.

Tabach, E. (1999). Activity theory and the concept of integrative levels. In Engeström, Y., Miettinen, R., and Punamaki, R.-L., editors, *Perspectives on Activity theory*. Cambridge University Press, Cambridge.

Tempelaar, D. T., Gijselaers, W. H., and van der Loeff, S. S. (2017). Puzzles in Statistical Reasoning. *Journal of Statistics Education*, 14(1):129.

The British Educational Research Association [BERA] (2011). Ethical guidelines for educational research.

The British Psychological Society (2012). Accreditation through partnership handbook Guidance for sport and exercise psychology programmes. Technical report, The British Psychological Society.

The Engineering Council (2013). UK Standards for professional engineering competence. Technical report, The Engineering Council.

Thomas, S. (2011). *An activity theory analysis of linear algebra teaching within university mathematics*. PhD thesis, Loughborough University.

Thompson, A. G. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, 15(2):105–127.

Tintle, N. L., Chance, B. L., Cobb, G. W., Rossman, A. J., Roy, S., Swanson, T., and VanderStoep, J. (2014). *Introduction to Statistical Investigations*. John Wiley & Sons, preliminary edition edition.

Tintle, N. L., Topliff, K., and VanderStoep, J. (2012). Retention of statistical concepts in a preliminary randomization-based introductory statistics curriculum. *Statistics Education Research Journal*, 11(1):21–40.

Tintle, N. L., VanderStoep, J., and Holmes, V. L. (2011). Development and assessment of a preliminary randomization-based introductory statistics curriculum. *Journal of Statistics Education*, 19(1).

Tishkovskaya, S. and Lancaster, G. A. (2010). Teaching Strategies to Promote Statistical Literacy: Review and Implementation. In Reading, C. E., editor, *International Conference on the Teaching of Statistics*, Ljubljana, Slovenia.

Tishkovskaya, S. and Lancaster, G. A. (2012). Statistical Education in the 21st Century: a Review of Challenges, Teaching Innovations and Strategies for Reform. *Journal of Statistics Education*, 20(2).

Tofan, D., Galster, M., and Avgeriou, P. (2011). *Capturing tacit architectural knowledge using the repertory grid technique*. ACM, New York, USA.

Tolman, C. W. (1999). Society versus context in individual development: Does theory make a difference? In Engeström, Y., Miettinen, R., and Punamaki, R.-L., editors, *Perspectives on Activity theory*, pages 70–85. Cambridge University Press.

Treffert-Thomas, S. and Jaworski, B. (2015). Developing mathematics teaching: what can we learn from the literature? In Grove, M., editor, *Transitions in Undergraduate Mathematics Education*, pages 259–276. University of Birmingham; Higher Education Academy.

Trowler, P. and Knight, P. T. (2000). Coming to Know in Higher Education: Theorising faculty entry to new work contexts. *Higher Education Research & Development*, 19(1):27–42.

Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist*, 24(2):83.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Pearson.

Tversky, A. and Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. *Utility*.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4):293–315.

Tversky, A. and Kahneman, D. (2009). Belief in the law of small numbers. Cambridge University Press, Cambridge.

Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction*. University of Chicago Press, Chicago.

Urquhart, N. S. (1971). Nonverbal Instructional Approaches for Introductory Statistics. *The American Statistician*, 25(2):20.

Valsiner, J. (1987). *Culture and the development of children's action: A cultural-historical theory of developmental psychology*. John Wiley, New York.

Van Den Berg, M. N. and Hofman, W. H. A. (2005). Student Success in University Education: A Multi-measurement Study of the Impact of Student and Faculty Factors on Study Progress. *Higher Education*, 50(3):413–446.

van der Veer, R. (1998). From concept attainment to knowledge formation. *Mind*, 5(2):89–94.

van der Veer, R. (2007). Vygotsky in Context. In Daniels, H., Cole, M., and Wertsch, J. V., editors, *The Cambridge companion to Vygotsky*, pages 20–48. Cambridge University Press, New York.

Viirman, O. (2014). *The function concept and university mathematics teaching*. PhD thesis, Karlstad University Studies.

Viirman, O. (2015). Explanation, motivation and question posing routines in university mathematics teachers' pedagogical discourse: a commognitive analysis. *International Journal of Mathematical Education in Science and Technology*, 46(8):1165–1181.

von Glasersfeld, E. (1998). Cognition, Construction of Knowledge, and Teaching. *Synthese*, 80(1):121–140.

Vos, P. (2011). What Is 'Authentic'in the Teaching and Learning of Mathematical Modelling? In Kaiser, G., Blum, W., Ferri, R. B., and Stillman, G., editors, *Trends in Teaching and Learning of Mathematical Modelling*, pages 713–722. Springer, New York.

Vygotsky, L. S. (1978a). Interaction between learning and development. In *Mind and society*, pages 79–91. Harvard University Press, Cambridge, Massachussets.

Vygotsky, L. S. (1978b). *Mind in Society: The development of higher psychological processes.* Harvard University Press, Cambridge, Massachussets.

Vygotsky, L. S. (1981). Genesis of the higher mental functions. In Wertsch, J. V., editor, *The concept of activity in Soviet psychology*, pages 144–176. M E Sharpe, New York.

Vygotsky, L. S. (1987). Thinking and speech. In Rieder, R. W. and Carton, A. S., editors, *The Collected works of L. S. Vygotsky.* Plenum Press, New York.

Vygotsky, L. S. (1997). *Educational Psychology.* CRC Press.

Vygotsky, L. S. (1998). Infancy. In Rieber, R. W., editor, *The Collected works of L. S. Vygotsky*, pages 207–242. Kluwer Academic, New York.

Vygotsky, L. S. (2012). *Thought and Language.* The MIT Press.

Wagner, J. (1997). The unavoidable intervention of educational research: A framework for reconsidering researcher-practitioner cooperation. *Educational Researcher*, 26(7):13–22.

Ware, M. E. and Chastain, J. D. (1991). Developing Selection Skills in Introductory Statistics. *Teaching of Psychology*, 18(4):219–222.

Wartofsky, M. W. (1973). *Models.* D.Reidel, Dordrecht.

Washburn, G. N. (1994). Vygotskian Approaches to Second Language Research. In Lantolf, J. P. and Appel, G., editors, *Vygotskian Approaches to Second Language Research*, pages 69–82. Ablex Publishing, Westport, CT.

Watkins, C. and Mortimore, P. (1999). Pedagogy: what do we know? . In *Understanding Pedagogy and its Impact on Learning*, pages 1–19. Paul Chapman/Sage, London.

Watson, J. M. (1997). Assessing statistical thinking using the media. In Gal, I. and Garfield, J. B., editors, *The Assessment Challenge in Statistics Education*, pages 107–122. The International Statistical Institute, Amsterdam.

Watson, J. M. (2000). Statistics in Context. *The Mathematics Teacher*, 93(1):54–58.

Watson, J. M. (2001). Profiling Teachers' Competence and Confidence to Teach Particular Mathematics Topics: the Case of Chance and Data. *Journal of Mathematics Teacher Education*, 4(4):305–337.

Watson, J. M. and Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics education Research Journal*, 2(2):3–46.

Watson, J. M. and Moritz, J. B. (2000). Development of understanding of sampling for statistical literacy. *The Journal of Mathematical Behavior*, 19(1):109–136.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *The Journal of Mathematical Behavior*, 23(2):115–133.

Wells, G. (2007). Exploring the Relationship between Brain and Culture. *Human Development*, 49(6):358–362.

Wenger, E. (2000). *Communities of Practice: Learning, Meaning, And Identity*. Cambridge University Press, Cambridge.

Wertsch, J. V. (1988). *Vygotsky and the Social Formation of Mind*. Harvard University Press.

Wertsch, J. V. (1992). The Voice of Rationality in a Sociocultural Approach to Mind: : Instructional Implications and Applications of Sociohistorical Theory. In Moll, L. C., editor, *Vygotsky and Education*, pages 111–126. Cambridge University Press, Cambridge.

Wertsch, J. V. (1993). *Voices of the Mind: a Sociocultural Approach to Mediated Action*. Harvard University Press, Cambridge.

Wertsch, J. V. (1995). Sociocultural Research in the Copyright Age. *Culture & Psychology*, 1(1):81–102.

Wertsch, J. V. (1998). *Mind as Action*. Oxford University Press, Oxford.

Wertsch, J. V. (2000). *Intersubjectivity and Alterity in Human Communication*. Ablex Publishing Corporation, Stamford, Conneticut.

Wertsch, J. V. and Stone, A. (1985). The concept of internalisation in Vygotsky's account of the genesis of higher mental functions. In Wertsch, J. V., editor, *Culture, Communication and Cognition: Vygotskian Perspectives*, pages 162–182. Cambridge University Press, Cambridge.

Wertz, F. J., Charmaz, K., McMullen, L. M., Josselson, R., Anderson, R., and McSpadden, E. (2011). *Five Ways of Doing Qualitative Analysis: Phenomenological Psychology, Grounded Theory, Discourse Analysis, Narrative Research, and Intuitive*. The Guildford Press, first edition.

Westbury, I. and Travers, K. (1990). Second International Mathematics Study. Technical report, Illinois University, Urbana College of Education, Washington, DC.

Wild, C. J. and Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67(3):223–248.

Williams, J. and Wake, G. (2006). Black Boxes in Workplace Mathematics. *Educational Studies in Mathematics*, 64(3):317–343.

Wulff, S. S. and Wulff, D. H. (2004). "Of Course I'm Communicating; I Lecture Every Day": enhancing teaching and learning in introductory statistics. *Communication Education*, 53(1).

Yin, R. K. (1994). *Case study research*. Sage Publications, London.

Yorke, D. M. (1978). Repertory Grids in Educational Research: some methodological considerations. *British Educational Research Journal*, 4(2):63–74.

Zieffler, A., Garfield, J. B., Alt, S., Dupuis, D., Holleque, K., and Chang, B. (2008). What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature. *Journal of Statistics Education*, 16(2).

Zieffler, A., Garfield, J. B., DelMas, R. C., and Bjornsdottir, A. (2010). Development of an instrument to assess statistical thinking. In Reading, C. E., editor, *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)*, Ljubljana, Slovenia.

Zieffler, A., Park, J., Garfield, J. B., DelMas, R. C., and Bjornsdottir, A. (2012). The Statistics Teaching Inventory: A Survey on Statistics Teachers' Classroom Practices and Beliefs. *Journal of Statistics Education*, 20(1).

Zinchenko, V. P. (1985). Vygotsky and units for the analysis of mind. In Wertsch, J. V., editor, *Culture, Communication and Cognition: Vygotskian Perspectives*, pages 94–118. Cambridge University Press, Cambridge.

Appendices

A | A definition of statistics

Although no formal definition of statistical thinking, reasoning and literacy has been agreed in the literature, in my pilot study I used the following definition from Ben-Zvi and Garfield (2004, p. 7)

- **Statistical literacy** includes basic and important skills that may be used in understanding statistical information or research results. These skills include being able to organize data, construct and display tables, and work with different representations of data. Statistical literacy also includes an understanding of concepts, vocabulary, and symbols, and includes an understanding of probability as a measure of uncertainty.
- **Statistical reasoning** may be defined as the way people reason with statistical ideas and make sense of statistical information. This involves making interpretations based on sets of data, representations of data, or statistical summaries of data. Statistical reasoning may involve connecting one concept to another (e.g., center and spread), or it may combine ideas about data and chance. Reasoning means understanding and being able to explain statistical processes and being able to fully interpret statistical results.
- **Statistical thinking** involves an understanding of why and how statistical investigations are conducted and the "big ideas" that underlie statistical investigations. These ideas include the omnipresent nature of variation and when and how to use appropriate methods of data analysis such as numerical summaries and visual displays of data. Statistical thinking involves an understanding of the nature of sampling, how we make inferences from samples to populations, and why designed experiments are needed in order to establish causation. It includes an understanding of how models are used to simulate random phenomena, how data are produced to estimate probabilities, and how, when, and why existing inferential tools can be used to aid an investigative process. Statistical thinking also includes being able to understand and utilize the context of a problem in forming investigations and drawing conclusions, and recognizing and understanding the entire process (from question posing to data collection to choosing analyses to testing assumptions, etc.). Finally, statistical thinkers are able to critique and evaluate results of a problem solved or a statistical study.

B | Pilot study

In this section, I include the list of actual statistics modules included in the data analysis of learning outcomes (Pilot 1).

B.1 Factor anlaysis (Pilot 1 and 2)

To analyse the repertory grids in my pilot study, I used factor analysis plots, which represented the elements, constructs and factors graphically, to link the elements to the construct(s) that were close to them (Manly, 2005). Using Idiogrid 2.4, I then classified each repertory grid element together with the construct(s) that explained it. Using the same categories, I compared actual module learning outcomes, the elements produced by participants and the element-construct(s) labels. In this section, I exemplify one such analysis of the grid in Table B.1.

In this analysis, I applied open categories to the extracted factors in order to identify emerging themes. Their properties were further categorised into sub-themes. In order to identify and characterise the collection of constructs that attracted most interest, a frequency analysis identified how often each theme occurred. The results were not absolutes, but a set of relationships between elements and their constructs in the repertory grid. The literature was not emphasised in this case as it could constrain the data analysis (i.e. classifying). Instead, the literature was accessed as it became relevant and used to refine the findings in light of it. Below, I exemplify the Factor analysis using data from one participant.

The first step in the quantitative analysis was to undertake a separate Factor analysis for each grid, with the aim of deriving a smaller set of uncorrelated (i.e. independent) factors, which can be used to describe the larger number of statements in a grid. Hence, the statements are treated as the 'variables' in a classical Factor analysis, whilst the elements serve the role of individual observations. As there were typically more constructs than elements, the factor analysis was used as an exploratory study (Chatfield, 1988). The result is a set of new variables or **Factors** that are smaller in number than the number of constructs, such that each construct can be expressed as a *linear combination* of these Factors.

Table B.1: An example of a repertory grid from one participant

Construct (positive pole, C_+)		Elements* E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	Construct (negative pole, C_-)	
C_{1+}	E_3 is interpreted using E_5	1	7	1	7	2	7	6	1	6	qualitative interpretation	C_{1-}
C_{2+}	description and interpretation of data	7	6	4	1	6	1	4	4	1	inferential stats	C_{2-}
C_{3+}	analysing and interpreting data	3	7	4	1	3	1	3	2	1	sampling	C_{3-}
C_{4+}	methods of analysis	1	7	1	3	1	4	1	1	4	sampling	C_{4-}
C_{5+}	interpretation of data and stats, what it means	7	6	6	1	2	5	4	7	1	analytical method, mechanics of doing it	C_{5-}
C_{6+}	qualitative description & interpretation	7	2	6	1	7	1	4	7	1	quantitative description	C_{6-}
C_{7+}	techniques	1	2	1	6	7	2	4	1	7	how to interpret results of analysis	C_{7-}
C_{8+}	inferential	1	4	2	6	1	7	3	2	4	descriptive	C_{8-}
C_{9+}	interpreting results and what they mean	7	5	6	1	2	5	4	7	1	applying techniques	C_{9-}
C_{10+}	simple methods	7	2	7	1	6	1	2	4	2	more complex methods	C_{10-}
C_{11+}	method/design	1	1	1	7	7	2	4	1	7	interpreting results	C_{11-}
C_{12+}	method	1	3	1	6	4	1	3	1	7	being critical	C_{12-}
C_{13+}	understanding data	1	7	1	1	1	1	1	1	1	collecting data	C_{13-}
C_{14+}	low level	4	2	5	3	5	1	3	4	6	high level	C_{14-}
C_{15+}	least demanding year 1	5	3	NA	6	5	3	3	5	7	most demanding year 1	C_{15-}
C_{16+}	least demanding year 2	4	2	5	6	5	2	3	4	6	most demanding year 2	C_{16-}
C_{17+}	least demanding year 3	4	2	5	5	5	1	3	4	5	most demanding year 3	C_{17-}

*The participant elicited 17 bipolar **constructs** (positive C_+ and negative C_-), nine **elements** (E_1 to E_9) and 152 **ratings**.

E_1: be able to use simple inferential statistics;

E_2: appreciate different sampling methods - when each might be used ;

E_3: have a basic appreciation of multivariate methods;

E_4: be able to interpret graphs and tables and draw reasonable conclusions;

E_5: understand confidence intervals;

E_6: be able to describe data using descriptive statistics and appreciate graphical methods;

E_7: be able to apply correlation methods and interpret results;

E_8: be able to apply regression analysis;

E_9: be critical when interpreting data and stats given to students (e.g. in media).

Table B.2: Eigenvalues for all factors (unrotated)

Factor	Eigenvalue	% Variance	Cumulative%
1	**7.42**	**43.63**	**43.63**
2	**5.98**	**35.18**	**78.80**
3	**2.06**	**12.14**	**90.94**
4	0.88	5.20	96.14
5	0.23	1.38	97.52
6	0.19	1.10	98.63
7	0.16	0.94	99.57
8	0.07	0.43	100.00

Idiogrid generated the output in Table B.2, for the repertory grid data shown in Table B.1. This relates to what are referred to as unrotated factors, which I first consider here. Rotating the factors is an approach that can produce an equivalent set of factors which explain the same amount of information in the statements as in the unrotated case, but rotation can often make interpretation of the underlying meaning of the factors easier. Each factor is derived from one of the eigenvectors associated with the correlation matrix relating to the original repertory grid.

Table B.2 provides information about each factor in descending order of the size of the eigenvalue associated with the eigenvector from which it was formed. The eigenvalues also indicate how much of the variation in the statements is explained by that factor and so the factor with the highest eigenvalue is the most informative. The total variance in the statements explained by the factors is given by the sum of the eigenvalues and so Idiogrid also displays the eigenvalue in terms of percentage of variance explained (e.g. factor 1 explains 43.63% of total variance). For repertory grid data, a factor that explains 50% of variance is considered high.

Next I made a decision regarding the number of factors that should be extracted using the output in Table 6 based on how many eigenvalues were greater than unity (Manly, 2005). In this case, there were three factors with eigenvalues greater than one, ranging from 7.42 down to 2.06, which together explained 90.94% of variance. In the context of Factor analysis, this represents a very good solution. I interpreted this to mean that the three factors provided the best explanation of the information and inter-correlations contained amongst the different statements and hence best indicated a summary of that participant's views of the statistical abilities students ought to learn.

The eigenvectors of the correlation matrix are used to determine factor loadings, which are the coefficients used to represent each construct as a linear combination of the factors. These are shown in Table B.3 for three factors. The output also generates communalities for all constructs, which represent the proportion of variance in a construct that is explained by the factors. In this example, the communalities are all quite high (close to one) which again suggests the solution is a good one. Most of the variance for the seventeen constructs is therefore accounted for by the three common factors. Factor loadings that are 0.50 or more (regardless of the valence) are considered large or moderate loadings and indicate how the constructs are related to the factors. In B.3 it can be seen that eight constructs depend

Table B.3: Factor loadings and communalities (unrotated)

Variable	F1	F2	F3	Communalities
C1+	-0.52	0.83	-0.07	0.97
C2+	0.69	-0.33	-0.54	0.88
C3+	0.58	0.19	-0.76	0.95
C4+	-0.18	0.79	-0.43	0.84
C5+	0.96	0.05	0.12	0.95
C6+	0.60	-0.76	-0.03	0.93
C7+	-0.90	-0.15	-0.25	0.89
C8+	-0.45	0.79	0.33	0.94
C9+	0.95	-0.03	0.23	0.96
C10+	0.51	-0.79	-0.13	0.91
C11+	-0.92	-0.19	-0.13	0.91
C12+	-0.93	0.00	-0.35	0.98
C13+	0.28	0.58	-0.73	0.95
C14+	-0.28	-0.86	-0.21	0.85
C15+	-0.63	-0.59	-0.01	0.75
C16+	-0.64	-0.70	-0.02	0.89
C17+	-0.41	-0.85	-0.15	0.91

on more than one factor, shown in bold (e.g. C6, C13, C15 and C16).

Next, I used Idiogrid to carry out a varimax rotation with these three factors to find a new solution that rotates the factors which are thus easier to interpret. The text output in Table B.4 contains the eigenvalues for the varimax rotated factors ($F*$) and the factor loadings for all statements that give the coefficients for this new solution. The communalities are, as expected, the same and the rotated factors are still uncorrelated. By comparing the relative factor loadings of the various constructs on the factors, I could identify the constructs that had the highest factor loadings on this factor. This rotated solution seems better than the unrotated factor model since only four of the constructs are dependent on more than one factor. In the case of repertory grids, varimax factor loadings greater than 0.7 or smaller than -0.7 can be considered substantial in interpreting them since this would indicate at least 50% (0.702) overlap between the constructs and the factors (Grice, 2007). In analysing the grids, I therefore used this criterion as it made the data easier to interpret, represented in bold in Table B.4.

Factor analysis uses the pattern of correlations among constructs to estimate how much the ratings on each construct depend on the hypothetical Factor (i.e. the Factor loadings for each Factor). Examining the content of each construct that loads on each Factor provides the substantive meaning of a Factor. For example, consider that the factor loadings range from -1 to +1. If all the constructs that positively load on a Factor involve ‘using statistical techniques’ (e.g. C7-, C11- and C12- using the same notations as in Table B.1 for these constructs), that negatively load on a Factor involve ‘interpreting data’ (e.g. C7+, C11+ and C12+) but other constructs about say ‘statistical reasoning’ have low or zero loadings on the Factor, the Factor could be identified with ‘interpreting data versus analysing data’. This interpretation therefore is developed *inductively*, based on a combination of formal mathematical modelling and subjective judgements that make the link between observable data and hypothetical Factors (Kane, 2006).

Table B.4: Eigenvalues of factors and factor loadings (varimax)

Factor	Eigenvalue	%Variance	Cumulative%	Construct	Factor loadings 1	2	3
1	6.54	38.48	38.48	C1+	-0.22	**0.94**	-0.19
2	6.06	35.64	74.13	C2+	0.30	**-0.68**	-0.57
3	2.86	16.81	90.94	C3+	0.29	-0.22	**-0.90**
				C4+	-0.06	**0.70**	-0.59
				C5+	**0.90**	-0.32	-0.17
				C6+	0.28	**-0.92**	0.06
				C7+	-0.92	0.18	0.06
				C8+	-0.02	**0.96**	0.17
				C9+	**0.91**	-0.36	-0.04
				C10+	0.15	**-0.94**	-0.01
				C11+	**-0.92**	0.17	0.19
				C12+	**-0.94**	0.30	-0.07
				C13+	0.15	0.26	**-0.93**
				C14+	-0.59	**-0.70**	0.16
				C15+	**-0.74**	-0.28	0.35
				C16+	**-0.78**	-0.37	0.38
				C17+	**-0.68**	-0.63	0.24

In order to understand the patterns defined by factor analysis and identify the relationships among constructs, elements and factors graphically, for each grid, the elements, constructs and factors were plotted against each other. Distances between and among elements, constructs and factors suggest how they may be related to each other (Curtis et al., 2008). Idiogrid generates a plot for each pair of factors that simultaneously represent the elements, constructs and factors. The plot relating to Factors 1 and Factor 2 is shown in B.1. Each factor can be thought of as defining a coordinate axis of such a plot. In Figure 5, Factor 1 is the X axis, Factor 2 is the Y axis and the blue dots inside the plot indicate the Factor loadings for Factor 1 and Factor 2, e.g. S16- is plotted using the factor loading in Table 8, such that in I have +0.78 on the X-axis and +0.37 on the Y-axis vs. S16+ with -0.78 on the X-axis and -0.37 on the Y-axis.

The constructs that loaded on the first Factor are C16-, C15-, C11-, C7-, C12-, C5+ and C9+ versus C9-, C5-, C12+, C7+, C11+, C15+ and C16+, represented along the right and left hand side of the plot. In this case, the sets of constructs 'interpretation of data and stats - what it means', 'interpreting results and what they mean' versus 'techniques', 'method/design', 'method', 'most demanding' represent Factor 1. The grey lines represent the distance between the construct factor loadings (blue dots) and origin.

The red crosses represent the Factor score for each element that indicate which elements have higher or lower Factor scores for each Factor and hence are similar or dissimilar to each other with respect to the Factors. In these Factor analysis plots, groups of statements that are close together (i.e. groups of blue dots) are assumed to be similar in meaning as are elements that are close together (e.g. E1, E3 and E8 which seem to be about statistical techniques versus E4 and E9 in the opposite quadrant which appear to be about interpretation and critical analysis, i.e. statistical reasoning).

Further, elements in a particular quadrant are explained by the construct represented

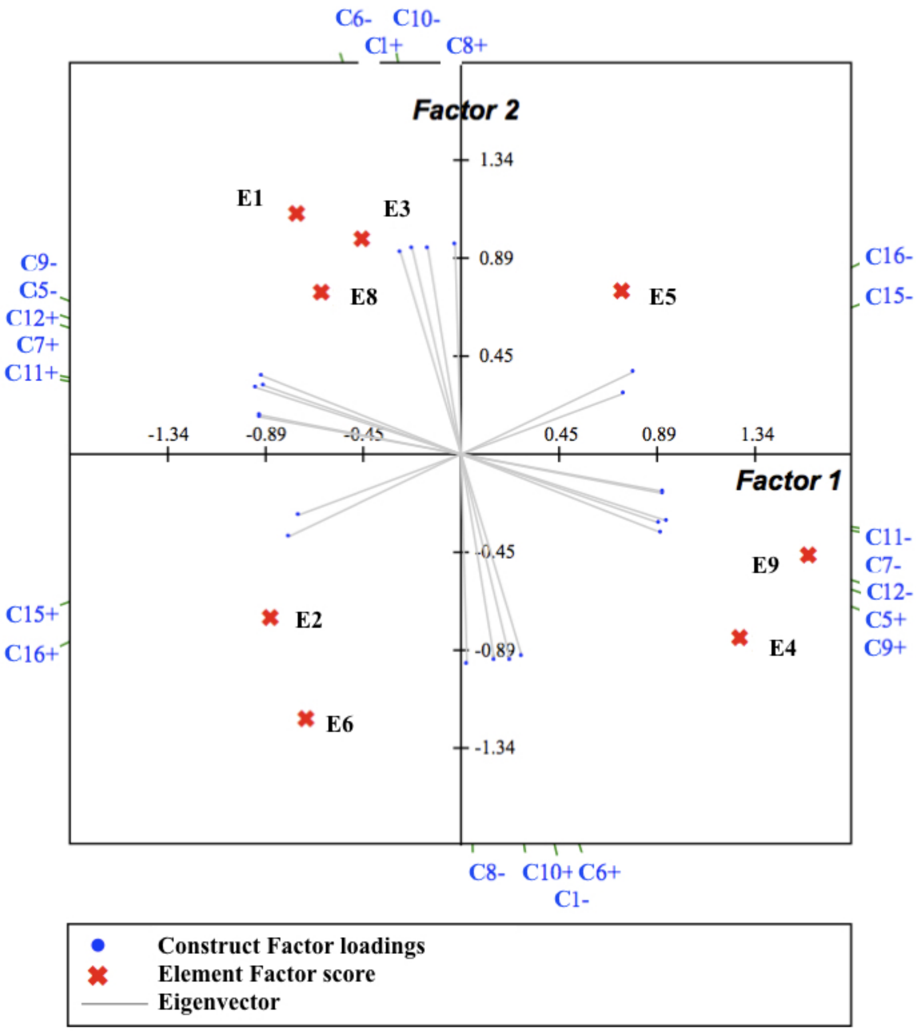


Figure B.1: Factor analysis plot

alongside that quadrant. For instance, E1 'inferential statistics' and E8 'regression analysis' are about C6- 'quantitative description', C1+ 'E3 is interpreted using E5', C10- 'more complex methods' and C8+ 'inferential' as well as C9- 'applying techniques', C5- 'analytical method, mechanics of doing it' and C12+ 'method'.

Once all the grids were analysed following the same procedures as shown here, I was able to interpret each factor based on the set of constructs with the highest Factor loadings and each element based on the Factors that defined it, using multiple classification schemes. In total, the factor analysis reduced the 143 constructs provided by the participants in Design 1 to 32 factors, shown in Appendix 3 for each participant. Note that the analysis presented previously related to Participant 6 in Appendix 3, overall factor numbers 18 to 20.

B.2 List of statistical modules consulted for the Pilot study

Table B.5: The statistics modules used to clasify the types of learning outcomes in Section 3.4.2

N	Domain(s)	Level	Credits	Teaching
1	Geography	4	10	2h contact time/week and 1h lectures
2	Engineering	5	5	24x1h lectures
3	Psychology	4	10	2h contact time/week including lectures & tutorials
4	Psychology	4	10	2h contact time/week including lectures
5	Psychology	4	20	12x1h lectures
6	Psychology	5	10	10x2h lectures
7	Sociology, Psychology, Human Biology	4	10	14x1h lectures
8	Psychology	5	10	22x1h lectures
9	Psychology	5	10	11x1h lectures
10	Psychology	6	20	11 x 2h lectures
11	Psychology	4	10	10 x 1h weekly lectures
12	Engineering	4	10	11 x 2h lectures
13	Mathematics	4	10	24 x 1h lectures
14	Business & Economics	4	10	2h lectures per week
15	Business & Economics	4	20	3h lectures per week
16	Business & Economics	4	20	3h lectures per week
17	Business & Economics	4	10	2h lectures per week
18	Business & Economics	4	10	2h lectures per week

B.3 Interview data analysis (Pilot 1)

In this Appendix, in **Table B.6** I include the list of repertory grid elements with the constructs that explain them based on the SVD plots of all pairs of factors (NA means that there were no constructs close to an element) for each participant. Next, **Table B.7** includes the Factor analysis values for repertory grid interviews (Pilot 1, 32 Factors).

Table B.6: List of elements and constructs using Factor analysis (9 participants)

N	Participant	Element	Constructs (relating to element)
1	1	creativity	might not need to use advanced maths; not always required; most demanding
2	1	good report writing skills	needed to analyse and present complex reports; least important; low level
3	1	confidence to apply math to different areas	not necessarily applicable to teams; not necessarily applicable; better to have strong knowledge so better equipped to apply it to different areas; need decision making skills to decide when and how to apply mathematical techniques; most demanding at introductory level
4	1	ability to handle data in spreadsheets	needed to analyse and present complex reports; demanding introductory level; numbers might not apply
5	1	Strong numeracy skills	needed to analyse and present complex reports; least year 3; needed to succeed with maths problems; better to have strong knowledge so better equipped to apply it to different areas; need decision making skills to decide when and how to apply mathematical techniques
6	1	ability to lead and be lead	flexibility when working within teams; not always required; might not work with others; only applies if within a team
7	1	able to handle team work	flexibility when working within teams; not always required
8	1	Ethics	creativity to have broad ideas but constrained by ethics
9	1	quick decision making skills	needed to succeed with maths problems; numbers might not apply; needed to analyse and present complex reports;
10	2	ability to write a report using statistics	ability to interpret analysis, use statistics in practice, skill rather than knowledge, low level, understanding on intuitive level, knowledge of basic methods

Table B.6 continued from previous page

N	Participant	Element	Constructs (relating to element)
11	2	ability to estimate econometric models and interpret them	ability to use and interpret statistics (skill, technique)
12	2	understanding of classical view regression model	knowledge
13	2	knowledge of time series models	knowledge of advanced methods and statistical theory
14	2	ability to understand and read statistical papers in journals	ability to interpret and use statistics, low level
15	2	knowledge of statistical inference procedures	knowledge
16	2	knowledge of advanced cross-section methods, e.g. panel data	interpret analysis
17	2	understanding of problems with CLRM, e.g. serial correlation, heteroscedasticity	understanding statistical theory
18	2	knowledge of probability theory	knowledge
19	3	statistical inference	econometric theory uses statistical inference
20	3	exploring data	straight forward, low level
21	3	graphical displays	straight-forward procedures, produced by software, stat inference not required, exploratory technique require software
22	3	econometric theory	thrown up by regression and correlation
23	3	time series techniques	straight forward problems, no econometric problems
24	3	econometric problems	uses statistical inference
25	3	statistical software	techniques require software, does not require statistical inference
26	3	regression and correlation	underlie time series
27	3	probability concepts	straight forward, least years 1 and 2, low level
28	4	how to do a t-test	tedious detail; practical application
29	4	what is a p-value	most demanding; meaning of data; population; output; inferential; most demanding; high level
30	4	ways of describing distributions (mean, SD, mode, etc)	structure of data; sample; input; descriptive; least demanding; low level; practical application
31	4	assumptions of various different statistical tests	NA

Table B.6 continued from previous page

N	Participant	Element	Constructs (relating to element)
32	4	understanding difference between main effects and interactions in an ANOVA	NA
33	4	difference between inferential and descriptive statistics	conceptual understanding; important concepts
34	4	difference between dependent and independent data	structure of data; sample; input; descriptive; least demanding; low level
35	4	relationship between effect size, power and significance	meaning of data; population; output; inferential; high level; conceptual understanding
36	4	how to interpret results of statistical tests	NA
37	5	associations	regression needed to understand relationships;
38	5	regression	significance testing not important; randomisation not important; high level; most year 1
39	5	multivariate statistics	high level; significance testing not important; most demanding
40	5	bi-variate relationships	low level, least demanding; need descriptive statistics to understand associations
41	5	numerical theory	need descriptive statistics to understand association; understanding numerical theory is fundamental to understanding associations
42	5	descriptive statistics	low level, least demanding; need descriptive statistics to understand associations
43	5	significance testing	difference between variables identified by significance testing; no requirement to understand relationships; understanding numerical theory fundamental to understanding associations
44	5	population & sample	low level; good descriptive statistics result from randomisation
45	5	randomisation	good descriptive statistics result from randomisation; low level; no requirement to understand relationships
46	6	appreciate different sampling methods - when each might be used	collecting data; sampling; least demanding year 1, 2

Table B.6 continued from previous page

N	Participant	Element	Constructs (relating to element)
47	6	understand confidence intervals	method of analysis, most demanding year 1, 2
48	6	have a basic appreciation of multivariate methods	quantitative description, complex method, inferential
49	6	be able to describe data using descriptive stats and appreciate graphical methods	low level, simple methods, qualitative description and interpretation
50	6	be critical when interpreting data and stats given to students (e.g. in media)	interpreting data, results, what they mean
51	6	be able to use simple inferential statistics	applying technique, method, design; analytical method, mechanics of doing it; techniques; method/design
52	6	be able to interpret graphs and tables and draw reasonable conclusions	interpret results and what they mean; most year 1, 2
53	6	be able to apply regression analysis	applying technique, method, design
54	6	be able to apply correlation methods and interpret results	qualitative interpretation, qualitative description, simple methods, descriptive
55	7	team work	not essential in statisticians, social skills
56	7	t-tests	essential of good statisticians, value on the job market high, least demanding in year 2 and 3
57	7	one sample tests	misunderstood by students, low value in the job market, high level, purely statistical, logic of research process
58	7	relationship between effect size, power and significance	misunderstood by students, low value in the job market, high level
59	7	statistical inference	essential of good statisticians
60	7	chi-squared	essential of good statisticians
61	7	ability to understand and read quantitative research papers in journals	communication skills, non-statistical skill, value in the job market high
62	7	difference between descriptive statistics and inferential	essential of good statisticians, value on the job market high, least demanding
63	7	ability to write a report using statistics	communication skills, non-statistical skill, value in the job market high
64	8	dangers of 'blind' statistics	application-based concept, year 3
65	8	understanding linear and non-linear	one component of maths that defines one section/specific concept; specific requirement, functional

Table B.6 continued from previous page

N	Participant	Element	Constructs (relating to element)
66	8	decision making	global concepts; why we do what we do vs. what we do; application of the skill in the real world context
67	8	relationship vs. difference	related to linear algebra; specific
68	8	confidence with material	education-based concept, can be learnt; least year 3
69	8	philosophical underpinning	education-based concept, can be learnt; least year 3
70	8	numeracy	education-based concept, can be learnt; least year 3; specific
71	8	functional application	application of the skill in the real-world context; global concept
72	8	report writing	application of the skill in the real-world context; global concept
73	9	statistical considerations need to be included at all stages of a study	low, easy concepts for students to grasp; knowledge of concepts, right/wrong answers; concept used to communicate statistical procedures, findings and recommendations
74	9	know meanings of results of tests, such as p-values and confidence intervals	easy to use textbooks to guide the procedure; pure, theoretical statistics
75	9	understand principles of best statistical practice in terms of interpreting and reporting results	difficult, challenging concepts for students; high level; what test to run and how to interpret the results
76	9	facility with statistical software	easy concepts; applied statistics
77	9	understanding the theoretical underpinnings of tests to understand their assumptions/ constraints	difficult, challenging concepts for students; high; pure, theoretical statistics
78	9	report writing in statistical context	high level; difficult, challenging concepts for students; skills or knowledge that can only be learnt by doing, hands on experience is essential
79	9	relationship between power, significance, effect size and sampling size	easy to use textbooks to guide the procedure; easy concepts for students to grasp
80	9	understanding how descriptive statistics/graphical methods complement formal inference procedures	skills or knowledge that can only be learnt by doing, hands on experience is essential

Table B.6 continued from previous page

N	Participant	Element	Constructs (relating to element)
81	9	critical evaluation of a study	concept used to communicate statistical procedures, findings and recommendations

Table B.7: Repertory grid Factor analysis values (Design 1)

Factor N	Participant	Factor	E-value	Var%	Cum%	Within participant Factors +	Within participant Factors -
1	1	1	4.45	27.81	27.81	creativity to have broad ideas but constrained by ethics	required to succeed with quantitative (maths) problems; least year 3; needed to analyse and present complex reports
2	1	2	3.48	21.78	49.59	flexibility when working with teams; need to be creative to take on different roles; need decision making skills in order to lead	[NA]
3	1	3	3.47	21.72	71.31	need decision making skills in order to decide when and how to apply mathematical techniques; better to have strong knowledge so better equipped to apply it to different areas	[NA]
4	1	4	3.08	19.23	90.53	low level, least required in year 1, 2, 3	[NA]
5	2	1	10.38	61.03	70.6	knowledge and understanding of advanced techniques, statistical theory, statistical models, intuitive understanding of statistical theory (e.g. econometric models)	skills in using or interpreting statistics, ability to use and interpret statistics, skill rather than knowledge (e.g. report writing)
6	2	2	5.35	31.47	92.5	knowledge of basic methods, understanding on intuitive level, low level, least demanding across years of study	understanding of statistical theory using mathematics, knowledge of advanced methods, high level
7	3	1	5.39	31.71	31.71	econometric theory uses statistical inference, no software required (e.g. statistical inference, probability concepts, econometric theory)	graphical displays produced by software, graphical displays important for identifying time series models, exploratory techniques require statistical software (e.g. exploring data)
8	3	2	3.93	23.14	54.85	no econometric problems, straight forward procedures (e.g. time series techniques) [NA]	using regression and correlation may lead to econometric problems, need software to calculate regression and correlation

Table B.7 continued from previous page

Factor N	Participant	Factor	E-value	Var%	Cum%	Within participant Factors +	Within participant Factors -
9	3	3	3.34	19.67	74.52	high level	low level
10	3	4	2.14	12.59	87.11	regression and correlation underlie time series techniques, concepts that examine time series properties of data	not about time series [NA]
11	3	5	1.72	10.11	97.23	Not required in regression and correlations, most Year 3 (e.g. time series techniques) [NA]	statistical inference required to interpret regression/correlation, least year 3
12	4	1	4.07	33.91	33.91	low level	high level [NA]
13	4	2	3.76	31.32	65.22	structure of data; sample; descriptive; input	output; population; meaning of data; inferential [NA]
14	4	3	2.4	20.01	85.23	tedious detail; practical application;	important concepts; important conceptual distinctions; conceptual understanding [NA]
15	5	1	4.75	27.95	27.95	description of data relies on theory; multivariate not required; concepts not dependent on bi-variate relations; no regression testing relevant	associations stem from bi-variate relationships; multivariate needed for association; bivariate relationships explained by regression; multivariate stats needed to understand associations [NA]
16	5	2	4.66	27.43	55.39	need descriptive stats to understand association; understanding numerical theory fundamental to understanding associations; low level; least year 1, 2, 3	no need; not dependent on understanding; high level; most year 1, 2, 3[NA]
17	5	3	4.05	23.8	79.19	good descriptive stats result from randomisation; difference between variables identified by significance testing; understanding population and sample required for descriptive statistics; no requirement to understand relationships	regression needed to understand relationships; significance testing not important; randomisation not important [NA]

Table B.7 continued from previous page

Factor N	Participant	Factor	E-value	Var%	Cum%	Within participant Factors +	Within participant Factors -
18	6	1	6.54	38.48	38.48	interpretation of data and stats - what it means; interpreting results and what they mean; being critical	techniques; method, design, analytical method-mechanics of doing it, low level (least demanding across years of study)
19	6	2	6.06	35.64	74.13	multivariate methods are interpreted using confidence intervals; methods of analysis; inferential; quantitative description; more complex methods	qualitative description and interpretation; simple methods; descriptive; sampling, low level
20	6	3	2.86	16.81	90.94	collecting data, sampling (e.g. when sampling methods may be used)	analysing and interpreting data, understanding data
21	7	1	4.25	24.99	24.99	purely statistical, essential of good statisticians, logic of research process that students need, high level (e.g. statistical techniques and procedures)	communication and interpretation skills, social skills, organisational skills, low level, not essential in statisticians (e.g. team work; ability to write a report using statistics)
22	7	2	4.08	24	48.99	low value in the job market, misunderstood by students, Most demanding year 2 and year 3 (e.g. relationship between effect size, power and significance)	easier to grasp, crucial to the curriculum, value on the job market high, low level, least demanding in years 2 and 3 (e.g. descriptive and inferential statistics)
23	7	3	3.95	23.23	72.21	not essay style (skill), less requirement of English language skills, not related to report writing/structure least demanding in year 1 (e.g. statistical techniques)	report structuring, language skills required, essay skills, most year 1 (e.g. ability to write a report using statistics, ability to understand and read quantitative research in journals)
24	7	4	2.59	15.26	87.47	theoretical underpinning of statistics, presence within lectures high, repeated	hand computations required

Table B.7 continued from previous page

Factor N	Participant	Factor	E-value	Var%	Cum%	Within participant Factors +	Within participant Factors -
25	8	1	4.93	29.01	29.01	education-based concepts that can be learnt; relationship and linear algebra are related, different way of conceptualising; specific requirement functional, different ways of achieving the same result; implicit knowledge and understanding, least demanding in year 2 and 3 (e.g. statistical techniques, knowledge, numeracy, philosophical underpinning)	application-based concept; global concepts, usage of the skill in the real-world context; most demanding in year 3 (e.g. report writing, functional application, decision making, dangers of blind statistics)
26	8	2	4.54	26.68	55.69	emotional response to the material; most year 1 (e.g. confidence with material)	knowledge-driven response to the material, analytical components; application, presentation of material, report writing, least year 1 (e.g. report writing)
27	8	3	2.16	12.68	68.37	core knowledge that one should have of statistics	making a decision is based on interactions and their nature (e.g. z-scores)
28	8	4	2.01	11.82	80.18	low level; generic skill	understanding of the concept (relationship vs. difference) leads to good decision making; high level
29	8	5	1.8	10.58	90.76	why we do what we do vs. what we do; global, harder to define concepts, needed for employment	one component of maths that defines one section, specific concept
30	9	1	3.72	46.46	46.46	easy concepts for students to grasp; easy to use textbook to guide the procedure; low demand	judgment, decisions about what test to run and how to interpret the results; difficult, challenging concepts for students; skills or knowledge that can only be learnt by doing, hands on experience is essential; high
31	9	2	1.77	22.08	68.54	best practice in statistical analysis, broad statistical ideas; pure, theoretical statistics	applied statistics;

Table B.7 continued from previous page

Factor N	Participant	Factor	E-value	Var%	Cum%	Within participant Factors +	Within participant Factors -
32	9	3	1.45	18.08	86.61	concept used to communicate statistical procedures, findings and recommendations (in writing, as a consultant, orally);	concept used to carry out technical procedures that are not necessarily reported [NA]

B.4 Pilot 1 and 2 emerging themes

Table B.8: Emerging themes and sub-themes in the repertory grid data

Theme description	Sub-theme code	Sub-theme description	Factors	Frequency
1. Statistical knowledge and understanding vs. application of statistical methods and techniques	1.1	Knowledge of theory, conceptual knowledge and understanding vs. applying techniques in the real world context	5, 14, 25, 30, 31	5
	1.2	Knowledge and understanding of statistical methods, techniques and procedures vs. statistical thinking and reasoning (evaluate or reflect on statistical procedures, integrate and synthesise the context knowledge with statistical knowledge, work with models to draw inferences from data and communicate a statistical argument about the real situation)	8, 13, 18, 23, 25, 29, 32	7
	1.3	Statistical (exploratory techniques, inferential statistics, graphical representations, probability, estimation, significance tests, ANOVA, regression, correlation, bi-or multivariate methods, time series analysis) vs. non-statistical outcomes (mathematics, confidence with statistics, team work, logical thinking, communication skills – report writing, critical thinking, creativity, using resources – information technology)	1, 2, 3, 6, 7, 8, 21, 23, 25, 28	10
2. Statistical process components	2.1	Collecting data vs. analysing and interpreting data	20	1
	2.2	Analysing data (applying techniques) vs. interpreting data/models (communicating with data)	5, 7, 18, 19, 21, 23, 32	7

Table B.8 continued from previous page

Theme description	Sub-theme code	Sub-theme description	Factors	Frequency
3. Student learning and curriculum design	3.1	Easy (basic, straightforward, intuitive, understood, learnt from textbooks) vs. difficult (demanding, challenging, complex, advanced, hard, misunderstood, mathematically-based, learnt by doing) ideas/concepts/methods	5, 6, 8, 19, 22, 25, 29, 30	8
	3.2	Knowledge or skill important (relevant, essential, core) vs. not important (not essential) to learning particular aspects of statistics	7, 14, 17, 21, 27, 30	6
	3.3	Need to cover in the module (presence within lectures, high value in employment) vs. don't need to cover (low value in employment or in carrying out particular statistical analysis)	10, 11, 15, 16, 17, 19, 22, 24, 29	9

B.5 Pilot study coding system

Interviews from the Pilot 1 and 2 analysis using quantitative and qualitative data analysis was produced using the emerging topics, the constructs and elements provided to participants and the repertory grid data. All interview summaries were coded, although I produced a more detailed analysis of the interviews provided by the two lecturers who participated in the main study analysis (as described in Chapter 4). I assigned themes or statements to various descriptive and analytical categories. The code system is made up of themes, classified as *parent codes* in Transana, which contained *codes* (sub-themes or categories) within them. In total there were ten parent codes and sixty five codes in the code system. The code system is summarised in Table B.9. Statements within the interview summaries were qualitatively reviewed and coded to as many parent codes and codes as applicable.

Table B.9: Repertory grid code system (interview design II)

Theme (parent code)	Sub-theme (code)
Curricular	Assessment
	Lecturer beliefs about what students should be taught
	Purpose of module
Didactic	Student engagement
	Student learning
	Student short/long term goals
	Teaching approaches
	Value of experience on the job
Institutional	Control over module
	Institutional pressures
	Lecturer's experience
	Outside department
	Within department
Pedagogy	Easy for students
	Hard for students
	Lecturer beliefs about statistics
	Lecturer beliefs about students
Skills	Calculations
	Communication
	Contextual knowledge
	Descriptive statistics
	Gosset
	Hypothesis testing
	Interpretations
	Intuition
	Probability - not statistics
	Problem solving
	Software
	Speak the language of maths
	Statistical analysis procedures
	Statistical literacy
	Translate real world context into maths
Statistical analysis	Creative
	Interpretation
	Knowing what it means

Continued on the next page...

Theme (parent code)	Sub-theme (code)
	Maths language
	Practical
	Process
	Routine applications of techniques
	Technique
	Terminology
Statistical applications	Concrete
	Context of what is being measured
	Decision making process
	Experimental
	Narrative
	Non-statistical
	Particular skills
	Statistics needs to make sense in context
Statistical theory	Abstract
	Equation
	Mathematics/statistics
	Next step in mathematics
	Tool
	Variables
Relevance	Context
	Field of study
	Statistical
Other (KRG Technique)	Build scale
	Construct
	Did not rate cards/constructs
	Explain technique
	Feedback to participant
	Rate cards
	Value of ratings

C | Main study

C.1 Interview guide

Background

For how many years have you run the modules?
What is your academic background?
What is your involvement in developing the module resources?

Teaching on the module

As you look back on the module you taught this term, are there any lectures that stand out?
Could I ask you to describe the most important ways to teach statistics?
Could I ask you to describe how you designed/came up with the tasks and examples you used (a task is an example, exercise, lab sheet, question)?
How did you produce the tasks you showed students in your lecture?
What are the links to other modules or to what students will do with the material later on?

Learning resources

Why did you select the different resources to present to students?
Why did you select particular context of statistical tasks? (context is a background scenario, research background, dataset)
What do you think about the context of these tasks?
How do you think the tasks are relevant to the students?
How do they compare to other tasks you may have used or are used on other modules?

Student evaluation and assessment

What are the links between the tasks and examples and the way you assessed the module?
Why did you assess your students and what do you consider assessment to be?

Mathematical and statistical content

In your view, do you think the formulas help students understand the statistical tests?
In your view, how important is mathematics to students' understanding of the content of your module?
How important is statistics to these students' academic and professional development?
What is the role of software in teaching and learning statistics?
What is your students' previous experience with maths? How did they cope with it?

Conclusions

What would you do differently with this module?
Is there something else that occurred to you during this interview?
Is there something else you would like to ask me?

C.2 The content of Modules A & B

In Table C.1, I summarise the lectures I observed for the modules included in the data analysis.

Table C.1: Summary of observations from Modules A and B

	Module A		Module B	
Lecture	Topic	Context	Topic	Tasks in Context
1	Inferential statistics	Students' attention during stats lectures	Probability	N/A
	Descriptive statistics			
	Inferential statistics	SPSS output comparing males/females		
	Standard deviation			
	Types of variables			
	Error and significance	Guilty verdict		
2	t-test 1	Gosset	Normal distribution	Throw dice
		Stats exam results (1)		Examination test scores
		Drug A vs. drug B		Human height
		Driving and alcohol		Strip of wire
				Assembly component
3	T-test 2	Diet	Descriptive statistics, Sampling	Employee days off
		Stats exam (2)		THE University league tables
		Family therapy (1)		Volt cells
				Optical cutting
				Brick cutting
				Airport arrivals
4	ANOVA 1	Family therapy (2)	Hypothesis testing 1	Better driver (1)
		Masking witness		Dishwasher powder
		Death of parent and depression		Car assembly plant
		Age and memory		Fuel formulation
				Rugby scores
5	ANOVA 2 (Two-way)	Family therapy (3)	Hypothesis testing 2	Paint factory
		Teaching methods		Rubber material
		Biscuits and age		Iron in compounds
				Exam marks
				Marmite and gender
				Machine breakdown
6	ANOVA 3 (Two-way)	Context-dependent memory (Gooden & Baddeley, 1975)	Contingency tables 2	Darts and handedness
		Student revision conditions		Sound track preferences

Finished on next page

Table C.1: Summary of observations from Modules A and B (*Finished*)

	Module A		Module B	
Lecture	Topic	Context	Topic	Tasks in Context
7	ANOVA (Revision)	Family therapy (4)	Regression	Heart disease and blood pressure Car acceleration and tyre diameter Years of education and earnings Temperature and electronic components Driver's age and sign legibility distance Leaning Tower of Pisa Blood sugar levels and diabetes
8	ANOVA (Post Hoc)	SATs and revision length	Correlation	Driver's age and sign legibility distance Sewage in the Lake Gas-Electricity bill Mother-daughter heights Purity of oxygen
9	Review	Revision worksheet	ANOVA	Better driver (2) Alloy spacers Four trucks Happiness, political party and football team Students' self-esteem and academic success
10	Revision	Selected response questions	Experimental design	Tuition and student learning CW experiement Inglis & Mejia-Ramos, 2009
11	Mock test	Selected response questions	Revision	None

D | Examples of Teaching Resources

In this section, I provide examples of lecturing resources used on Modules A and B.

D.1 Module A

Example 1 (Module A, Practical tutorial task)
Below are data from a study on the effect of family therapy as a treatment for anorexia. In this study 17 adolescents were given treatment. They were weighed before and after. The following scores are the weights in pounds.

No therapy	CBT	Family therapy
−0.5	1.7	11.4
−9.3	0.7	11.0
−5.4	−0.1	5.5
12.3	−0.7	9.4
−2	−3.5	13.6
−10.2	14.9	−2.9
−12.2	3.5	−0.1
11.6	17.1	7.4
−7.1	−7.6	21.5
6.2	1.6	−5.3
−0.2	11.7	−3.8
−9.2	6.1	13.4
8.3	1.1	13.1
3.3	−4	9.0
11.3	20.9	3.9
0	−9.1	5.7
−1	2.1	10.7
−10.6	−1.4	
−4.6	1.4	
−6.7	−0.3	
2.8	−3.7	
0.3	−0.8	
1.8	2.4	
3.7	12.6	
15.9	1.9	
−10.2	3.9	
	0.1	
	15.4	
	−0.7	

To run a related groups $t - test$ is very simple. You should do the following:

1. Enter the data in columns. Don't forget to label the variables clearly.
2. Click on Analyze>Compare Means>Paired Samples t test

3. A dialogue box comes up. It will list your variables. Click on both the Before and After variables and then the arrow to move them into the right hand box. They come up as one pair. Click on OK.

Now consider the output.

4 The first box gives you some useful descriptive stats for each set of scores (Ns, means and standard deviations).

5 The second box gives you the correlation between the two sets of scores (ignore this for the moment).

6 The third box tells you the most important stuff with respect to the paired t test. In particular it gives you the t value, the df and the p value – all of which are required for the report.

Activities

Paste your output into your word file and answer the following:

1. What is the null hypothesis for this study?
2. Does the $t - test$ allow you to reject it or not?
3. Do a scatterplot of the two sets of scores – give your graph a title, e.g. 'Fig 1 Graph showing the relationship between family therapy and weight gain in anorexic girls'. Fit the regression line only if you think it is a significant result. Paste the scatterplot and correlation output (above) into your word file. What does this tell you? Would you expect there to be a correlation?
4. Is the test significant? Represent with appropriate notation, i.e. $t(df) =?$, $p > 0.05$ or $p < 0.05$. Also report your correlation $r(N) =?$, $p >$ or $p < 0.05$
5. Produce a bar graph of the two sets of scores. Under Graphs/Legacy Dialogs select Bar and then 'Summaries of separate variables' and 'Simple' then 'Define'. Transfer both sets of scores to the 'Bars Represent' box and give the graph a title, e.g. Fig. 1 Graph showing the effect of. . . . Click OK. What does the graph tell you?
6. Why did you need to do the t-test as well as the correlation? What more did the t test tell you?
7. Is this a matched pairs design or a repeated measures design?

Example 2 (Module A, t-tests, Lecture 2)

Statistics Exam example data outputs and plots comparing males and females version A data – descriptive statistics

Group Statistics

	male and female groups	N	Mean	Std. Deviation	Std. Error Mean
stastscore	female	20	11.0000	1.77705	.39736
	male	20	6.2000	1.50787	.33717

version B data- descriptive statistics

Group Statistics

	males and females	N	Mean	Std. Deviation	Std. Error Mean
scores on a stats test	male	20	6.0500	7.49368	1.67564
	female	20	11.0000	8.31612	1.85954

(a) Two computer-generated analysis outputs for version A and version B data

t(38)= -8.90, p<0.001

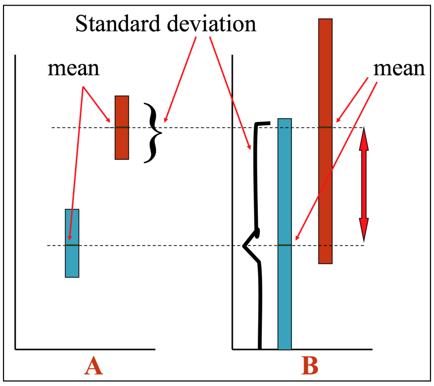
Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
stats test scores	1.00	20	6.0000	1.7770	.3974
	2.00	20	11.0000	1.7770	.3974

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
stats test scores	Equal variances assumed	.000	1.000	-8.898	38	.000	-5.0000	.5620	-6.1376	-3.8624
	Equal variances not assumed			-8.898	38.000	.000	-5.0000	.5620	-6.1376	-3.8624

(b) SPSS output of the independent t-test analysis



(c) Two plots for Version A and Version B data

Example 3 (Module A, ANOVA, Lecture 7)
Slide 1: SPSS Exercise

- 1 way ANOVAs
  - "This is a one-way ANOVA with 3 levels, W, X, and Y. Scores are measured on Z"
- 2 way ANOVAs
  - "This is a two way ANOVA with a $3 * 4$ between subjects design, Factor A relates to (whatever) and has 3 levels, X, Y and Z, Factor B relates to (the other thing) and has 4 levels, A,B,C and D"
- Report 3 F Values!

Slide 2

- Is there a significant difference?
- Answer the question, + report appropriately e.g. giving F/t, df, p values
- Visual displays
  - Graphs with appropriately numbered label, e.g. Fig. 1, 2, 3 etc throughout your document.

Slide 3: Tables

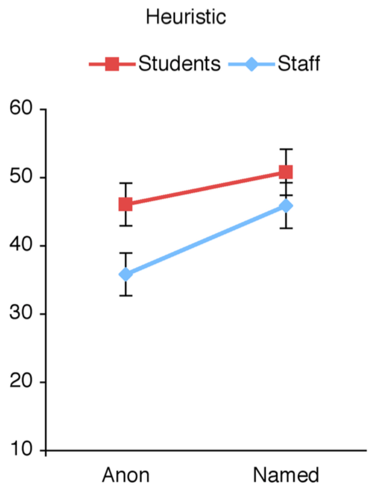
For each study, paste in your output tables and give each one a label – e.g. 'Table 1, 2, 3 etc, Table showing. . . '
...THEN do a description! Make sure all descriptions illustrate your understanding of the findings as clearly as possible
Golden Rule: More is more!

D.2 Module B

Example 4 (Module B, Experimental design, Lecture 10)
Slide 1
Inglis and Mejia-Ramos (2009) were interested in exploring the extent to which an authority figure influences how mathematics students and mathematics lecturers react to mathematical arguments.
Research Question: Does knowing that an argument was written by an expert mathematician make it more persuasive for students and lecturers?

- The research design
- Reliability
- Validity (esp. construct validity and external validity: both of which were criticised by reviewers)

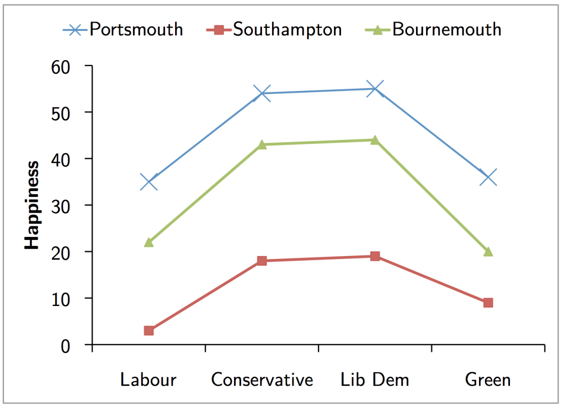
Slide 2
Participants
Undergraduate mathematics students (from three "highly ranked" UK universities; and research-active mathematicians from universities throughout UK and Australia.
Method
Internet study. Participants were asked to read a claim, an argument in favour of that claim, and then were asked to "say to what extent [they were] persuaded by it". Responding via a sliding 0-100 scale.
Groups were randomly split into two conditions: those in the experimental condition were told the author of the argument's name, and those in the control condition were not.
Author was Professor Tim Gowers (University of Cambridge).
Slide 3
Screen shot of on-line survey
Slide 4



Slide 5

1. Draw a diagram to explain the design of the experiment.
2. Do you think that the design is reliable?
3. What do you think the issues were that the reviewers raised to do with: construct validity? external validity?

Example 5 (Module B, ANOVA, Lecture 9)
Question
A study looked at whether happiness (measured as a %) was related to how an individual votes, and which football club they support:



- What is the design? (e.g."2 × 3 factorial")
- Will the main effects of football-club and voting-behaviour be significant?
- Will the football-club × voting-behaviour interaction be significant?

Describe the findings of the study in words.

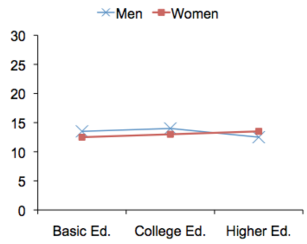
Example 6 (Module B, ANOVA, Lecture 9)
Worked example - equal sample sizes
Slide 1
The better driver example

If we want to answer the question: Who drives better? Drivers with basic education, college education or higher education?
and
Does the gender of the driver make any difference?

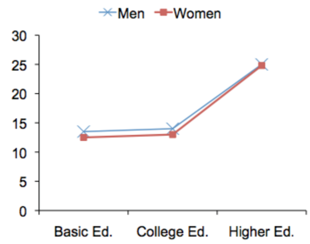
Slide 2
Two way
Using a two-way ANOVA we can interrogate the influence of more than one factor. For example, we may be interested in both how education affects driving, and in how drivers sex affects driving.
Suppose we categorise each one of "education" samples into two groups: "men" and "women". So we have a design where we want to know how two factors influence how good driver a person is:

- Education: basic, college, higher (three levels)
- Sex: Men, Women (two levels)

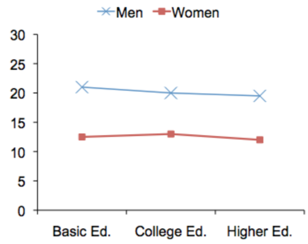
This is called a 3×2 factorial design.
Slides 3-8



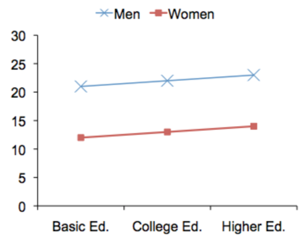
(a) Two way - no effects



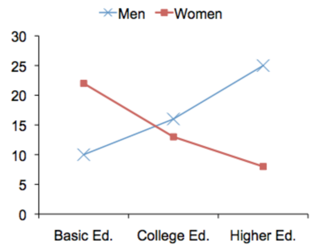
(b) Two way - main effect of education



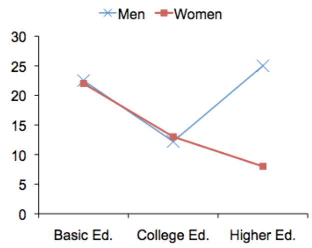
(c) Two way - main effect of sex



(d) Two way - main effect of sex and education



(e) Two way - interaction



(f) Two way - interaction

Figure D.2: Module B ANOVA, Better Driver example

Slides 9
An interaction occurs when the effect of one factor changes depending on the level of another factor. Graphically, this is reflected by the lines on a profile plot (of means) not being vertically shifted copies of each other.

Example 7 (Module B, A worked example)
Slide 1, The problem

The table below shows the respective heights of a sample of 12 mothers and their adult daughters in inches (in).

Height of mother	Height of daughter
65	68
63	66
67	68
64	65
68	69
62	66
70	68
66	65
68	71
67	67
69	68
71	70

1. Calculate the Pearson's correlation coefficient (r) for these data.
2. Is this correlation statistically significant (use $\alpha = 0.05$)?
3. Interpret the results of part (b) regarding the relationship between the heights of a mother and her daughter.
4. Fit a least squares linear regression model to predict the height of a daughter from the height of her mother.
5. If the height of a mother is 75in, predict the height of her adult daughter.

Slide 2, Answer A1 Calculate the Pearson's correlation coefficient (r) for these data.
Let x be mother's height and y be daughter's height.
First calculate the sums $\sum x = 800$, $\sum y = 811$, $\sum x^2 = 53,418$, $\sum y^2 = 54,849$ and $\sum xy = 54,107$. Also $n = 12$.
Then

$$\begin{aligned} r &= \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y\right)^2\right)}} \\ &= \frac{12 \times 54,107 - 800 \times 811}{\sqrt{(12 \times 53,418 - 800^2)(12 \times 54,849 - 811^2)}} \\ &= \frac{484}{\sqrt{1016 \times 467}} \approx 0.703. \end{aligned}$$

Slide 3, Answer A2

Is this correlation statistically significant (use $\alpha = 0.05$)?
We want to test $H_0 : r = 0$ against $H_1 : r \neq 0$ using $\alpha = 0.05$

The test statistic is

$$\begin{aligned} r_{\text{test}} &= \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}} \\ &= \frac{0.703 \times \sqrt{10}}{\sqrt{1 - 0.703^2}} \approx 3.126. \end{aligned}$$

The critical value is $t(n-2, \alpha/2) = t(10, 0.025) = 2.228$.
As $3.126 > 2.228$ we reject H_0 in favour of H_1 and conclude that there is a significant correlation between the heights of mother and daughter.
Slide 4, Answer A3

Interpret the results of part (b) regarding the relationship between the heights of a mother and her daughter.
The correlation coefficient r is positive and close to 1.
We can conclude that there is a moderately strong increasing linear relationship between the heights of mother and daughter.

Slide 5, Answer A4

Fit a least squares linear regression model to predict the height of a daughter from the height of her mother.
The regression line of y on x has equation $y = a + bx$, where

$$b = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n}\frac{\sum y}{n}}{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \frac{\frac{54,107}{12} - \frac{800}{11} \times \frac{811}{12}}{\frac{53,418}{12} - \left(\frac{800}{12}\right)^2} \approx 0.47622$$

Slide 6, Answer A5

If the height of a mother is 75in, predict the height of her adult daughter.
If $x = 75$ then $y = 35.8353 + 0.47622 \times 75 \approx 71.6$in.

Example 8 (Module B, ANOVA, Lecture 9)
Worked example - equal sample sizes
Slide 1
Four machines make alloy spacers for use in the assembly of a microlight aircraft. The spacers are supposed to be identical, but the four machines give rise to the following varied lengths in mm:

	Machine A	Machine B	Machine C	Machine D
	46	56	55	49
	54	55	51	53
	48	56	50	57
	46	60	51	60
	56	53	53	51
$\overline{x}$	50	56	52	54

We can work out the mean of means as $\overline{\overline{x}} = \frac{1}{4}(50 + 56 + 52 + 54) = 53$.
Slide 2
The variation between samples (or treatments or factors) can now be worked out (let a be the number of treatments):

$$
\begin{aligned}
S_{Tr}^2 &= \frac{1}{a-1}\sum_{i=A}^{D}\left(\overline{x}_i - \overline{\overline{x}}\right)^2 \\
&= \frac{1}{4-1}\left((50-53)^2 + (56-53)^2 + (52-53)^2 + (54-53)^2\right) \\
&= \frac{20}{3} = 6.67
\end{aligned}
$$

Slide 3

We can also work out the variation within samples (due to chance errors unrelated to the treatment):

A: $\sum(x-\overline{x}_A)^2 = (46-50)^2 + (54-50)^2 + (48-50)^2 + (46-50)^2 + (56-50)^2 = 88$

B: $\sum(x-\overline{x}_B)^2 = 26$

C: $\sum(x-\overline{x}_C)^2 = 16$

D: $\sum(x-\overline{x}_D)^2 = 80$

To get a pooled estimate of the total variance within samples (the variance due to errors), we use the formula (compare with a similar method for the independent samples t-test):

$$
\begin{aligned}
S_E^2 &= \frac{\sum(x-\overline{x}_A)^2 + \sum(x-\overline{x}_B)^2 + \sum(x-\overline{x}_C)^2 + \sum(x-\overline{x}_D)^2}{(n_A-1)+(n_B-1)+(n_C-1)+(n_D-1)} \\
&= \frac{88+26+16+80}{4+4+4+4} = 13.13
\end{aligned}
$$

Slide 4

We have now worked out the variation between samples, S_{Tr}^2 and the variation within samples S_E^2 and can start to ask which is bigger.

If the variance between-samples is bigger than the variance within-samples we can probably conclude that the treatment makes a difference. It can be shown that if H_0 is true then the ratio

$F = \dfrac{nS_{Tr}^2}{S_E^2}$

will be approximately 1. (n here is the number of observations per sample). This ends up being the test statistic and follows an F distribution with the following degrees of freedom:

- Numerator: Number of Samples − 1, $a-1$.
- Denominator: Total Sample Size − Number of Samples, $a \times n = a(n-1)$

Slide 5

Ratio:

$F = \dfrac{nS_{Tr}^2}{S_E^2} = \dfrac{5 \times 6.67}{13.13} = 2.54$

Degrees of freedom:

- Numerator: Number of Samples − 1 = 3.
- Denominator: Total Sample Size − Number of Samples = 4 × 5 − 4 = 16.

The critical value (using $\alpha = 0.05$) is $F(3, 16) = 3.24$. Since $2.54 < 3.24$ we cannot reject the null hypothesis. There is no evidence of a systematic difference between the groups. (The observed differences between the sample means are plausible under the null hypothesis that the underlying group means are the same.)

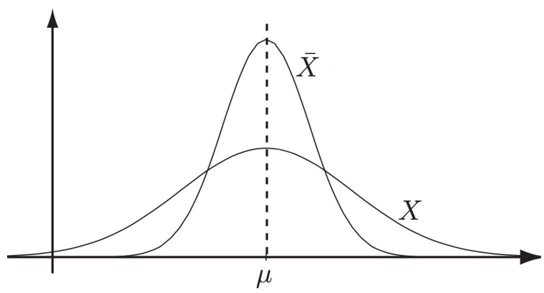
Example 9 (Module B, Central limit theorem, Sampling, Lecture 3)
Slide 1: Cenral Limit Theorem
If we take large samples of independent observations of size n with mean $\overline{X}$ from a population that has a mean μ and standard deviation σ then the distribution of sample means $\overline{x}$ is normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$
$\frac{\sigma}{\sqrt{n}}$ is the **standard error of the mean**.

`http://onlinestatbook.com/stat_sim/sampling_dist/index.html`

Slide 2: Some conditions
For smaller samples:

- if individual observations vary according to a normal distribution, then so will the sample means but
- if the observations don't vary normally, the sample means probably will, **given that the sample size is big enough**.

If the sample comes from a population which is not infinitely large, the observations are not independent. In this case, the standard error is given by the formula: $\dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$.
Slide 3: Distribution of the mean
Suppose $X \sim N(\mu, \sigma^2)$, then the means of samples from this distribution are distributed $\overline{X} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$.



So the bigger your sample, the smaller the variance of your sample mean, and the more accurate your understanding of the population mean.

Example 10 (Module B, Confidence intervals, Lecture 3)

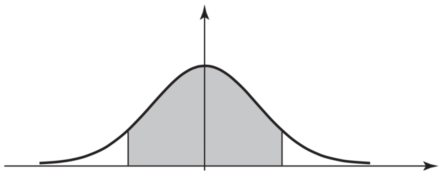
Slide 1: Confidence Intervals

A confidence interval gives an interval estimate rather than just a point estimate for parameter values. With a pre-determined level of confidence, the true value of the parameter lies within the calculated interval.

- We will use CIs for the mean and the standard deviation.
- Often calculate a 95% confidence interval so that 95% of sample estimates will capture the population value.
- Although the 95% level of confidence is the most common, other percentages can been used, e.g. 90% or 99%.
- A CI can be thought of as a range of plausible values for a population parameter, given some data.

Slide 2: Interval estimation for the Mean

Suppose $X \sim N(\mu, \sigma^2)$, the means of samples from this distribution are distributed $\overline{X} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$.



Slide 3: 95% confidence interval for the Mean

$$\mathbb{P}(-1.96 < U < 1.96) = 0.95$$
$$\text{(Why? Because } \mathbb{P}(U > 1.96) = 0.025\text{)}$$
$$\mathbb{P}\left(-1.96 < \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < 1.96\right) = 0.95$$
$$\mathbb{P}\left(\mu - 1.96\frac{\sigma}{\sqrt{n}} < \overline{x} < \mu + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$
$$\mathbb{P}\left(\overline{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

Slide 4: 99% confidence interval for the Mean

$$\mathbb{P}(-2.58 < U < 2.58) = 0.99$$
$$\text{(Why? Because } \mathbb{P}(U > 2.58) = 0.005\text{)}$$
$$\mathbb{P}\left(-2.58 < \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < 2.58\right) = 0.99$$
$$\mathbb{P}\left(\mu - 2.58\frac{\sigma}{\sqrt{n}} < \overline{x} < \mu + 2.58\frac{\sigma}{\sqrt{n}}\right) = 0.99$$

$$\mathbb{P}\left(\overline{x} - 2.58\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + 2.58\frac{\sigma}{\sqrt{n}}\right) = 0.99$$

Slide 5: In general...
So, in general
The $(1-\alpha)100\%$ confidence interval is equal to $\overline{x} \pm SE \times u_{\frac{\alpha}{2}}$ The value $u_{\frac{\alpha}{2}}$ is such that $\mathbb{P}(U > u_{\alpha/2}) = \frac{\alpha}{2}$ (from the normal distribution tables).
The above formula assumes that we know the population variance. In practice this is often not the case and we have to estimate the population variance from the sample.
For this we use the sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$.
If n is big enough ($n \geqslant 30$), the above CI formula remains valid.

Example 11 (Module B, General advice, Lecture 11)
Slide
When performing statistical tests make sure you:

- Explicitly state the null hypothesis, the alternative hypothesis, the alpha value, the test statistic and the critical value.
- Write a short conclusion of what the result of the test means, e.g. ”therefore we can reject the null hypothesis and conclude that $\mu_1 \neq \mu_2$; i.e. that machine 1 is performing differently to machine 2.”
- Write things out in detail: much easier to mark.
- Make sure you are aware of the assumptions of the different tests – you might be asked about this.
- Make sure you understand not just how to do the various tests, but also when they are appropriate and how to interpret the result/s.
- Show your working.