The Novice Mathematician's Encounter With *Mathematical Abstraction:* Tensions in *Concept-Image Construction* and *Formalisation*

Elena Nardi

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Linacre College

Trinity Term 1996

To my parents,
Nicole and Christophoros

Abstract

Elena Nardi **D.Phil.**
Linacre College Trinity Term 1996

The Novice Mathematician's Encounter *With Mathematical Abstraction:* Tensions in *Concept-Image Construction* and *Formalisation*

Mathematics is defined as an abstract way of thinking. Abstraction ranks among the least accessible mental activities. In an educational context the encounter with mathematical abstraction is the crucial step of the transition from informal school mathematics to the formalism of university mathematics. This transition is characterised by cognitive tensions. This study aimed at the *identification and exploration of the tensions in the novice mathematician's encounter with mathematical abstraction.*

For this purpose twenty first-year *mathematics undergraduates* were *observed* in their weekly *tutorials* in four Oxford Colleges during Michaelmas and Hilary Term of Year 1. Tutorials were tape-recorded and fieldnotes kept during observation. The students were also *interviewed* at the end of each term of observation. The recordings of the observed tutorials and the interviews were *transcribed* and submitted to an analytical process of *filtering out episodes that illuminate the novices' cognition.* An analytical framework consisting of cognitive and sociocultural theories on learning was applied on sets of episodes within the mathematical areas of *Foundational Analysis*, *Calculus*, *Linear Algebra* and *Group Theory*. This *topical analysis* was followed by a *cross-topical synthesis* of themes that were found to characterise the novices' cognition.

The novices' encounter with mathematical abstraction was described as a *personal meaning-construction process* and as an *enculturation process*: the new culture is Advanced Mathematics introduced by an expert, the tutor. The novices' *interaction with the new concept definitions* was obstructed by their *unstable previous knowledge. Concept image construction* was described as a *construction of meaningful metaphors* and an *exploration of the 'raison-d'-être'* of the new concepts and the new reasoning and was characterised by the *tension* between the *Informal/Intuitive/Verbal* and the *Formal/Abstract/Symbolic* — which was discussed in terms of *semantics* and *reasoning*. The novices were in *difficulty with the mechanics of formal mathematical reasoning* as well as with applying these mechanics in a contextualised manner. This *decontextualised behaviour* was linked to the *fragility of their knowledge with regard to the nature of rigour in formal mathematics.*

Acknowledgements

I have always been fascinated by the mental journey inherent in the simplest mathematical thought. By allowing me to share some of their journeys, the students and the tutors that participated in this study — whose names I have promised not to reveal but I will never forget — helped me in this attempt to translate my overwhelmed gaze at mathematical cognition into a coherent discourse. For this I am profoundly grateful.

This work was completed under the supervision of Dr Barbara Jaworski whom I deeply thank. Throughout these years I felt constantly that I could wish for no more *challenge* and *support* than those provided by Barbara. Her formative influence on my way of thinking about research will continue well beyond the completion of this thesis.

A D.Phil. is, I think, an all-consuming, mind-stretching activity the intensity of which would have been unbearable without the affective presence of empathetic friends. I feel immense gratitude both to those who endured painstakingly long discussions about work and to those who blissfully insisted on distracting me from work. I am especially grateful to my friends in Greece for tolerating our holidays and travelling all these years revolving around the various stages of this study — data collection / transcribing / analysis /... ; also for mastering the skill of sustaining a friendship through long-distance phonecalls.

This piece of research was partly supported by the Economic And Social Research Council and by the British Federation of Women Graduates whom I sincerely thank. However it would have been simply impossible to start, continue and complete it without the support of my parents who, wholeheartedly, made the resources of the family available to me and to whom this work is dedicated: I shall always consider myself extremely fortunate to have had the opportunity — all due to them — to discover the seductive powers lying in the pursuit of knowledge.

CONTENTS

Chapter 3: The Pilot Study

Chapter 4: Data Collection

Chapter 5: Data Analysis

Chapter 6: Foundational Analysis

Chapter 7: Calculus

Chapter 8: Linear Algebra

Chapter 9: Group Theory

Chapter 10: Synthesis

Prologue:
A Chronological Account of the Study and a Guide Through Chapters 1-10

This study is embedded in the area of research on the Psychology of Mathematics Education currently known as Advanced Mathematical Thinking and is an inquiry on the novice mathematicians' cognition with regard to their encounter with mathematical abstraction in their first year of university studies. It is a qualitative piece of research and its scope covers a wide range of mathematical topics as well as of learning phenomena related to the cognition of advanced mathematics. The theoretical background of the study, regarding research on PME-AMT and Methodology, is presented in Chapters 1 and 2.

Work on the Pilot Study — presented in Chapter 3 — was initiated in Michaelmas 1992. The Pilot was carried out in parallel with the Educational Research Methodology Training for the Research Students of the Department of Educational Studies and lasted until Trinity Term 1993. During this time tutorials to ten first-year undergraduates in one Oxford College were observed. The analysis of the fielnotes kept during observation revealed the potential richness of tutorials as a source of evidence related to the cognition of advanced mathematics.

In the first two weeks of Michaelmas 1993, negotiations with the Mathematical Institute and several Colleges — tutors and students — led to the formation of the body of participants to the Main Study: four tutors and twenty first-year undergraduates. Observation and recording of the tutorials lasted two terms: Michaelmas 1993 and Hilary 1994. During this time the students were also interviewed twice. The procedures of Data Collection are presented in Chapter 4.

Analysis of the collected material started during Data Collection: generally it aimed at the extraction of learning episodes related to the novices' conceptual and reasoning difficulties in their encounter with mathematical abstraction in a range of mathematical topics and at a cross-topical synthesis on these difficulties. Transcribing of the recordings was completed during Trinity and Michaelmas 1994 and the analysis of the material was completed between HIlary 1995 and the submission of the thesis. The procedures of Data Analysis are presented in Chapter 5.

The part of data analysis presented here — and introduced in the Interlude between Chapter 5 and Chapter 6 — is grounded on a selection of learning episodes that are paradigmatical cases of the themes that emerged during analysis. The presentation — Chapters 6-9 — is topical: so Chapter 6 contains a selection of episodes from the area of Foundational Analysis, Chapter 7 from Calculus, Chapter 8 from Linear Algebra and Chapter 9 from Group Theory. The cross-topical synthesis of the findings is presented in Chapter 10 and in the Epilogue I indicate briefly how the themes that emerged in this study may lead to more focused research.

I note that in the Appendices — that are provided in a separate volume — I present Extracts from the data, as well as other auxiliary material, on which the analysis in Chapters 6-9 is based.

Due to the limitations of space in the thesis, the material presented here is only a part of the collected material: so, for instance, the material relevant to Topology — the fifth area of mathematics explored in this study in addition to the aforementioned four — has been left out. However, as mentioned above and explained in detail in the Methodology Chapters and the Interlude, the material presented here is representative of the collected material.

Chapter 1
A Review of Relevant Literature: *Philosophical* and *Psychological Assumptions*. Linking the Study with Current Research in *Advanced Mathematical Thinking*

Introduction

Mathematics, both as a way of knowing (Bishop 1991) — the activity of doing mathematics — and as a body of knowledge — the outcome of this activity — grows in a physical, a sociocultural and a self-referential context. By the first I mean the physical world which mathematics takes into account and explains; by the second I refer to the particular social and cultural habitat in which mathematical activity is embedded; and by the third I refer to the mathematical community of all individuals engaged in the creation and dissemination of mathematics. These three dimensions reflect, interdependently and not exclusively, the complex environment in which mathematics develops.

The learning of mathematics takes place also in this tripartite context. If a global understanding of the process of learning mathematics is sought, one ought to add to this socio-cultural context the individual nature of learning. The dialectic consideration of both contextual (environmental) and mental (psychological) aspects is then likely to illuminate the process of mathematical learning.

Education is an institutionalised form of learning. Hence it serves as the milieu within which learning takes place and in which learning can therefore be studied. In the dialectics suggested above, Education, as an institution, is an inextricable part of the learning environment and therefore, in an inquiry regarding learning, contextual considerations are important.

One implication from the above can be that no study of the learner's thinking processes can be undertaken that is void of the impact of epistemological, psychological and educational theoretical assumptions. In this chapter I thus present a declaration of the principles underlying the study which is located within the area of research on the Psychology of Mathematics Education recently known as Advanced Mathematical Thinking (PME-AMT; also a Working Group of the International Group for the Psychology of Mathematics Education since 1985).

PME-AMT is at a Kuhnian pre-paradigm stage (Kuhn 1962) or, in R.B. Davis' words (1989), is a not 'data-poor' but a 'metaphor-poor' field. The meaning of this will be elaborated in Chapter 2 but briefly it means that PME-AMT is still in search of unified theoretical frameworks, of explanatory systems within which its researchers can work equivocally and unambiguously. I quote from *The Structure of Scientific Revolutions* in order to illustrate what makes the declaration of principles in this Chapter necessary:

> *When the individual scientist can take a paradigm for granted, he need no longer, ..., attempt to build his field anew, starting from first principles and justifying the use of each concept introduced...The creative scientist can begin his research where it leaves off and thus concentrate exclusively upon the subtlest and more esoteric aspects of the...phenomena that concern his group.*
>
> (Kuhn 1962, p.19)

Exactly because a researcher in PME-AMT can currently 'take no paradigm for granted', in what follows, an account is given of the *Theoretical Background* of the study, i.e. a declaration of its underlying philosophical (Part I) and psychological (Part II) principles. Subsequently the study is embedded within current PME-AMT developments (Part III).

PART I The Principles of the Philosophy of Mathematics Education Espoused in this Study

According to Paul Ernest (1991) 'philosophical schools of thought have a direct bearing on educational issues'. More specifically, a researcher's influences on her approach to mathematics education research include her perspective on mathematical learning, namely on their beliefs on how mathematical knowledge grows as a discipline as well as within the individual learner. Here I present the views espoused in this study about the phylogenetic (mathematics as a discipline) as well as the ontogenetic (mathematics as learned by an individual, cognizing subject) growth of mathematical knowledge. Subsequently a resulting philosophy and psychology (Part II) of mathematics education is presented. I note that the purpose of this presentation is to highlight the philosophical views that relate to the formation of the theoretical underpinnings of the study: therefore it is concise and does not have the structure — or the aspirations — of a philosophical debate.

Ia. The Phylogenesis of Mathematical Knowledge: Fallibilism, Relativism and the Role of Language and Culture

Johann Von Neumann (1983) describes Hilbert's theory of mathematical activity as involving 'an internally closed procedure which operates according to fixed rules known to all mathematicians and which consists basically in constructing successively certain combinations of primitive symbols which are considered "correct" or "proved" '. This identification of mathematics with its formal, axiomatic abstractions is known as the formalist school of thought and has been the target of Imre Lakatos' fallibilist polemic in *Proofs and Refutations* (1976). Lakatos accuses formalism of disconnecting the history of mathematics from the philosophy of mathematics:

> *Formalism denies the status of mathematics to most of what has been commonly understood to be mathematics, and can say nothing about its growth. None of the 'creative' periods and hardly any of the 'critical' periods of mathematical theories would be admitted into the formalist heaven, where mathematical theories dwell like the seraphim, purged of all the impurities of earthly uncertainty.*
>
> (Lakatos 1976, p.2)

He then quotes Dieudonné who insists on the 'absolute necessity imposed on any mathematician who cares for intellectual integrity to present his reasoning in axiomatic form'. Formalism then is derived from logical positivism whose main contention, according to Rudolf Carnap (1983), is that a statement is meaningful if it can be proved/ disproved or shown unprovable.

This principle, however adequately it describes the finalised outcome of mathematical activity, is not helpful in describing the process of mathematical growth. If mathematics

> *does not grow through a monotonous increase of the number of indubitably established theorems but through the incessant improvement of guesses by speculation and criticism, by the logic of proofs and refutations,*
>
> (ibid., p.5)

then formalism is a historically inadequate account of mathematical activity; or in Karl Popper's words (1959), not revealing enough of its logic of discovery. If mathematics follows the evolutionary pattern of scientific paradigms then, in tune with Thomas Kuhn (1962), mathematical knowledge grows through revolutionary transitions fuelled by refutation and conflict, and the history of mathematics is written, not as a static succession of breakthroughs and successes, but as a time-relevant activity which is majorly determined by its informality, intuitions and logical uncertainty. In brief, mathematics is as strongly determined by the power of its constructions as by the climactic events through which the constructions were completed. Paraphrasing Goethe — quoted in (Kline 1982) — 'the history of mathematics is mathematics itself'.

This anti-positivism can be reinforced by the metamathematical relativism that certain 20th century developments regarding the foundations of mathematics have sparked off: as Kurt Gödel's demonstration of the incompleteness of logical-mathematical systems shows (1947), the axiomatic method, favoured by logicoformalist approaches to mathematics, is only relatively valid. Moreover the development of mathematical areas, such as non-Euclidean geometries, pinpoints that a slight modification of the axioms of the theory can lead to different mathematics.

Another dimension to this relativism is added by the realisation that mathematical activity, as well as time and era-relevant, is also culture-relevant, as suggested by the anthropological revision of mathematical history in the works of d'Ambrosio (1985), Ascher (1991), Cole (1971; Gay & Cole 1967), Gerdes (1986; 1988), Joseph (1992), Kline (1962; 1972), Lancy (1983), Pinxten (1983) and Wilder (1981) to mention but a few. Moreover, like language and music, mathematics unfolds in a completely original way whose objective rationalisation is not definitely possible. Wittgenstein (1978) has shown this inexorably: the logic of mathematical discovery is a language game whose rules are severely characterised by the fragility of the ambiguity which is at the same time the glory and the plague of language. As Michel Foucault (1973) established, a historicised approach to mathematics reveals, not only the conscious acts of conflict that propel progress, but also the unconscious level of discomfort and difficulty that determined conflict. In other words, an account of mathematical knowledge that incorporates its fragility is, almost paradoxically, a more powerful account.

In sum mathematics here has been described as a discipline whose collective consciousness cannot be memory-void. The achievements of mathematical growth cannot be treated as rootless entities, as detached abstractions from the thought processes — often imperfect and uneconomical — that brought them into existence. Mathematical knowledge, as a genuine product of language and culture, is constantly modified and challenged and contains time in a catalyst-like fashion. In Part Ib, this dynamic perception of mathematical growth is extended in order to describe the views on the ontogenesis of mathematical knowledge, the mathematical growth of the individual, espoused in this study.

Ib. The Ontogenesis of Mathematical Knowledge: Epistemological Obstacles

Alan Bishop in his description of mathematical activity as essentially cultural (1991) quotes George Kelly who claimed that 'we grow cognitively by handling contrasts'. Whether these contrasts are between the individual's view of the world and their experience of the world — in which case contrast possibly leads to personal conflict — or are embedded in the complexity of experiences — in which case contrasts are only multiple facets of experience — individual growth is not possible without reconciliation and accommodation of the contrasts. In this Part cognitive growth of the individual is described in terms of a parallel between how mathematics grows and how mathematics is learned: or in terms of whether ontogenesis imitates phylogenesis (Pinxten 1987) in that learning takes place as the constant overcoming of Epistemological Obstacles.

This parallel was first suggested within the discourse on the epistemological growth of mathematics. Poincaré (1946) and Pólya (1962), as indicated by Lakatos (1976), propose the application of Haeckel's fundamental biogenetic law according to which ontogeny recapitulates phylogeny to mental mathematical development. Freudenthal (1983) refined this proposition by suggesting that ontogenesis recapitulates phylogenesis in a modified way, that is the individual experiences the obstacles that have been phylogenetically experienced but the experience of the individual is modified by the added value of the developments bequeathed by the mathematical culture. So Freudenthal brings our attention to a non too-literal interpretation of the suggested parallel and suggests instead a cautious approach, that is conscious of the constraints and the crudities of absolutely adopting the parallel. From the perspective of Part Ia, if the growth of mathematics is the outcome of the constant refutation, falsification and modification of previous theories, if the evolution of mathematical knowledge is a never-ending overcoming of hindrances, then there must be considerable benefits to be earned, regarding the learning process of the individual, from the awareness of these hindrances. Works of Luciana Bazzini, Hans Niels Jahnke, Marta Menghini, Georges Glaeser, Francesco Speranza, Hans- Georg Steiner, Horst Struve (all mentioned in (Nardi 1990)), Evelyn Barbin (1989), Paolo Boero (1983) and others are attempts to embed an application of the parallel into specific mathematical areas.

What seems to me however most relevant to this study is the parallel between phylogenesis and ontogenesis as substantiated in the theory of Epistemological Obstacles, a notion introduced by Gaston Bachelard (1983/1938) and resurrected by Guy Brousseau (1986) in his theory on the foundations and methods of mathematics education. In the notion of Epistemological Obstacle Kuhn's notion of paradigm shift (during which a paradigm commits a leap of growth and progress in a process of challenge, conflict and reform) and Piaget's genetic epistemology (Part IIb) seem to converge.

Brousseau and Sierpinska (1994) quote Bachelard in describing Epistemological Obstacles as a category useful for the 'psychoanalysis of scientific thought', for explaining the shift from ordinary to scientific thinking through the tendency of the human mind to generalise or to construct tentatively universal laws often without establishing the legitimacy of the generalisation; or to accommodate experiences by describing them in terms of familiar but not necessarily adequate metaphors. Learning then occurs as an incident of Bachelard's 'intellectual repentance' and in this sense it is not an accumulative process.

The novice mathematician's cognition as observed in this study is explored in terms of the discourse on Epistemological Obstacles (Part II and Chapter 2). As explained there in more detail, the emphasis is put on the novice's cognitive growth through the enculturating experience of overcoming a series of epistemological obstacles related to their shift from concrete to abstract mathematical thinking.

Ic. A Constructivist Approach to Mathematics Education

In parts Ia and Ib the growth of mathematical knowledge was described, phylogenetically and ontogenetically, in terms of a dynamic, fallibilist epistemology. Either as a paradigm or as a learner's cognitive activity, this growth involves the constant challenge of refutation and conflict and the overcoming of epistemological obstacles. Metamathematical relativism, language, as the syntax of communication, and culture, as the milieu of communication, were highlighted as important influences on this growth.

In the previous parts mathematical learning was described in terms of the ontogenesis of the individual learner. This assumption of learning as an idiosyncratically individual process as well as the underlying relativism of the description in parts Ia and Ib are the main philosophical links between the epistemology of mathematics and the philosophical ideas on mathematics education espoused in this study. The dominant characteristic of these ideas is that they can be located within the area of a constructivist perspective on mathematics learning.

Constructivism in mathematics education, as most solidly advocated by Ernst Von Glasersfeld (1983; 1987; 1991 and 1995) describes mathematical learning as an inward and idiosyncratic personal construction process during which the learner seeks to assign meaning to mathematical knowledge. Knowledge in the constructivist way of thinking is not taken as a static body that exists platonistically outside the knower. Knowledge is rather a dynamic set of notions whose acceptance results from constant negotiating within the cognising community. It does not reside outside the knower and therefore knowing is a construction of experience. Given the strictly personal character of learning, knowledge cannot be transmitted since the interpretation that each learner constructs may differ substantially from the interpretations of other learners.

As Paul Cobb et al (1992) note from an educational point of view, interaction and negotiation are key notions because they suggest a promising teaching alternative to traditional transmissive models. Communication then, in the sense of a consensual construction of meaning, is based on the construction of intersubjective knowledge, namely knowledge that has transcended from the personal to the taken-as-shared and has now been institutionalised. The status of institutionalised knowledge, once acquired, does not remain unchallenged but operates temporarily in an almost objective fashion.

In this study the above interpretation of constructivist views on mathematical learning encompasses the notion that social interaction is paramount. That knowledge cannot be distilled or conveyed is a matter more directly linked with the ontological question whether objective reality exists. From an educational point of view, it might be more important however to address the issue of knowledge growth through communication. Recognising learning as a unique, for the individual, neuropsychological and sociocultural mental process seems to me to be more educationally central. I certainly recognise that it is strategically important to decide whether something 'is' or 'is agreed to be' because 'is' implies natural acceptance and 'is agreed to be' implies acceptance after negotiation and persuasion. From a psychological point of view however it is almost irrelevant whether a group in Abstract Algebra is defined or a group exists. Anything can begin to exist once defined. I also note that from an educational point of view, the failure of transmissive models of learning suggests that perhaps 'negotiated existence' is more likely to achieve acceptance in the learner's mind than 'realistic existence'. In this vein the constructivist notion of the ownership of knowledge, that is the intensity of the sense of understanding generated by construction (as opposed to transmission/ reception) of knowledge is crucial.

The constructivist views of mathematical learning are mostly founded on Piaget's genetic epistemology and psychology of learning and Vygotsky's theories on the influence of language and culture on learning. I note that the Piagetian theories constitute the cognitive dimension of the psychological influences of the study; the influences from developments in cultural psychology originate in the Vygotskian perspective as well as linguistics. In Part II I present an account of the psychological principles of the study, as they stem from an intention to embrace the cognitive (Part IIb) and the sociocultural (Part IIc) dimension of mathematical learning. In order to situate this psychological discourse within advanced mathematical cognition, which is the focus of this study, I first describe the perspective espoused here with regard to the nature of Advanced Mathematical Thinking (Part IIa).

PART II The Principles of the Psychology of Mathematics Education Espoused in this Study

In Part I a description was given of how mathematics grows either as a discipline or within the individual learner. Subsequently a philosophy of mathematics education was outlined. In this Part the focus is on outlining the psychology of mathematical thinking espoused in this study. Given that the theme of this study is the novice mathematician's thought processes, mathematical thinking is explored at its advanced level and this exploration mainly encompasses the novice's encounter with and induction into mathematical abstraction. Mathematical thinking at this advanced, abstract level is referred to as Advanced Mathematical Thinking (AMT), a name borrowed from David Tall's synonymous book (1991c) as well as from the synonymous PME Working Group. The nature of AMT and its cognitive and sociocultural dimensions are discussed in the following three sections of Part II.

IIa. The Nature of Advanced Mathematical Thinking

Here the description of the mental process of AMT is a condensed account based mostly on the works of Poincaré (1946), Hadamard (1954) and some of the authors of Tall's *Advanced Mathematical Thinking*. Given that different authors have different purposes and address different audiences (mathematicians, psychologists, educators), I have tried to homogenise the description favouring a psychologised tone. I also note that my focus is on formalisation and abstraction which are the features that distinguish elementary from advanced mathematical thinking and are the aspects of the novice mathematician's thinking mostly explored in this study. This account is not developmental but it aims at capturing some pre-eminent features of advanced mathematical thinking.

Apparently different people think differently; even among mathematicians it is hard to find two individuals who give identical accounts of their mathematical thinking. Hadamard favoured the classical distinction between the logical-analytical and the intuitive-geometric mind but even within this schematic polarisation the variations are numerous. The human mind, as the developments in the neurosciences indicate, seems to work in a complex way in which both hemispheres participate. In fact the final outcome of this mental process is greater than the sum of the analytic and the geometric contributions. In this sense the dichotomy between intuition and rigour is another rather misleading polarisation since it omits what David Tall calls logical intuitions (1991a), that is intuitions at an already abstract level.

From the earlier works on the nature of mathematical thinking (Poincaré and Hadamard), AMT emerges as an idiosyncratically complex mental process of analysis and synthesis in which theory construction and testing play a major part. Contrary to the impression fostered by traditional curricula, proof is only a final stage in the process of establishing mathematical truth. In the following, I present the genesis of mathematical insight (Part IIa.i) and proof (Part IIa.ii) as the two substantial features of advanced mathematical thinking. I note that in this part I am mostly concerned with describing the thought processes in AMT and not with the sociocultural powers that motivate and form it: these are discussed in Parts IIb and

IIc (for instance, in the light of these parts, proof, as opposed to insight, is seen as a sociocultural necessity rather than an intrinsically human cognitive need). Before describing the processes of insight and proof, I briefly refer to two terms (as used by Dreyfus in (1991)) that inform substantially my description of AMT: Representation and Abstraction.

Here Representation is taken in two senses:

- symbolic representations, that is notation, involving relations between signs and meanings.
- mental representations which are personal and idiosyncratic.

The former are external and used for communication, the latter refer to internal schemata. Kaput (1987) claims that mental representations are created in the mind on the basis of concrete representation systems; hence the competing representations of a concept that a learner might acquire. Success in mathematics relies heavily on the existence and competent manipulation of different representations. Switching between different representations and translations is a characteristic of mathematical fluency. Modelling, that is the act of finding a mathematical representation for an object or process, is one common way of representing.

Prerequisite to representing are generalising (expanding domains of validity) and synthesising (compressed merging into a single picture of previously isolated facts) which are processes that often are hard to distinguish from abstraction. Abstraction however requires the constructive act of shifting attention from the objects to the structures of their properties. The notion of abstraction is central to the research presented in this study and it will be extensively discussed in various chapters.

In sum representing and abstracting, that is the mental act of constructing, manipulating, linking and transcending representations, emerge as dominant features of AMT. This terminology strongly informs the following description of mathematical insight and proof.

IIa.i The Genesis of Mathematical Insight

From the preliminary arithmetical or algorithmic phases up to the creative and constructive ones, the creation of mathematics (by a mathematician or an individual learner; in a sense a learner re-creates for herself the didactically transformed mathematical insights of other mathematicians) can be seen as the intermittent genesis of mathematical insights. The insights that specifically, but not exclusively, determine AMT are, according to Ervynck (1991): creating a useful concept, discovering an unnoticed relation and constructing a useful ordering. This creativity is motivated by understanding (grasp, familiarity), intuition (fruitful selection, imagination), insight (reorientation towards what is important), generalisation and is characterised by selectivity, that is the power to condense and objectify. Contrary to the outcome of this process which is traditionally expected to be rigorous and precise, this creative process can be erratic, circular and deeply revisionistic (the fallible character of mathematical cognition was outlined in Part I in the account of the epistemological underpinnings of the study).

IIa.ii The Genesis of Mathematical Proof

As mentioned above, traditional formalist views of proof are the norm in mathematical community. A more careful look at the mechanisms of acceptance of a proof however reveals that, similar to the outcomes of all human activities, a mathematical proof is submitted to a context-dependent scrutiny.. In other words acceptance of a proof is a sociocultural process, some of the characteristics of which are dealt with subsequently (Hanna 1991).

Axiomatic deduction, hailed by Hilbert (1918) as the most rigorous form of mathematical proof, originates in the 19th and 20th century development of the discourse on foundational issues in mathematics. Since then however, and in spite of the ephemeral inclusion of deductive proof in most mathematical curricula, further developments have unveiled the sociocultural embeddedness that characterises proof (see also Part I). Formal validity differs qualitatively in different contexts. Above all, proving is convincing and the rhetorics of conviction are subject to a large number of communicational conventions. Moreover on the forefront of mathematical creativity, new mathematical proofs are often presented in elliptic, condensed forms that require a certain amount of suspense of disbelief from the reader; or in other words the author transfers the responsibility of discovering the exact anatomy of a theorem to the reader. In that sense formalism is more an aspired-to ideal, a paradigmatic force than a fully-fledged common practice. Formal proof in other words is the driving force and the aim of official mathematical communication but it is materialised on the basis of a number of conventions; these conventions are characteristic of the formal mathematical culture and their adoption is synonymous to a learner's advanced mathematical enculturation.

In the subsequent chapters one of the strongest tensions in the novice mathematician's induction into the mechanisms of advanced mathematical thinking is between the persistent intuitive practices of the novices and the affiliation to rigour that is equally persistently instilled in them. This brief account of advanced mathematical thinking is partly aimed at re-establishing informal or semiformal mathematical insight, intuition, with the appreciation it deserves. In what follows the focus is on the cognitive and sociocultural theories regarding mathematical cognition that have influenced this study.

IIb. Psychological Theories Relating to the Cognitive Nature of Advanced Mathematical Thinking

As mentioned in Part I the psychological theory that is coherent with the views on the growth of mathematical knowledge espoused in this study is most dominantly the developmental theory of Jean Piaget's genetic epistemology. In the following the relevant aspects of this theory are outlined followed by brief accounts of learning theories that relate specifically to advanced mathematical cognition.

I note that the psychological theories presented in this part of the chapter will be supported by concrete examples in Part III where the psychological developments in particular areas (conceptual and reasoning) of advanced mathematical cognition are explored.

IIb.i Jean Piaget's Genetic Epistemology

Like Kant in *The Critique of Pure Reason*, Piaget (1970) presents the knower as the constructor of their own knowledge of the world. The construction takes place mediately through the various (bio-neurological, cognitive) capacities of the knower. Knowledge is in this sense 'phenomenal'. Piaget expands the Kantian 'what we directly experience is a sensory manifold' to an explanation of the genesis of mechanisms through which experience is acquired. Central to the investigation of these mechanisms are his concepts of Equilibration and Reflective Abstraction and the focus of his thinking is the dynamics of epistemology (1975; Piaget & Inhelder 1963).

Piaget encapsulates in his notion of Equilibration the impact on cognitive development of maturation and experience of the physical and social world. Consciousness tends constantly to what Gestalt psychology (Ellis 1938) calls a harmonious Equilibrium. The Piagetian Equilibration is a series of cognitive actions: *disequilibration, assimilation* and *accommodation* through which this equilibrium is attained or regained. The integration of novelty can be achieved with greater or lesser degrees of internal reorganisation and reconstruction. Since the experience of dis- and re-equilibrating resides in the knower, it is possible for the knower to make constructions that in the cognising process may prove inadequate. Attempts at re-equilibration are particularly dependent on the *adequacy* of the already existing cognitive structures to adapt and on the *degree of recognition* of the inadequacy by the learner. Conflict between extant mental structures and new information triggers off accommodation. Therefore conflict generates a reorganisation of the cognitive structures; conflict breeds creation. In this sense the Piagetian notion of Equilibration resonates with the theory of Epistemological Obstacles (Part Ib) in that they both emphasise the crucial role of the scheme construction-conflict-reconstruction.

The most powerful form of equilibration is that of re-equilibration of the cognitive structures to a disturbance by undergoing reconstruction; this process is Reflective Abstraction. This has two facets: *reflecting* cognitive structures to a higher level and *reconstructing* them to accommodate the needs in this higher level. Below I explain Reflective Abstraction in more detail following mostly Ed Dubinsky's account.

Ed Dubinsky has extended the Piagetian idea of Reflective Abstraction (1986 and 1991) to describe the epistemology of various concepts and to illuminate advanced mathematical cognition. He presents Reflective Abstraction as the complex process through which the knowledge drawn from actions on objects (such actions are for example empirical or pseudo-empirical abstractions) is interiorised into new objects. What differentiates Reflective Abstraction from these earlier forms is that Reflective Abstraction is a general co-ordination of actions, and does not focus on the actions themselves. The property drawing involves a cognisance or consciousness of actions and whatever is abstracted is projected onto a higher plane of thought where other actions are present as well as more powerful modes of thought. So beyond abstraction lies construction during which new combinations are constructed by a conjunction of abstractions.

Construction in Reflective Abstraction can have various forms, mainly *interiorisation, co-ordination, encapsulation* and *generalisation*. Interiorisation is the 'translation of a succession of material actions into a system of interiorised operations'; co-ordination is the composition of two or more processes to construct a new one; encapsulation is the conversion of a dynamic process into a static object; and generalisation is the extension of the domain of application of existing schemas. Dubinsky also includes reversing an original process as another form of construction in Reflective Abstraction. I note that this concise reference to the Piagetian dynamic epistemology is embedded in theories relevant to advanced mathematical thinking (Part IIb.ii) and to particular mathematical topics (Part III).

IIb.ii Brief Accounts of Learning Theories Influential to This Study Relating to Advanced Mathematical Cognition

In the following I present a brief account of a few influential learning theories related to AMT that have been directing the analytical thinking of this study.

<u>Concept Image and Concept Definition</u>. Concept Image is a term coined by Tall and Vinner (Tall & Vinner 1981; Vinner 1983) and describes the learner's mental structure associated with a concept. Unlike formal concept definitions, concept images need not be coherent since their various parts are not evoked at the same time and at times contradict each other. Contradictions are revealed when a conflict occurs and cognitive structures are forced to re-equilibrate in ways that accommodate the contradictory parts.

Concept Definitions are traditionally seen as the main source of the learner's concept acquisition (Vinner 1976) and learners are expected to use definitions once they have been introduced to them. Their elegance, conciseness and descriptive power are considered the most efficient tools for introducing a concept even though these characteristics do not reflect accurately the learners thought processes. As substantiated with examples in Part III, the name of a concept does not necessarily evoke the concept definition in a learner; in most contexts people do not even consult definitions. Even when learners are expected to adopt more systematic behaviours, they do not. What they do is employ a less technical but more common-sense approach: what a concept name evokes is a conglomeration of non-verbal associations with the concept: visual representations, impressions or experiences. These associations are called concept images and are personal and circumstantial. The concept definition is only, if at all, a scaffolding device for the formation of this image. The controversial element of this theory (concept image as a more powerful mental structure than concept definition) lies in the surprising realisation that learners uncritically transfer commonsensical, everyday behaviour to mathematical thinking. The power of cognitive habits however seems to overshadow issues related to meta-cognition of which the learner is not necessarily consciously aware.

Consider the following figure from (Vinner 1991):

concept image ⟷ concept definition

Either or both the cells of concept image and concept definition may be empty, or, if not, they might interact. Given the definition the learner, who already has a concept image, might

- accommodate the image to include the definition
- keep the same image and forget the definition
- keep both and evoke them independently

The interplay between the two cells can take any form.

<u>Theories Related to the Encapsulation of Mathematical Processes into Conceptual Entities.</u> A type of Reflective Abstraction that is frequently recognised in accounts of mathematical cognition is the formation of mathematical objects that contain procedural elements and, in particular, originate in mathematical processes. In other words mathematical processes are often encapsulated into a new mathematical entity and in turn become the objects upon which cognitive actions are applied. This reification, as it is called by Sfard, condenses and consolidates information facilitating thus its retrieval from memory. Facilitated evocation can accelerate comprehension in a Gestalt manner, that is in a global and comprehensive way.

Anna Sfard, who has written extensively ((1989; 1991; 1994 and Sfard & Linchevski 1994)) on the dual nature of mathematical concepts (process and object), sees reification as the birth of metaphor. For Sfard, who draws on a variety of classical mathematical texts as well as contemporary semiotic studies, meaningful abstraction is a result of the construction of appropriate metaphors, namely 'figurative projections from the tangible world onto the universe of ideas' (Sfard 1994). According to a number of psychological studies, learning proceeds with the transformation of constantly renewed experience into metaphors for a generalised, condensed and simultaneous organisation of this experience: with regard to mathematics these metaphors are reified processes. Sfard (above), Dubinsky (1991) and others contend that, perhaps unlike mathematicians who are trained in the activity of abstraction, operational familiarity with a concept precedes the structural one. This idea is not to be taken in its absolute sense since thinking is an obscure and chaotic activity during which the possibility of having first a fuzzy, elusive, object-like grasp of a concept (and then elaborating on its procedural aspects) cannot be ruled out. In this sense Metaphor Construction may be a superior descriptive tool to Reification which can often be taken as implying that operational stages precede structural ones. A powerful version of Reification can be one that does not exclude the dialectic interaction of the operational with the structural aspects of the concept.

Tall (1991b) and Tall & Gray (1994) introduced the term Procept (from process and concept) to describe the 'amalgam of three components: a process that produces a mathematical object, and a symbol that represents either the process or the object'. The ambiguity that is inherent in a procept, contend Tall and Gray, depending on the cognitive maturity of the learner, can become a facilitator of flexibility in switching between the operational and the structural aspects of the concept.

In the above the Piagetian Equilibration and, in particular, Reflective Abstraction were embedded within the cognitive discourse on AMT in that the theories of concept image-concept definition, of reification (and of metaphor construction) and of procepts draw — more or less explicitly— on the Piagetian theoretical framework. In Part IIc this perspective which has been exclusively from the individual learner's point of view, is modified in order to encompass the notion that there is no such thing as learning in a clinical, controlled-experiment sense: learning takes place in a sociocultural context of which language (and, in advanced mathematics, notation too) is an inextricable part. In fact AMT is a mental activity where the distinction between thought processes and language structures becomes significantly blurred. In the following AMT is described in terms influenced from cultural psychology — mostly the Vygotskian perspective — and linguistics.

IIc. Psychological Theories Relating to the Sociocultural and Linguistic Nature of Advanced Mathematical Thinking

In Parts IIa and IIb mathematical thinking was described as a mental process. However unlike behaviourism and, to a lesser degree, Gestalt psychology, that seem to underestimate the sociocultural embeddedness of cognitive activities, this mental process is not seen as taking place in an environmental vacuum. Vygotsky (1986 and 1978) acknowledges language, an element of intrinsic sociohistorical nature and in Bruner's words (1960a and b) a major cultural amplifier, as a dominant element in thought processes. Language functions as the essential facilitator of the transition from unmediated sensory reflection to mediated rational thinking and as such it is the fundamental psychological element bearing the features of the cognitive development of the learner. In this part I deal with the sociocultural and linguistic dimension of cognition.

IIc.i The Vygotskian Perspective on the Sociocultural Nature of Cognition

Vygotsky develops a theory of mind in which thought and language develop together: conceptual growth is grounded on linguistic experience and social interaction. While his theory adopts what the symbolic interactionists like Mead

(1934) have introduced as a focus on the nature of this social interaction, Vygotsky contends further that social interaction forms learning and thinking fundamentally. Meaning is constructed in a continuous social process of negotiation.

So for Vygotsky the psychological process of cognition lies in the subjective discourses as the acts of signification and meaning-making of the individual; he thus does not separate the individual cognising subject from the processes of communication that form its consciousness. In a sense Vygotsky here meets the Lacanian subjective self (Lacan 1977), constantly formed by the discourses about it. This emphasis on the formation of consciousness through communication does not necessarily entail a reduction of the idiosyncratic and personal to a homogeneous social: communication is always of a situational character. Discourse is strongly characterised by intersubjectivity and, as Lerman stresses (1996), we cannot ignore the claim that is 'common to Piaget and Vygotsky that objects or tools become a part of an individual's life when they are acted upon; they cannot become the individual's in a "transmitted" sense'.

Beyond the Piagetian recognition of language as a 'manifestation of the symbolic function, Vygotsky rejects a representational view of the mind and its functions. Language is not a mirror of consciousness but a correlative of consciousness, and consciousness takes form and significance when it is articulated. If, then, the formation of the individual's consciousness is seen as a social and a cultural process, a Davydovian mastering of cultural tools (Davydov & Radzikhovskii 1985), then the discourses on learning and teaching merge into one closely interrelated discourse because learning is then seen as an enculturation process. The mathematical language creates a reality, mathematics: it then describes this reality. By the same token then a study of the novice's learning is a study of the process of appropriation of the individual's meaning making constructions to the ones of the mathematical culture.

In the above it is necessary to emphasise the attempt for a co-ordination of two possibly conflicting perspectives: the individuality of the Piagetian cognising subject and the sociocultural embeddedness of the Vygotskian subjective self. For this study the resolution of the possible conflict lies in clarifying that the Piagetian epistemology is employed only as so far as it does not segregate the signified from the signifier. From this point of view the Piagetian self departs from where the Vygotskian self arrives in a relay-race manner: language in the form of the intersubjective discourses of mathematical enculturation provides the internal architecture of consciousness which then develops through an internalised discursive processing. In this study the Vygotskian perspective is employed as a descriptor of the enculturating discourse that determines the learners' mathematical experience; the Piagetian perspective is introduced as the theoretical basis of the advanced mathematical learning theories employed in this study (Part IIb) and are used as aids in the interpretation of the learners' internalisation of some mathematical concepts. The distinct and complementary use of the two perspectives is exemplified in the analytical chapters (Chapters 6-10).

IIc.ii The Anthropological and Linguistic Perspective on the Enculturation Into Advanced Mathematical Practices

In Part IIc.i the study of mathematical cognition was discussed partly as a process of enculturation of the novice learner to the discursive practices of the mathematical culture. I think it is necessary to emphasise at this point that the notion of enculturation employed in this study departs from what is commonly thought in cultural psychology and anthropology as transmission of cultural practices (Bishop 1991). On a grand anthropological scale culture has been traditionally seen in an objectified sense and this objectifying metaphor has possibly helped in an initial and schematic understanding of human civilisation. Contemporary cultural theories however move critically beyond this simplistic transmissive perspective: to explain the notion of enculturation employed in this study I present briefly Sierpinska's (1994) use of the cultural theories of E T Hall (1981/1959) and Michel Foucault (1973) in order to describe the systemic conventions of the mathematical culture — semantic, linguistic and logical — as major determinants of the learner's cognition. This concise description is exemplified in more detail in Part IIIa of this Chapter and resonates with the theory of Epistemological Obstacles that to a large extent directs the thought of this study: for Hall Epistemological Obstacles are a deeply cultural phenomenon and, in particular for Foucault, they form an unconscious-to-become-conscious part of the scientific discourse within which the individual cognising subject operates.

For Sierpinska, 'in Foucault, *Man*, the abstract construct, the '*sujet épistémique*', becomes finally *a man*, unique, individual, only forced to accept the binding rules and categories of the *épistémê* he happens to find himself historically tied within under pain of appearing as mad'. She then suggests a co-ordination of Foucault's *épistémê* with Hall's *cultural triad*. I first outline Hall's theory and then draw the parallel between the two.

Hall recognises 'three types of consciousness, three types of emotional relations to things': the 'formal', the 'informal' and the 'technical'. In the context of mathematical culture the 'technical' level is the level 'of mathematical theories, of knowledge that is verbalised and justified in a way that is widely accepted by the community of mathematicians. At the 'formal' level, our understanding is grounded in beliefs; at the 'informal' level - in schemes of action and thought; at the 'technical' level — in rationally, justified explicit knowledge'.

Central to the purposes of this study are processes taking place within the informal level of Hall's triad. This is, in Sierpinska's words, 'the level of tacit knowledge, of unspoken ways of approaching and solving problems. This is also the level of canons of rigour and implicit conventions about how, for example, to justify and present a mathematical result'. The novice mathematician's enculturation is seen in this study as taking place at the informal level: through the accumulation of mathematical experience shared with the expert and in the process of appropriation by an internalising imitation of the expert's cultural practices.

Foucault's *épistémê* is described by Sierpinska as a triad of 'related categories, rules of sense and rules of rationality prevalent in a given epoch and culture' and is proposed as parallel to Hall's formal, informal and technical levels of cultural consciousness. In particular relevance to the focus of this study is her parallel between Foucault's rules of sense and Hall's informal level as the not always fully-articulated, possibly unconscious guides of our sense-making, our ordering of the world. The significance of the parallel for this study lies in that Foucault's archaeology of knowledge in *The Order of Things* is in more direct resonance with the philosophical tradition of Epistemological Obstacles since they both provide a diachronical perspective of enculturation, whereas Hall's theory is more of a synchronical character (a diachronical perspective in this sense is one that looks at one culture in a dynamic, evolutionary sense; synchronical is one that attempts to look spatially at various cultures and express their current states in a unified discourse).

In this part I briefly discussed the notion of Enculturation which was used as a major descriptor of mathematical cognition in Part IIc.i. So in Part IIc I have presented the aspects of cultural psychology that have influenced this study as well as attempted to co-ordinate them with the aspects of cognitive psychology that have influenced this study and are described in Part IIb. In Part III, then, this co-ordinated perspective is employed as the descriptive tool for the presentation of relevant literature on the novice mathematician's difficulties in specific conceptual areas as well as with aspects of mathematical reasoning. Part III in this sense is designed to bridge the psychological theoretical background of the study with its materialisation which is presented in the subsequent chapters.

PART III Developments in PME-AMT Relevant to this Study

In this study the mathematical learners that are the focus of the inquiry, the first year mathematics undergraduates, are often called *novices* or *novice mathematicians*. The term is used emphatically as opposed to the *expert mathematicians* that teach or tutor the students so that the juxtaposition is simply referring to the amount of mathematical experience that the students and their teachers have acquired. However, given that the novices' learning, apart from a psychological mental process, is seen also as an enculturation into formal mathematical practices, in the study the differences between the two groups are also often discussed in terms of the mathematical expertise that the novices aspire to and the experts demonstrate.

Some of the features of this mathematical expertise are linked to the emotion of the mathematical experience. Rosamond (1994) reports that 'emotional elements interact with cognitive processes and exert strong influences' during problem solving. Her study is mostly concentrating on the exploration of these emotional elements over a wide spectrum of mathematicians — experienced or not — and not on the differences between novice and experts but it seems that the role of emotion, even though strong in both groups, especially in the form of anxiety or lack of confidence, is very important as far as the novices are concerned. Mc Leod (1987) reports however that experts manage their emotions better and do not allow them to interfere with their cognitive processes. In a sense they are in more control of their cognition. Moreover they seem to have strong aesthetic considerations — for instance they prefer to avoid powerful mechanisms for elementary problems and in general they try to match solution method with problem level or they favour elegant and economic solutions versus tedious case analyses — and often this attitude proves constraining (Rosamond 1994).

On the other hand Schoenfeld & Hermann (1982) found that novices rush to an answer, use known procedures uncritically, believe there must be a formula for each problem and go on 'mathematical wild goose chases'. They classify problems according to their superficial, rather than their deep structural characteristics in contrast to experts who seem to categorise more by principles used in the solutions. Also experts possess organised and integrated knowledge of a domain, including how domain-specific principles, general concepts and reasoning processes are related as well as the ability to recognise patterns of problem features. Selden & Selden and Mason (1994) report that novices think all problems can be completed in at most five minutes — possibly influenced by their previous classroom experiences — and this affects their ability to solve non-routine problems. Becker and Pence (1994) also found students to be learning either as splitters (analyse information logically and break it down into smaller parts) and lumpers (look for patterns and relationships).

Underlying the above juxtaposition between expert and novice practices is the discourse on a wide variety of issues related to advanced mathematical cognition. In this part of the chapter I address some of these issues with regard to the developments that are relevant to this study. So the three sections of this part are about the novice mathematician's difficulties regarding

- language, notation and visualisation,
- reasoning and proof, and
- particular mathematical topics and concepts.

IIIa. Developments in PME-AMT Regarding Mathematical Language, Notation and Visualisation

Language was described in Part IIc as a major determinant of consciousness. Moreover the mathematical culture is strongly characterised by its forms of expression — mathematics is often described as the activity of constructing metaphors that reveal the answers to extraordinary questions relating to abstraction which in turn is a form of imagination (Bullock 1994). So since language is traditionally seen as an expression of imagination, mathematics can be seen as the language of abstraction. Moreover given that the learners' cognition in this study is explored in the context of the mathematical discourse developed in tutorials, it is also significant that language is also seen as a major determinant of communication: whether it is ordinary language, or the formal syntax and notation of mathematics, or visualisation, language is the architecture of thought as well as its carrier. In this part I review some current research on the communicational aspects of AMT which relates to the purposes of this study.

Laborde (1990) notes that oral mathematical discourse has not been extensively researched even though there is evidence that suggests its importance as far as the learner's cognition is concerned. She contends that the rarity of research in the field

> *reflects the practical and theoretical difficulties of research on oral language: the transcription of spoken language needs time to be done very accurately, and the analysis of such transcripts is generally complex because the degree of implicitness of oral discourse is greater than in a written communication and elements of the context constituting the enunciative situation play a more important role. These constraints prevent the analysis of long pieces of dialogues.*
>
> (Laborde 1990, p.60)

She then refers to David Pimm's (1987) notion of *teaching gambit*, that is the discourse strategies employed by teachers in order to transfer some of the learning control to the learners themselves and stresses that the most common of these strategies, the very closed questioning, 'allows a very narrow scope for answers and [denies] students practice in formulating long explanations'. In this study the effects of closed questioning on the novice's cognition are frequently discussed and juxtaposed to more open dialectics as well as some forms of student-student interaction.

Also in this study what Laborde calls the *enunciative situation* has been a serious consideration in the sense that the students' cognition is explored, not in spite of the different contextual situations in which this cognition is expressed, but exactly as a transition from one form of expression to another. Clement et al (1981), Ghosh and Giri (1987) and Burton (1988) discuss the novices' difficulties in translating from ordinary language into mathematics and in resonance with the theoretical position of this study that learning difficulties are revealing as far as cognition is concerned, the fact that students express differently on the blackboard, in their writing and in an exam is used here as illumination of their cognition and not as a restraint.

Like Clements et al, interference from everyday logic and language (student-professor problem) is also mentioned by other researchers such as Janvier (1987) and Bjorkqvist (1993) who studied the students' personal conceptions of logical necessity and possibility and found influence from everyday conceptions and a partial dependence on key elements in the structure of the sentence; particularly relevant here are his findings on double modalities, the categorical form of propositions that contain two negations. Students were found to be deeply confused with these linguistic structures. By implication, when confronted with similar structures, that in addition are embedded as syntactic and semantic content of mathematical expressions, students are expected to be equally or possibly more confused. The influential role of content and form in mathematical cognition and the need to introduce it explicitly to the novice has been discussed also in (Byers & Erlwanger 1984; Abkemeier & Bell 1976 and Davis & McKnight 1984).

In the above the communicational aspects of advanced mathematical cognition have been discussed in terms of the learners' discursive practices and in particular their linguistic practices. The reason that research on visualisation is marginally mentioned here is that — with the exception of Linear Algebra and the persistently repeated metaphor of the plane — the students' and tutors' literally visual references are not frequent. So visual representations are not a large part of their repertory and when they appear they are often verbally described by the students as part of their intuitive access to some new concepts. At various points this metaphorical discourse is discussed as variably beneficial to the learners' development. Also it seems that the abstract nature of the mathematics discussed in the observed tutorials did not allow a substantial demonstration of the visual element of the learners' cognition. I note here that one important point made in the literature on visualisation (Bishop 1989; Davis 1989; Janvier 1987 and Presmeg 1986) is that while there is a strong visual element in mathematical cognition at all levels, when the students find difficulty in connecting different representations (for instance formal definitions and visual representations), they often abandon visual representations — which tend to be personal and idiosyncratic — for other more socially acceptable tentative ones.

Finally I discuss some work in the area of cognition as seen from the perspective of understanding mathematical texts. The reason I close this section with these brief references is that the mathematical discourse explored in the context of tutorials in this study is largely based on the students' — and the tutors' — understanding of mathematical texts (problem sheets, lecture notes). Furinghetti and Paola (1991) have discussed the understanding of mathematical texts in relation to the students' difficulties with formal proof. Other researchers have explored particular aspects of the students' responses to written mathematical text: so for instance Dee Lucas and Larkin (1991) found that proofs written in a verbal, ordinary language produced better performance than equation-based proofs on problems related to both equation and nonequational proof content. They also explain that equations cause students to shift attention away from non equational content and learners have more difficulty processing equations than verbal statements of the same content. Similarly MacGregor (1990) noticed that writing sentences helps students write correct equations and contrary to expectations the most successful students were those who used common idiomatic forms of English that could not be directly translated into mathematical notation. Finally Perrenet and Groen (1993) studied hint effectiveness and found that hints are effective when in a written question they stimulate concrete action and they are not when they are simply warnings against certain mistakes. In this sense signalling to the learners, without justification, certain outcomes of their action before this action takes place, seems not to be very convincing. Similar comments to the latter are made generally on the next section in relation to the students' response and difficulties with mathematical proof.

IIIb. Developments in PME-AMT Regarding Mathematical Reasoning and Formal Deductive Proof

David Reid who presided over the AMT sessions in PME19 (1995) stressed that the transition from informal to formal reasoning is 'the basis of both the concept of proof, developed out of informal deductive reasoning, and transformation of the structure of mathematics itself into objects of mathematical investigation. Informal reasoning, in the form of intuitions and insights, also plays an important role' in creative thinking.

Formal proofs have gradually disappeared from the school curricula in the last decades because it is a complex activity of the human mind and school needed to unload some of its complexity in order to guarantee a more democratic access to knowledge (and qualifications).

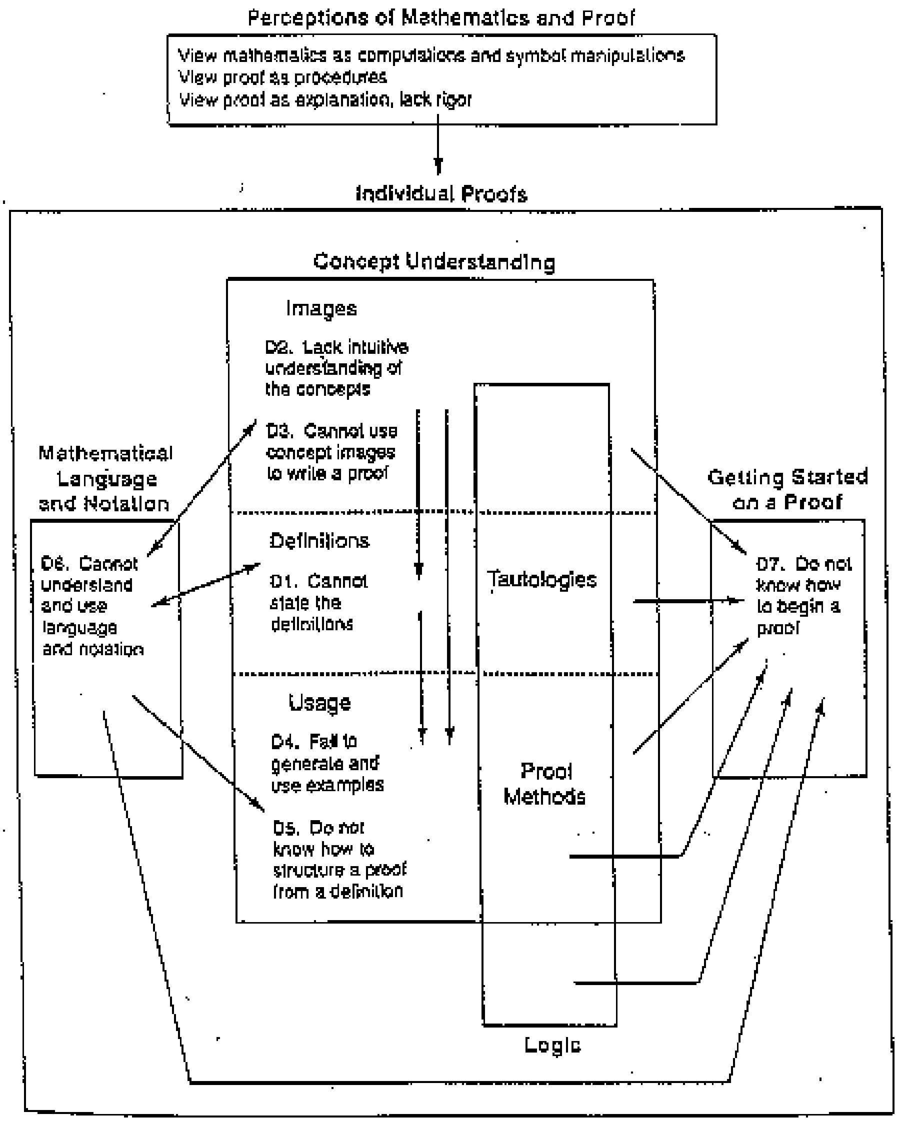
In this part the novice's difficulties with advanced mathematical reasoning are presented: in particular the transition from the concrete to the abstract mode of thinking and notions of formal mathematical proof (deductive and inductive) are dealt with as the dominant features of advanced mathematical reasoning. As Robert and Schwarzenberger note (1991),

- the increase in the quantity of material taught in advanced mathematics courses as well as
- the reversal in the path of presentation of this material (from the 'experimental' to the 'axiomatic - deductive')

are the basic differences between elementary and advanced mathematics as didactically processed knowledge (Brousseau (1989) and Chevallard (1985) distinguish between mathematics as a discipline and taught mathematics and note that the former undergoes a didactical transposition, in order to be used as the latter). Robert and Schwarzenberger also claim that the capacity for distinguishing between mathematical and metamathematical reflection as well as the capacity for the latter are inherent characteristics of AMT. The former is mostly hampered by conceptual difficulties with the subject (Part IIIc of this Chapter); the latter by the novice's inflexibility in adopting a wide range of perspectives when approaching a mathematical problem. In this study the need for versatility and for personal control over learning coupled with the Piagetian theories on the mathematical growth of the learner offer the theoretical framework within which the novices' difficulties with advanced mathematical reasoning are dealt with.

Here I present a brief compilation of works related to the novices' difficulties with formal mathematical proof and with a more global approach to the difficulties of advanced mathematical reasoning. Underlying the presentation is the assumption that a great deal of the significance of mathematical reasoning lies in the semantics and mechanics of proof and that, even though this is widely recognised, it comes through mathematics courses implicitly rather than as an explicit aim.

The transition from the experimental and intuitive habits of school mathematical reasoning to the formal requirements of advanced mathematical thinking is abrupt. Novices are not prepared for what Bell (1976 and 1979) called 'crossing the barrier between empirical awareness of a generalisation and a deductive insight'. Moreover the students' understanding of proof seems to suffer from the cumulative effect of their conceptual as well as their expression difficulties. Robert C Moore (1994) consolidates his findings on the students' sources of difficulty in doing formal proofs in the figure given below.



Model of the major sources of the students' difficulties in doing proofs.

Moore employs the Vinner-Tall-Dreyfus (Dreyfus 1990; Tall & Vinner 1981; Vinner 1983 and Vinner & Dreyfus 1989)) model of *concept image and concept definition*, reinforced with what he calls *concept usage* (the way a concept is used in order to generate examples or proofs) in order to describe the sources of difficulty he identified (D1 ÷ D7). The arrows indicate the interrelations among the various sources. In Part IIIa mathematical language and notation was described as one of these sources and in Part IIIc the implications of impeded understanding in particular mathematical topics are discussed. Here the focus is on

> the learners' perceptions of the nature of proof as well as
> the required logic and methodology of proof

as the major determinants of the novices' difficulty. I note that the largest number of studies mentioned in this section refer to research relevant to secondary education but since the participants of this study are first year mathematics undergraduates it is reasonable to assume that the findings of these works reflect the students' predispositions towards proof as they enter university. In the course of the presentation in subsequent chapters the actual and potential departures of the novices from these predispositions will be exemplified and discussed.

There is substantial evidence in research findings according to which students find proof difficult, unnecessary and meaningless. Also they view empirical evidence as proof and in fact prefer empirical arguments over deductive arguments.

Balacheff (1986) distinguishes between *pragmatic and conceptual proofs*: the former 'have recourse to actual action or showings' and the latter ' do not involve action and rest on formulations of the properties in question and relations between them'. Pragmatic proofs are either *naive empiricisms* ('asserting the truth of a result after verifying several cases') or *crucial experiments* ('experiments whose outcomes allow a choice to be made between two hypotheses'). Other findings (for example Martin & Harel 1989; Porteous 1990; Williams 1980; Yerushalmy et al 1990) also point at the students' view of empirical argument as sufficient proof.

Even when introduced to deductive proof students do not seem to appreciate its 'generic' aspect, namely, according to Balacheff (1990a), one that is relying on an object that is 'not there in its own right, but as a characteristic representative of its class'. Harel and Tall (1991), Tall (1989) and Mason and Pimm (1984) have discussed the difficulties that learners have in abstracting from a generic example these elements that constitute its genericity and this implies, as Williams says, that students then 'do not understand the generalisation principle of deductive proofs'.

Martin and Harel (1989) also found that 'use of the special features of a non-generic figure does not appear to influence judgements of the proof' and that proofs are seen as figure-particular.

Another hierarchical model on the development of proof is Van Dormolen's (based on Van Hiele's levels of geometric development): Van Dormolen (1977) distinguishes among *specific*, *common-properties* and *reason-about-reasoning* proofs. Bell, following his triad of functions for a proof (verification, illumination, systematisation) also acknowledges that the learners' proving practices seem to develop through *regularity and rationality* to *explanatory quality* and *logical sophistication*. Within the empirical/deductive dichotomy Bell also refines empiricism: from *failure to generate correct examples or to comply with given conditions*, to *extrapolation* and *non-systematic/partially systematic/systematic checking out of full finite set of cases*. His refinement of deductive behaviour follows a similar pattern of progressively becoming able to make connections and present connected arguments with explanatory qualities. Adopting Bell's and Dormolen's hierarchies, Galbraith (1981) also presented a levelled list of components of proving behaviour: a large number of these components were missing and Galbraith also notes that learners do not exhibit 'an objective and detached view, but rather exhibit restricted and on occasions psycho-emotional approaches to the use of process skills'.

As Fischbein (1982) notes, students are possibly not aware of the distinction between empirical and deductive arguments. Even when they are, says Schoenfeld (1987) who calls learners at this stage *pure empiricists*, they decline using deduction as a constructive tool for problem solving. Coe (1992) and Coe & Ruthven (1994) also demonstrate that

> *students' proof strategies are primarily and predominantly empirical, with a very low incidence of strategies that could be described as deductive...Students' primary concern is to validate conjectured rules and patterns [by]... testing them against a few examples'.*

As Chazan (1993) schematically proposes, there seem to exist two types of problematic predispositions towards proof: students either see empirical evidence as proof or deductive proof simply as evidence. Elaborating on the above distinction, Chazan notes that students seem to recognise, especially when their attention is drawn to it, that empirical arguments rely on examples which are special, measurements are not exact and there may be counterexamples. Instead of becoming sceptical about empirical arguments they prefer to modify their empirical strategies in order to accommodate some of the limitations and, as it will be demonstrated in this study too, counterexamples do not disturb their notion of mathematical truth which is not characterised by universality. Fischbein showed that proving is 'completely outside mainstream behaviour' and 'there is a difference between accepting a proof and accepting the universality of the statement proved by it'. There is 'a need for a complementary intuitive acceptance of the absolute predictive capacity of a statement which has been formally proved'.

Moreover in the cases where students object to deductive proof as mere evidence they also point out, ignoring the universality of a deductive proof, that it doesn't provide safety from counterexamples: a deductive proof is based on one example, is 'written in the singular' and is also 'based on assumptions'. Miyazaki (1991) too notes how the students tend not to recognise deductive proof as establishing the generality of propositions or conjectures. Finally degrees of persuasion vary given that depending on where the proof comes from and how it is presented, learners may feel obliged to accept a proof which they do not necessarily believe. This reflects the reality in the mathematical community where acceptance of a proof is often a result of a variety of sociocultural factors other than its sheer formality. Sekiguchi (1992) in his thesis points at the social dimension of proof presentation as a communication process and Hanna (1989a and b) enumerates a variety of dimensions in the social process of accepting a proof.

Duval (1991) attributes the learners' difficulties with mathematical proof to their confusion of deductive thinking with ordinary argumentation. But despite the fact that they both use similar linguistic forms and propositional connectives, these are two different modes of thinking. Implication by inference is inherent in deductive thinking as opposed to summatively connected arguments. This mechanism of inferential implication constitutes the basis of the logic of mathematical proof, of Math Logic as O'Brien (1972 and 1973) called it in his study of Modus Ponens. Students need to comprehend this mechanism in order to move beyond the persistent Child's Logic and engage in formal proof. Remarkably O'Brien also noted that the distinction between Child's and Math Logic is crude and does not accommodate a large percentage of student proving behaviours. But he also employs this dichotomy in order to identify the difficulty of the transition to formal proof.

The difficulty of dealing with the logic behind formal proof lies also in the fact that learners are overwhelmed by the content of the mathematical statements in a proof and are not able to move beyond content and into the realm of logical manipulation of the statements. Anderson (1994) characteristically refers to persistent inconsistencies in the students' proving behaviour when they fail to 'disentangle the logical elements from the analytic concepts'. A similar interference on the students' clarity and competence in logical manipulation was observed by Barnard (1995) in relation to the students' considerable inability to negate logical statements: the perceived 'meaning' of the statement, Barnard explains, interfered with the action of negation and prevented the students from acting on a logical statement regardless of its content.

The above difficulties with mathematical reasoning imply that students may, in view of these difficulties, avoid formalisation but they do not imply that there is no cognitive need for conviction and explanation. De Villiers (1991) demonstrates how both expert mathematicians and mathematics learners pursue the power of proving and Collis (1974) points at the development of a preference for logical consistency as a sign of cognitive maturity. The issue therefore is to establish the necessary connections between this need and the learners' conscious or subconscious decision making about the explanation means they prefer.

Significant to this study are findings that concern the students' attitudes towards proof in contrast with visual or intuitive arguments. Tall (1992), for instance, observed the students' preference for a generic proof (a generalisable example) of the irrationality of $\sqrt{2}$ over the classical proof by contradiction. Tall reports that the students prefer the generic proof because it demonstrates why the result is true. So for the students 'mathematical insight maybe more important than mathematical precision'. Also in Vinner's experiment with the Mean Value Theorem where the students were presented with a visual and an analytic proof, students found the visual more convincing and the ones who preferred the analytic one seemed to be more restrained by their fear that 'something was illegal' with the visual proof than by their believing in it.

This suspicion towards visualisation (see also Part IIIa) also has historical and epistemological roots. A characteristic example is Hilbert's aversion towards Euclidean geometry: his mistrust was based on claims that geometry depends on subtle intuitions about space whereas a proof must be universal and representation-independent. In fact what seems to breed the need for proof is the fear that something might elude intuition. Intuition cannot surely cover all possible concept images and can be misleading whereas proof is about conviction. Findings from experimental psychology also suggest that the learners prefer reasoning by operating on semantic images or using analogies rather than with ways that conform to formal logic (Greeno 1989). In the conflict between intuition and formal theory of logic-and-probability the former dominates.

IIIc. Developments in PME-AMT Regarding Particular Mathematical Topics and Concepts

This part broadly resembles the structure of the topical order in the presentation of data analysis in Chapters 6-9. Therefore, first, I deal with developments in PME-AMT with regard to learning difficulties related to real numbers, infinity, function, calculus and finally linear and abstract algebra. Proportionally Linear and Abstract Algebra are favoured in this presentation for two reasons: two of the four analysis chapters are about the mathematical cognition of Linear and Abstract Algebra; moreover among the mathematical topics related to this study, algebraic topics are the least researched. Therefore a review of recent developments in the area may be more crucial than one on, say, limit or function which are the 'older' advanced concepts in terms of the didactical interest they have raised. In fact the research on these concepts, especially on function, has been thoroughly reviewed in for instance (Dubinsky & Harel 1992). Judging from the consistent reappearance of reference to functions and calculus in all the recent reviews of developments within PME-AMT (Dreyfus 1990; Ferrini-Mundy & Graham 1991; Kaput & Dubinsky 1994; Langford 1987; Selden & Selden 1993 and Tall 1992) I would say that these two topics are the most talked-about and probably the only ones about which there seems to be a consolidating consensus as to where the field stands.

I note that the sources for the following reviews vary in some essential research characteristics: given the pre-paradigmatic state of the field, I draw equally on purely observational studies (sometimes carried out by mathematicians themselves who act out of a personal interest in how their undergraduates learn and what difficulties they encounter) to more psychologically and pedagogically grounded works. Most of the latter are based upon the learning theories outlined in Part II and so an opportunity is given to elaborate on these theories through their implementation on specific mathematical concepts. Regardless of their observational or explanatory potential the works reviewed here are selected on the basis of their broad compatibility with the theoretical underpinnings of the study as presented in Parts I and II. So for instance statistical studies on the students' performance in calculus, where the skills tested were derivatives of a content-centred and quantitative approach to mathematical learning, were not included. In sum this research review aims at establishing the links between the study and current developments in the field that have been its influences as well as placing the study within the PME-AMT context.

IIIc.i The Novice's Difficulties With the Notion of Infinity

The concept of infinity has been a long-standing mathematical concern since the times when Aristotle made the distinction between actual (for instance the infinity of *N*) and potential (for instance that for every number there is a larger one) infinity. Cantor's theory of infinite sets, in which infinity is defined within his theory of cardinality of sets (two sets are of equal cardinality if there can be defined a one-to-one correspondence between them, thus a set *A* is infinite if there is a one-to-one correspondence between *A* and *N*) stands against the mystique surrounding ∞ for centuries: mathematicians either treated it as a real number trying to evade the contradictions that this was leading to (for instance Euler) or, while recognising it, attributed to it literary and metaphysical properties (Descartes' 'recognisable but not comprehensible' infinity). In sum potentiality of infinity seemed to be more acceptable than actuality.

Understandably then infinity is a concept puzzling to the novice learner. Actually it appears as if the notion of infinity remains puzzling even as mathematical experience accumulates — see for instance Martin and Wheeler's findings (1987) on preservice teachers' conceptions of infinity: inconsistent between the various algebraic and geometrical settings and not generalisable. In Fischbein et al's (1979) terminology, potential infinity, which seems to dominate the concept image of infinity, is a primary intuition deeply embedded in our common epistemological heritage as opposed to Cantor's conception which is a secondary intuition, embedded in the specialists' experiences. The novice's learning seems to require a shift from the primary to the secondary perceptions of infinity and the evidence given below suggests that the shift is problematic.

Piaget, in his studies of the young learners' development of number, found (Piaget and Inhelder 1956) that 11 or 12 years of age children appeared to believe

> in the possibility of indefinite bisection of a figure,
> that the end result would be a point and
> that the original figure could be reconstituted from its 'constituent' points.

The first two reflect a parallel development of actual and potential infinity but the third shakes the evidence of that because it shows that children have not acquired the non-extensive notion of a point. Similar results, but more heavily context-bounded, were presented by Taback in 1969 where children did conceive of indefinite reiteration. This doubt expressed by Piaget's findings is also reflected in Langford

(1987) and Fischbein et al who also acknowledge potential infinity as a priority in the learner's development.

Tirosh (1991) also reports on the students' profound conflicts between their conceptions of infinity, based on finite events and objects, and the notion of actual infinity. These conceptions are deeply ingrained, resist change and are context sensitive. Most students' potential infinity conceptions come to conflict with the taught notions of actual infinity. The children's experiences are mostly about things growing larger or smaller, not with the notion of transfinite cardinal numbers. She then lists some of the students' intuitive criteria for comparing infinite quantities:

- students do not use the Cantorian idea of 1-1 correspondence to compare infinite quantities
- because of their notion of infinity as inexhaustibility they think of all infinite sets as the same
- they transfer finite methods to handling problems of infinity
- the above generate contradictions that most of the students are unaware of
- the generated conflict is between their intuitive conceptions and formal instruction.

Tall (1980) explains further the counterintuitive nature of Cantor's theory for the learner — comparing sets is not a large part of their experience — and stresses that except for Cantor's ordinal and cardinal infinity (extending measuring via comparison of sets) there is also the notion of infinity, originating in non-standard analysis, according to which infinity is defined as a generalised notion of measuring in an ordered field larger than the real numbers. He terms this measuring infinity and proposes, along with Fischbein, Tirosh and Hess (1979), that, as an extrapolation of the learner's arithmetical experience, measuring infinity can more easily lead to a perception of the potential infinity of the limiting process.

Tall (1981) also explored the students' various intuitions of infinity and reported that, within the context of limits for instance, students find it hard to reconcile ideas such as $-\infty$ and $+\infty$; or that both x^2+4 and $2x^3+3$ tend to infinity but the latter tends 'faster' since $(x^2+4) / (2x^3+3)$ tends to zero as x grows. Sierpinska (1987) also in her study of the students' conceptions of the limiting process found the students' perceptions of limit closely linked to their perceptions of infinity. She classified the students' according to their conceptions of infinity as

- unconscious infinitists — using the word infinity but in fact referring to very large or very small quantities,
- conscious infinitists — using the word infinity in a metaphysical and evasive manner. They act as if infinity does not really have a place in mathematics and they tend to avoid references to infinity.
- kinetic infinitists — infinity is connected to the idea of time as either theoretically achievable but not necessarily tangible (potentialists) or actually achievable (actual potentialists).

In the following I present a concise review of developments with regard to the novices' perceptions of the concept of function and of concepts related to calculus.

IIIc.ii The Novice's Difficulties With the Concept of Function

I note that most of the works referred to here are from research on function in an analytical, calculus context; also that the notion of function as encountered in this study goes beyond the calculus contexts to linear mappings and homo/isomorphisms in Topology and Linear/Abstract Algebra. However, as it will become obvious in coming chapters, the results prove valuable and possibly generalisable for all domains.

Function, as commonly known through the Bourbaki definition (a correspondence between two sets which assigns one element of the second set to each one of the elements in the first set), plays a central and unifying role in mathematics. However, as the evidence suggests in this section, most students do not generally associate function with this formal definition and in response to problem situations, they usually call up a variety of concept images. These images are numerous and I briefly outline them in the following.

Functions (Nardi 1992) are seen as rules with regularities; this is the 'function as formula' idea. Functions are often identified with just one representation - either the symbolic or the graphical. A change in the independent variable is seen as causing a change in the dependent variable, with the consequence that constant functions are often not considered functions. The vertical line test is used almost exclusively in determining whether a given example is a function and even functions expressed via polar co-ordinates are tested this way. Students are reluctant to employ graphs and interpret graphical information poorly. Monk (1990) points out that the novices can answer 'point-wise questions' fairly well but have trouble with 'across-time questions'. Many of these difficulties result from discrepancies between the mathematical definition of the function, the concept definition a student remembers, the concept image of the student and the part of the concept image evoked when solving a problem.

Function is a concept that has been evolving in the last 4000 years, from a complex network of conceptions as a graph, a formula, a relationship, an input-output machine to the unified and rigorous Bourbaki set-theoretic definition. Didactical interest in the concept was raised as a result of the learning difficulties with the Bourbaki definition that were observed when it was used as an implementation of the New Math approach to the mathematics curriculum. Malik (1980) identifies the gap between Bourbaki definition and the rule-based relationship between dependent and independent variables as a source of these difficulties and Sierpinska contended that the latter, as opposed to the former, is fundamental for understanding functions within the context of formulas and graphs.

Dreyfus and Vinner (1989), in their application of the *concept image - concept definition* schema on the learning of function, suggest that concept images are not simply formed by definitions but by experiences. This explains the diversity of concept images associated to the concept of function: a correspondence between two variables, a rule of correspondence, a manipulation or operation (put one number, take another), formula/algebraic term/equation, graph. Additional diversity and refinement of these images is suggested by the students' images when given graphs of functions (also in Barnes 1988; Markovits et al 1986 and Vinner 1983): a graph is continuous, otherwise it is considered that it 'changes its character', or the domain of the function splits, or there is an exceptional point. So continuity seems to be perceived as an inherent characteristic of a function and cases of discontinuity are seen as anomalies of an exceptional, hence secondary character. Markovits et al, for instance, blame the emphasis put by school mathematics on the rule or relationship that is part of the definition of the functions that students are mostly familiar with at the expense of the equally important notions of domain and range.

The same authors also account for the students' biased domination of the function concept image by straight lines as an effect of the extreme emphasis to linearity in the school mathematics curriculum. The students' persistence of linearity is evident in the experiment in which students were presented with two points on a Cartesian graph and asked them to define a function whose graph passes through these two points. Most students drew the line connecting the two points, and, when given scattered points, some of them still tried to fit linear functions. To Markovits et al's contextual interpretation of the students' persistent linear images (curriculum emphasis) I also add the tendency of the human mind, suggested by the gestaltist psychology, to prefer the obvious simplicity of harmonious and symmetrical images such as straight lines.

In Barnes' experiment students claimed that $y=4$ is not a function because x does not depend on y. When shown the graph of $y=4$, a line parallel to the x-axis, most changed their minds. Students from the same group then claimed that $x^2+y^2=1$ is a function because it's a circle; also that a function with a split domain is not really one function but several. In general 'strange' or broken, that is non smooth and continuous graphs, are not easily considered as functions. A large number of these conceptions reflect, for instance, that a function has a continuous graph, ideas that were held among mathematicians for a long time until concepts were axiomatically defined and terms started having a regulated meaning in the mathematical community. Some of these perceptions however transcend the regulated meaning and permeate through our concept image as tacit conventions. There is a number of these conventions, or as Tall calls them (1992), these 'creases in our minds that we attach to graphs and we tacitly expect learners to adapt to'.

A major source of the learners' difficulties with functions is their lack of flexibility in switching representations or working on the relationships between them. Dreyfus and Eisenberg (1983) demonstrated evidence of difficulties with several transformations, such as stretches and shifts, as well as with change of variables (or change of domain). In sum students tend to view algebraic data and graphical data as being independent.

In Ferrini-Mundy and Graham's study first semester calculus students (1991) 'were not able to provide any type of general definition of function but readily gave examples of functions by writing formulas. There is little evidence that the students see functions as objects of study in mathematics; rather, when a function is given, in equation form, usually one is expected to do something to it, such as substitute in a value. This part of studying functions (plugging in values) seems to be firmly established and becomes their way of working with other calculus concepts such as limit'. Remarkably similar tendencies are evident in the behaviour of more advanced learners such as prospective mathematics teachers, reports Even (1993): in her study the function is seen as always defined by a formula, continuous, and 'reasonable'.

In the above the major difficulty encountered by students seems to be to shift between different representations of function and establish connections between its static (object) and dynamic aspects of its nature. This duality (Sfard 1991) in the nature of the concept, which is part of its epistemological power, seems to be cognitively distracting and leading to conflicts or to disjoint perceptions, to a schism between the two. In the various contexts where the concept of function is explored in this study this discussion on the duality of its nature will reappear as essentially relevant.

IIIc.iii The Novice's Difficulties With Calculus

Calculus is the mathematical area that didactically has been the most efficiently researched and talked about in the context of the discourse on the Calculus Reform, the revitalised interest in advanced mathematical skills fuelled by the realisation of the importance of mathematics in the technology oriented, post-industrialist societies. Here I present very briefly some findings that relate to the parts of calculus mentioned in the analytical chapters of this study.

Ferrini-Mundy and Graham (1991) offer evidence that the students' understanding of central calculus concepts is exceptionally 'primitive': students demonstrate virtually no intuition about the concepts and processes of calculus; they diligently mimic examples and their attempts to adapt prior knowledge to a new situation usually result in very persistent and often inadequate conceptions whose change the students firmly resist.

The concept of limit, a concept which is central in the study of calculus and appears consistently in a variety of contexts, seems to be at the heart of the didactical enquiries in calculus. First of all the concept of limit seems to be accompanied by a certain amount of mystique: there are no computational recipes for finding limits. Then its understanding impinges upon a very complex network of ideas and an equally complex novel notation.

Jere Confrey as quoted in (Ferrini-Mundy & Graham 1991) argues that the students' introduction to calculus is likely to happen as follows: with a discrete understanding of continuous ideas (for instance, time as a series of instants, motion as a series of positions), with an independent transition to the idea of continuity, and with an algorithmic approach. Also due to the vernacular uses of the word *limit*, students find it hard to distinguish between limit and bound. Moreover *tends to, approaches, gets close to* are expressions underlain by the assumption that a limit cannot be attained. Frequently concept images of function also interfere in finding limits: for instance when 0 1 0 1/2 0 1/3 0... is taken as two functions, not one.

The words associated with the concept of limit seem to conjure up ideas that feed a rich and complex concept image. Davis and Vinner (1986) say about parts of the concept image of limit that they 'cannot be evoked instantaneously in complete and mature form'. As a result 'some parts of the idea will get adequate representations before other parts will' and this is their explanation of what they perceive as the inevitability of some obstacles. They then suggest that part of learning is about making these obstacles visible and conscious: concept images of limit are dominated by examples, for instance; convergent sequences are mostly seen as monotonic ones; another strong concept image is the dynamic notion of limit as a value where the terms of a sequence approach.

The latter has the implication of what Tall terms *generic limit property*: all the properties of the terms of the sequence also hold for its limit. I note here that Leibniz, as well as Cauchy, believed that the limiting function of a family of continuous functions must be continuous. As a result students believe lim0.999....<1 even when they can prove that $\sum 9/10^n=1$: 1 is not of the form $9/10^n$, so it cannot be the limit of $\sum 9/10^n$. Cornu reports that students in this case also claim that $\sum 9/10^n$ tends to 0.999..., but has the limit of 1. So in a sense a linguistically informal expression seems to reinforce the ambiguity of the notion in a way that the students can actually evade the confrontation with their contradictory ideas. Graphical representations such as the representation of limit through an illustration of the limiting procedure as a staircase have also been reported as prone to evoking ambiguous perceptions of the concept.

Similarly to linguistic metaphors, notational representations associated with limits seem to generate unanticipated ideas to the students: for instance ε, dx, δx also evoke ideas about numbers smaller than all positive reals but not equal to zero. There is evidence (Tall & Schwarzenberger 1978; Orton 1983a and b; Cornu 1980 and 1982)) that these symbols interfere in building up concept images from the definition.

Robert (1982) offers a taxonomy of the students' conceptions of the limit of a sequence — except the formal definition: monotonic and dynamic monotonic (limit associated with the monotonicity of the sequence), dynamic (limit associated with approaching, closeness, tendency), static (terms around the limit or close to the limit), limit as bound. She also reports that students seem to have a diversity of images that co-exist in their concept image: some of these images are stronger; they dominate and hence they are evoked more easily. Remarkably the concept definition is not the strongest of these images, even when the students seem to have acquired it. Tall and Vinner (1981) found that the students who recollect a dynamic informal definition are more likely to recollect it correctly than the ones who attempt to recollect the formal definition.

Schwarzenberger and Tall (1978) suggest that a common informal interpretation of the formal definition of $\lim s_n = l$ is 'we can make s_n as close to l as we please by making n sufficiently large' where 'close' means near but not coincident. Graham and Ferrini-Mundy also noted that students calculate $\lim f$ more successfully when f is continuous: they cannot give geometric understandings, they cannot use the graphs, they see the limiting process as an evaluation process and they think the δ–ε notation is redundant.

In her work on the epistemological obstacles related to limits, Sierpinska (1985) embedded more generally the students' difficulties in their attitudes towards mathematics and infinity. Sierpinska and Viwegier (1989) studied two young pupils' conceptions of infinity and found that 'concrete' conceptions of mathematical objects do not prevent one from being able to perform precise deductive reasonings based on assumptions not necessarily conforming to one's intuitions. The reasons she gives for this behaviour are the superficiality of intuitions in the learner's transition from the Piagetian concrete to the formal operational stages of development. Later when intuitions become linked to emotions, they turn into convictions or beliefs and possibly 'start to function as obstacles'.

Concrete procedures are reported as more comprehensively understood also in the context of integration and differentiation. Orton (1983a) showed that routine aspects of differentiation are often well-understood and that basic techniques of integration are applied with some facility. However errors like calculating the derivative of the quotient as the quotient of the derivatives still occur and they reflect gaps in conceptual understanding and predominance of some often inadequate intuitions. Orton claims that the idea of proportionality is fundamental in understanding the derivative and he attributes the students' limited understanding (of the tangent as the limit of a set of secants; of the rate of change of a line versus rate of change of a curve; and of the rate of change at a point versus the rate of change over an interval) to gaps in the students' conceptions of proportionality. As for the integral (Orton 1983b) students do not seem to understand the limiting process inherent in the dissection of area or volume in the calculation of integrals. In sum students tend to interpret differentiation and integration as processes of indefinite approximation. This reflects the influence of dominant conceptions of potential infinity as opposed to conceptions of actual infinity which would support conceptions of differentiation and integration in which a value for the integral and the derivative can be attained.

Finally in Artigue's 3-year experiment with differential equations (1987), students found it hard to reason with functions not given explicitly — as a formula —, to work simultaneously with two representations (namely seek graphical solutions when given algebraically presented differential equations) and to validate graphical solutions requiring what seemed to them as rather sophisticated tools from analysis.

IIIc.iv The Novice's Difficulties With Linear and Abstract Algebra

In Chapters 8 and 9 an opportunity is given to study the novice's cognition of a wide variety of algebraic topics: most predominantly related to Vector Spaces and Groups. Research interest in these topics is relatively new and a large number of the studies reviewed here are observational studies in which the teaching and learning of these topics merge into a whole of theorising practice: one or a few algebraic problems are given to undergraduates or a new teaching technique is implemented (usually related to current educational technology). Observations are made about the novices' learning and their difficulties. Finally empirical explanations are attempted. Despite their theoretical frailty these studies provide relevant and valuable illumination on the cognition of Algebra. In the following I cite some of these studies.

Orit Hazzan (1994), accounting for the students' repeatedly observed belief that in a group

$$x*x^{-1}=e \Rightarrow x=e,$$

notes that students 'borrow properties' from $\Re$. This possibly reflects their need of a familiar metaphor to associate with the axiomatic definition of operation *. Linguistically calling $a*b$ the 'product' of elements a and b in a group reinforces the illegitimate use of the $\Re$ metaphor (by an analogous linguistic token she also points at the similar sounding of the Hebrew words for 'keep' and 'preserve'. Students then reinterpret 'an isomorphism preserves the group operation' as keeping the properties of usual operations). Moreover representing elements a and b of a group with different letters may evoke the idea that a and b are two distinct objects. Finally Hazzan claims that students tend to confuse a theorem and its converse. So, because if $x=e$ then $x^2=e$, students also may think that if $x^2=e$ then $x=e$.

In a similar empirical vein Carducci (1993), benefiting from the visualising powers of Mathematica, attempted to question some of her students' persistent intuitions regarding properties of determinants. In the process she observed that, even when required to work in abstract environments, students demonstrate a partiality for some matrix operations (matrix inverse and scalar multiplication). In another IT experiment with ISETL, inspired by the students' frustration with lectures and the low retaining of knowledge even shortly after the lectures, Leron and Dubinsky

(1995) note that the construction of meaning 'is at the heart of students' difficulties in abstract algebra' and that the students' difficulty with 'fairly simple' relationships between certain mathematical objects (such as the Homomorphism Theorem or Lagrange's Theorem) should not be attributed to the complexity of the theorems but to the abstract nature of the mathematical objects involved. They then give the example of Lagrange's Theorem which is about 'easy' things such as 'one number dividing the other, two sets having the same cardinality or being disjoint, etc.':

> *The objects, on the other hand are 'complicated' in the sense of the many levels of abstraction and the great time and effort needed to 'construct' them in the mind of the student. For us the simplicity of the proof lies in our ability to have a clear image of the group as being partitioned into a disjoint union of cosets. But in order for the students to have such an image, they need to 'construct' in their mind not only 'group', 'subgroup' and 'coset' but also 'the set of all cosets of H in G'.*

So, they conclude,

> *the students' difficulty with understanding Lagrange's theorem may be largely due to their confusion about the nature of cosets. We find that they can understand the process of forming a coset, but often cannot take the next step of seeing these cosets as objects to be measured, counted and compared.*

Their didactical strategy then revolves around helping students to construct cosets and learn to manipulate them as objects. Harel (1989) a few years earlier also proposed that the teaching of Linear Algebra is based on false assumptions with regard to the students' ability to deal with abstract structures without extensive preparation or to appreciate the economy of thought characterising abstract representations. He then cites examples from school mathematics where the domains of algebraic applications are specific and embedded either in a numerical or a visually accessible geometric context. Like Dubinsky and Leron above, he also recognises the abstract nature of the mathematical objects involved in Linear Algebra as well as the not-always-emphasised multiplicity of its domains of application, as the sources of the students' difficulty. Even when introduced through a few examples, Harel concludes, the Linear Algebra abstract concepts rarely acquire a meaningful reason of existence in the students' minds.

Similarly to Harel, Robert and Schwarzenberger point at the concept of vector space and the concept of group as examples of new mental objects the construction of which causes fundamental difficulties in the transition from elementary to advanced mathematics. In tune with Tucker's historical account of Linear Algebra (1993), they note that 'the problems from which these concepts arose in an essential manner are not accessible to students who are beginning to study (and expected to understand) the concepts today': this historically decontextualised presentation deprives the novices of a potentially meaningful construction of these concepts.

Dorier et al (1994) elaborate on the above, recognising first that, because students do 'not really have to bring the concepts themselves into play', the results of final exams may very well hide the poverty of the students' concept image construction. They then suggest the following explanations:

- the specific epistemological nature of the concepts of Linear Algebra:

> *These concepts [vector space, linear operators, image, kernel, basis, dependence, dimensions, rank, etc.] took their formal shape only after numerous uses of linear methods in specific contexts without much unification. These concepts, in their present form, are the cause and results of unification and generalisation. Their relevance appears only a posteriori, as they first renewed in an economical manner the solving of old problems and only afterwards allowed new approaches.*

So the learners have difficulties learning these concepts because

> *they only have access to the final phase of the historical process: definition of the concept and systematic use in the solving of problems. Yet the problems they are asked to solve may often be solved with specific tools and methods, more familiar which do not necessarily imply the use of the new concepts. In this case the simplification and improvement induced, in the solving of the problems, by a change in the point of view cannot be foreseen by the students. They may only do this because they are asked to; this is an effect of the teaching contract.*
>
> *On a more theoretical basis, this implies that these problems, which can be solved in various settings, cannot yet be chosen as 'good' problems in the perspective of the tool/object dialectic [...], as the concepts to be taught are not indispensable for the solving. The only problems which would necessitate an absolute use of the formal concepts of linear algebra are all too complicated for our students (they involve non-countable infinite dimension vector space)...*
>
> *In other words for the students the concepts of Linear Algebra are above all objects before they can be used as tools; students are therefore deprived of the long progression, which brought mathematicians to express these concepts, step by step.*
>
> *...[Students have] elements of knowledge which initially are not well enough distributed in different settings (geometrical, analytical, logical, formal settings mainly). [some analyses of students' practices] reveal the lack of ability to change settings and points of view.*

Dorier et al's views can be generalised to other mathematical contexts. An example is Nicholson's paper on the historical development of the concept of quotient group which appears implicitly in the works of many 19th century mathematicians, including Galois and Jordan, but only as a context-embedded tool, not formally defined. Despite Dedekind's early understanding of the power and potential of the formalisation of the concept since the 1850s, it is not until the end of the century that Hölder systematically and explicitly defines the quotient group. It is noteworthy that Dedekind's lectures were lukewarmly received (members of the mathematically strong audience admitted they understood little); that is it seems that from a phylogenetic point of view the concept had not completed its cycle of fermentation. As a result professional mathematicians of the time resisted its significance and its acceptance through formalisation.

Harel and Kaput (1991) embed some of the difficulties cited above by Harel in their theory of Object-Valued Operators (an Object-Valued Operator is, for instance, the correspondence between parameters and functions in $f(x)=\sin(ax)$):

> *...students usually had difficulty dealing with such a correspondence, unless they were able to tag the outputs of the correspondence with familiar geometric figures, such as lines or planes [...]. These geometric figures, which were manipulable objects for the students, apparently helped the students to construct such a correspondence as an object-valued operator.*
>
> *Another common example involves the construction in abstract algebra of the quotient object associated with a 'normal' sub-object, e.g. in the case of groups. The cosets must be conceived as objects if they are to participate as elements of a group. However the existence of a 'representative element' for a coset, where the operation defined on cosets can be given in terms of an operation on their representatives, makes it possible to deal successfully with many aspects of the quotient group on a symbol manipulation level without treating the subsets of a group as objects, or even as subsets. Students' inadequate conceptions are revealed when one asks them to attempt to create a group using a non-normal subgroup's cosets - they often cannot understand why the subsets 'fall apart' when they attempt to multiply them together as 'sets', or by using representatives.*
>
> (Harel & Kaput 1991, p87)

Hillel and Sierpinska (1994) attribute the students' difficulties with learning Linear Algebra partly to their 'inexperience with proofs and proof-based theories. Moreover espousing Piaget's and Garcia's notion of intra-, inter- and trans- level of knowing something, they note that students learning Linear Algebra 'need to operate at the trans- level'. So generality and abstraction coupled with the students inter-level elementary experiences yields problematic cognitive behaviour. The researchers also include in their list of sources of difficulty, things that are not generic to Linear Algebra but relate to advanced mathematics in general, such as 'not understanding the need for proofs nor the various proof techniques, not being able to deal with the often implicit quantifiers, confusing necessary and sufficient conditions'. Below I present Selden and Selden's (1987) list of 'errors and misconceptions' regarding theorem proving in Abstract Algebra which reinforce Hillel's and Sierpinska's observation about the relevance of more general difficulties in advanced mathematical reasoning.

Selden and Selden's taxonomy probably fits better alongside some of the observational studies previously cited in this section, since they are based on their experience from a Moore-type Abstract Algebra course but I mention it here since it helps in placing the learning difficulties with Abstract Algebra in the wider advanced mathematical context. The major 'misconceptions' observed in their students include:

M1. Beginning with the conclusion.
M2. Names confer existence ('failure to distinguish between symbols for things whose existence is established and symbols for things whose existence is not').
M3. Apparent differences are real (failure to see things with different names are not necessarily different).
M4. Using the converse of a theorem.
M5. Real number laws are universal (mentioned also before).
M6. Conservation of relationships.
M7. Element - Set interchanges (difficulty in understanding statements involving sets results in substituting them with statements about the elements of these sets).

Selden and Selden also mention: overextended symbols (the use of the same symbol for two distinct things), weakening the theorem, notational inflexibility, misuse of theorems, circularity, locally unintelligible proofs, substituting with abandon, holes in the implication that links two statements and using information out of context. They then recognise that their taxonomy addresses a diversity of cognitive issues which they sum up as the learner's difficulty with: generalisation, use of theorems, notation and symbols, nature of proofs and quantification.

Back to Hillel and Sierpinska, specifically to Linear Algebra these authors add to Dorier et al's interpretations, cited previously, the 'shuffling back and forth' between different levels of description: within the discourse of Linear Algebra three languages coexist,

the language of the general theory (vector spaces, subspaces, dimensions etc.)
the language of $\Re^n$ (n-tuples, matrices, rank etc.)
the geometric language of $\Re^2$ and $\Re^3$ (orthogonality etc.),

that 'are often interchangeable but are certainly not equivalent'. Most emphatically Sierpinska and Hillel raise the problematic issue of 'representing a linear operator in a basis and moving from one such representation to another'. They contend that understanding this representation operation requires that the learner sees the language of Linear Algebra as a 'network of languages and rules of translation between them'. Following Foucault's distinction between language as part of the world it describes and language as a representative system of signs, they suggest that

> *the problem of understanding the language of linear algebra as a representation rather than as being part of the world lies in the fact that the world of linear algebra is indeed the world of simultaneously used systems of representation. How does one make sure that two externally different representations indeed represent one and the same 'thing'?*

In that sense when students encounter for the first time vectors in $\Re^n$, 'strings of numbers' become to them the 'primary representation', the 'thing'. This view of a vector as identical with one of its representations is profoundly shaken when the student reaches the realisation that this 'string of numbers' represents different vectors in different bases and that the same vector is represented by different 'strings of numbers' in different bases:

> *Strings of numbers, so familiar, so palpable, suddenly feel like 'ghosts of departed quantities'.*

Recurrent mistakes, such as taking the first column of the matrix of a linear operator as the image of the first basis vector under the operator (correct in the canonical basis), must be attributed, contend these authors, to a more deeply ingrained conceptual difficulty than simply to the Piagetian difficulty with internalising an activity as an operation. To this purpose they offer the linguistic scheme briefly outlined above, namely the need to distinguish vectors from their representations as well as reach the 'trans-level' of seeing linear operators as 'objects of inquiry themselves'.

Leron, Hazzan and Zazkis (1994), in a study preceding the first major effort of systematising learning difficulties with regard to Abstract Algebra (Dubinsky et al 1994), interviewed a small number of students on their conceptions of Group

Isomorphism. As Harel and others above, they stress that students, accustomed to the affectively secure routine algorithmic behaviour of school mathematics, find it difficult to engage in an existential process of constructing an abstract object. Moreover the existence of more than one possibility also clashes with their previous school experience of mathematical problems having one and only one solution. Decomposing the notion of an isomorphism, they note that it involves two especially complex notions: function and an existential quantifier. They identify three levels of internalisation of the existential quantifier $\exists$ that seem to influence their students' conceptions of isomorphisms: the personalised action level, the process level and the object level. The more objectified the conception of a quantifier is, the more appropriated the students' existential activity turns out to be. A similar tripartite levelling of the students' conception of function seems to influence the students' attitudes towards isomorphisms. In addition the authors note the importance of talking about isomorphisms necessarily in terms of their domain and range (which again clashes with the students' previous experiences of talking about functions mostly as procedures or formulas regardless of domain and range).

Finally they state that

> *the very concept of isomorphism is but a formal expression of many general ideas about similarity and difference, most notably the idea that two things which are different, may be viewed as similar under an appropriate act of abstraction.*

Remarkably the notion of function emerges as dominant in advanced mathematical cognition and indeed in much more abstract contexts than it has been mostly studied (see IIIc.ii of this Chapter). In a similar vein in Zazkis' experiment (1992), students, while finding the inverse of a given compound element, repeatedly claimed that $(XoY)^{-1}$ is $X^{-1}oY^{-1}$ instead of $Y^{-1}oX^{-1}$. Zazkis draws on misapplication of linearity and specifically over-generalisation of distributivity as the dominant sources of the error and cites classical algebraic and trigonometric examples of this behaviour. Given that her experiment was carried out in a computer environment, Zazkis also notes that learners do not always 'perceive the [computer] environment as mathematical and therefore they use their 'naive' knowledge and extrapolation techniques, rather than 'formal mathematical knowledge''. She finally employs a proceptual argument in order to indicate the flexibility in shifting from the process to the object aspects of a concept as a sign of mathematical maturity: this flexibility is a requirement for successful handling of tasks like finding the inverse of a compound element.

I close the section on Linear and Abstract Algebra with a reference to the above mentioned first major attempt at a systematic presentation of observations regarding learning difficulties with regard to the notions of group, subgroup, coset, normality and quotient group. As several researchers referred to in this section, Dubinsky et al (1994) also present their observations in terms of the action - process - object framework. In the following I summarise their findings.

Their didactical interest in Abstract Algebra, and in particular Group Theory, was triggered off by the reported failure of the novices to achieve good understanding of the concepts involved in introductory courses. In most cases the novice's formal encounter with the notion of group is the first one in a long series of abstract concepts. Accompanying the problematic novelty of the experience of abstraction is the historical antecedent of the epistemological complexity of the concepts, examples of which were given above in the reference to Nicholson's paper.

Developing the Concepts of Group and Subgroup. In sum the developmental view of these concepts presented in their paper is the following:

> *Our observations are consistent with a progression in understanding that moves through various intermediate (and incomplete) ways of understanding groups and subgroups... from seeing [them] as primarily sets of discrete elements, to a stage where the operations as well as the group elements are incorporated into the necessary definition [to]...a thorough understanding of a group as an object to which actions can be applied.*

In their study the participants seemed to develop the concepts of group and subgroup in parallel. The 'psychogenesis' of these concepts appeared to be 'linear' but also at many points 'in concert with the others'. Often progress with regard to one concept awaits developments in the others. Also other mathematical concepts, most notably Set and Function, are fundamental in these developments. More specifically: at first the most primitive conception of a group is 'based entirely on the student's conception of a set'. Progressively properties of this set are included in this conception; a binary operation maybe one of them and it is crucial when this property is singled out. At this moment the notion of function, and in particular of the two-variable function is needed in order to accommodate the notion of a binary operation. Ideally this process is completed with the encapsulation of the set and the binary operation into the notion of group and the reification of the pair so that the construction of the notion of an isomorphism is possible. It is noteworthy that the procedural aspects of the binary operation seem to attract more of the learner's attention than probably allowed in the above developmental path.

Moreover the notion of function appears strongly in the concept formation of subgroup as a restriction of a function to a subset of its domain. If the role of the binary operation has not been clearly understood it is likely that the learner confuses subgroup with subset.

Developing the Concepts of Coset and Normality. In general the construction of cosets in simple groups appears as a simple task. What seems to be extremely difficult is the conceptualisation of quotient group and normality; also understanding that the quotient group is isomorphic with another familiar group. Difficulty varies in the context of different groups but the overall impression is one of deep confusion. Normality is also confused with commutativity and the students' ignoring the condition of normality for the existence of the quotient group influences heavily their attempts to construct the cosets.

Probably the most problematic concept is the Quotient Group. In terms of the action - process - object schema, a major issue in the psychogenesis of the concept of coset and coset operation appears to be

> *the encapsulation of the process of forming cosets into objects which are to be the elements of the quotient group. Again the student's conception of function appears to be needed, but confusion seems to have occurred when he or she made an attempt to construct a binary operation on cosets before being able to manipulate these cosets as objects. An interesting point is that in some cases, formation of a coset did not guarantee that the student could not deal with this as an object. This suggests that the student's conception of sets may not have been adequate.*

In this last section Linear and Abstract Algebra were presented as two exceptionally difficult introductory topics for novices in which difficulties related to a wide range of basic concepts seems to be condensed. Issues of rigour and intuition, visualisation, notation and language were also raised. The findings of this study, as demonstrated in Chapters 6-10, complement or at some instances extend the findings reported in this Part of the Chapter. In Chapter 2 the theoretical background of the study's methodology is presented, so that in the first two chapters the study is embedded within the current thematic and methodological developments in the field.

Summary

The theoretical origins of the study lie in the realisation that an educational reform regarding mathematics teaching cannot take place in the absence of an awareness of the learner's thought processes. Coupled with the intrinsically idiosyncratic epistemological complexity of mathematics, the cognitive dimension of didactics arises as particularly significant.

With regard to learning advanced mathematics this study originates in the assumption, grounded on the relevant literature, that a novice mathematician faces a series of cognitive difficulties in the encounter with mathematical abstraction. Abstraction is meant both from a psychological perspective i.e. that the advanced mathematics learner has to build up knowledge in an axiomatic way and learn how to reason deductively; and from an epistemological perspective, i.e. that the nature of the objects of advanced mathematical learning can extend beyond the physical or the numerical.

In the above, learning is not seen as isolated in a cognitive vacuum but embedded in a sociocultural context. Therefore, in a constructivist strand of thinking, the learner's cognition, while being personal and individually interesting, is also emphatically seen as taking place in a learning environment.

If, as R.B.Davis contends (1989), 'theory building is the trademark of science' then this study of the learner's difficulties ought to be carried out in such a way that it leads to the enhancement of theory in the field of Didactics of Mathematics. This study seeks to construct a psychological profile of the novices' difficulties in their encounter with mathematical abstraction by probing into their expressions of learning. It is assumed here that the phenomena of cognition, an inaccessible and esoteric process, can only become visible and accessible through the learners' oral and written (in this study: oral) articulations of their mathematical thinking. In effect of the above in Chapter 2 the study is presented as a phenomenological study of advanced mathematical cognition.

Chapter 2

The *Methodological Approach* of the Study

Introduction

'Collegiate mathematics education' is, in Selden and Selden's words (1993), 'at a pre-paradigm stage', that is exploratory and seeking the construction of theoretical frames. Only recently has the field started being recognised as an autonomous field: Batanero et al (1994) report the teaching of mathematics at the undergraduate level as one of the fields attached to which there is a developing number of projects. Its publications appear scattered in general mathematics education journals and conferences and, as noted by Kaput and Dubinsky in the introduction of (1994), the field is in a transition: from identification of the phenomena of learning to interpretation and action, from the locality of studying particular mathematical topics/concepts to the global confrontation of the problematique of advanced mathematical cognition.

In other words the field seems currently to be shifting from the improvisational amateurism of quick diagnosis-prescription to informed professionalism in its confrontation of the phenomena of learning. Becker and Pence (1994) also point out the transitional stage of the field and subsequently categorise undergraduate students' learning as an increasingly autonomous area of research. This field of research, within which this study is located, has been denoted in Chapter 1 as PME-AMT.

Thomas Kuhn (1962), speaking of the 'route to normal science', offers an account which, I think, summarises vividly the state of the art in PME-AMT. To him 'normal science' means 'research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice'. These works 'define the legitimate problems and methods of a research field for succeeding generations of practitioners' and share two characteristics:

> *Their achievement was sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity. Simultaneously, it was sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve.*
>
> (Kuhn 1962, p.10)

He then calls the achievements that share these characteristics 'paradigms'. Paradigms relate closely to 'normal science' which he defines as 'some accepted examples of actual scientific practice' which 'include law, theory, application and instrumentation together' and 'provide models from which spring particular coherent traditions of scientific research'. Researchers then 'whose research is based on shared paradigms are committed to the same rules and standards for scientific practice. That commitment and the apparent consensus it produces are prerequisites for normal science, i.e., for the genesis and continuation of a particular research tradition'. In his further clarification of the concepts of normal science and paradigm, Kuhn notes that

> *there can be a sort of scientific research without paradigms, or without any so unequivocal and so binding...Acquisition of a paradigm and of the more esoteric type of research it permits is a sign of maturity in the developments of any given scientific field.*
>
> (ibid., p.11)

The usual developmental pattern for mature science is competition among paradigms followed by successive transitions via revolutions. This pluralism is indicative of how 'science develops before it acquires its first universally received paradigm'. These competing paradigms all possess 'components of real scientific theories, of theories that had been drawn in part from experiment and observation and that partially determined the choice and interpretation of additional problems undertaken in research'. Kuhn (and this work was first published in 1962) also notes:

> *...and it remains an open question what parts of social science have yet acquired such paradigms at all. History suggests that the road to a firm research consensus is extraordinarily arduous.*
>
> (ibid., p.15)

He then highlights some of the reasons for the difficulties in this road:

> *In the absence of a paradigm or some candidate for paradigm, all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. As a result, early fact-gathering is a far more nearly random activity than the one that subsequent scientific development makes familiar. Furthermore, in the absence of a reason for seeking some particular form of more recondite information, early fact-gathering is usually restricted to the wealth of data that lie ready to hand.*
>
> (ibid., p.15)

Kuhn then is sceptical about the possibility that this 'sort of fact-collecting' may 'produce a morass'. He is worried about whether 'facts collected with so little guidance from pre-established theories speak with sufficient clarity to permit the emergence of a first paradigm'. He however recognises that this complex and chaotic practice is essential to the building of scientific foundations. As a result of this practice it is

> *no wonder, then, that in the early stages of development of any science different men confronting the same range of phenomena, but not usually the same particular phenomena, describe and interpret them in different ways. What is surprising, and perhaps also unique in its degree to the fields we call science, is that such initial divergences should ever largely disappear.*
>
> (ibid., p.17)

The conditions in which these divergences disappear, if they do, are provided by the emergence of a theory 'better than its competitors' which 'need not, and in fact never does, explain all the facts with which it can be confronted'. Then both fact collection and theory articulation become highly directed activities. 'Truth' as Frances Bacon acutely noted, 'emerges more readily from error than from confusion'.

The reason I have so extensively quoted Kuhn is that, again in his words, 'it is hard to find another criterion that so clearly proclaims a field a science'. Also I find his account in complete resonance with the methodological and thematic ambience within PME-AMT which I consider as a field that is in search of its 'normal science' or, as worded earlier, 'at a pre-paradigm phase'. The methodology employed in this study reflects Kuhn's description of the 'fact-collecting' practice of a pre-paradigmatic field: the methodological tools of this study are chosen in a way that aspires at the identification, exploration and interpretation of the phenomena of undergraduate mathematics learning. The study is also underlain by an intention to submit these tools to evaluative testing (in tune with Jacob (1987) who acknowledges research 'on adapting qualitative traditions to the study of naturally occurring cognitive behaviour in classrooms' as one of the 'most exciting areas' of future methodological research). So in this sense research within a pre-paradigmatic field is required to work on these two levels: the thematic and the methodological.

As a result, this study is a piece of qualitative research and, in particular, it is a phenomenological study of advanced mathematical cognition. The theory that is generated from this study is the outcome of a data-grounded theory emergence process. In the following the theoretical underpinnings of this declaration are clarified. So:

- in Part I an account is given of the phenomenological character of the study, the cognitive nature of the phenomena to be explored and the learning environment within which they are explored,

- in Part II a theoretical justification is given of the methodological data collection techniques of observation and interviewing, and,
- in Part III a theoretical justification is given of the methodological data analysis techniques of data grounded theory.

PART I The Phenomenological Character of the Study, the Cognitive Nature of the Phenomena to Be Explored and the Learning Environment of the Exploration

In Chapter 1, particularly Part III, focus was placed on these developments within PME-AMT which are about the students' difficulties with advanced mathematical concepts and reasoning. Here I specify in more detail the nature of the learning phenomena that this study aims to investigate, the environment of the investigation and in the subsequent parts I describe the methodological tools that were employed in this investigation.

Ia. The Phenomenological Character of the Study

To achieve its purpose this study has adopted a strongly phenomenological approach. Here the *phenomena*, that is the things apprehended by the senses, on which *typifications* will be assigned are the novice's *verbal expressions* of their thought processes during tutorials and interviews.

Phenomenology (Burrell & Morgan 1979), that is the study of direct experience at face value, as a theoretical standpoint believes in the importance of subjective consciousness that is active and meaning bestowing. To comprehend the structures of consciousness Husserl encourages the act of *transcendence*, an *Epoché* during which the observer is freed from all preconceptions about the observed phenomena by questioning their taken for granted features. In this study, this transcendence is achieved by the constant clarification of the assumptions made and by adopting a very open series of data collection techniques such as unsystematic observation and loosely structured interviewing. In this sense Chapter 1 aims at declaring and clarifying the principles underlying the observation of cognitive phenomena and the techniques described in Part II of this Chapter are chosen because they allow the effortless emergence of the features of these phenomena.

Schutz (Burrell & Morgan 1979) suggests that, once this transcendence has been achieved, then meaning can be assigned to the observed phenomena retrospectively. That is, once observation of the phenomena (apprehension by the senses) is completed in the transcendental manner outlined above, then a meaning-bestowing process can begin through which the observer makes sense of the phenomena. In this study this means that theorising begins, once cognition has been observed in its full naturality and complexity and once the principles of observation and interpretation have been declared. This position is reflected in the theory emerging procedures that determine the data analysis.

An idea that is particularly significant in determining the methodological approaches of this study comes from the linguistic, ethnomethodological tradition (Garfinkel 1968). What is under investigation here is the novice's thinking processes from the perspective of the assumptions made, the tacit meaning attributed to external stimuli, the conventions utilised and the practices the novice adopts as these are reflected mainly in verbal and secondarily in written expressions. It is assumed here that the learners' indexical expressions in their interaction with peers, their tutor or the interviewer convey much more than it is actually said. *Indexical expressions* here is used as a term for the designations assigned to a stimulus so that this is located, labelled and interpreted. As seen in Chapter 1, these expressions are interesting and worth exploring because, in an interactionist frame of thought (Mead 1934) the learners continuously act — and in an educational study it is important to stress that they also *inter*act — on the basis of their psychology as formed partly by these attributed meanings. So access to the learners' thinking, and consequently a consideration of the didactical implications of such a study, can only be possible if the learners' expression of their thinking is understood.

The typifications over the learner's expressions in this investigation will be data-grounded generated theory in the sense Glaser and Strauss (1967) use the term (see Part III).

I note that the use of the term 'phenomena' in this study is not to be confused with its phenomenographic use. Unlike Piaget and the constructivists, phenomenographic research, for instance as advocated by Marton (1988), is underlain by the sharp distinction between the knower and the objects of knowing. Phenomenographic cognitive phenomena then, in Marton's terms, are the observable relations between knowledge and the knower. As a result cognition can be described in a limited and finite number of ways: students' conceptions can be presented as an ordered, hierarchical structure (outcome space) and are not psychological, that is not context-specific and generalisable. The epistemological, psychological and sociocultural characteristics of this study are in opposition to this description of mathematical cognition. Below I outline briefly how cognition is perceived in this study through the words of Jacques Lacan.

Brown, Hardy & Wilson (1993) in transferring some of Lacan's ideas to inquiries about learning seem to resist the finitist ideas of phenomenography according to which phenomena are describable in a finite number of categories. These categories are, for the phenomenographist, descriptions of the learner's images of a mathematical idea which act

> *...as the totalisation of the chain of signifiers...as a condensation of the symbolisations, abstractions, connections, illustrations, equivalences brought together within that key idea...From the image the journey through the mathematical activity can be evoked... This is symptomatic of the ideology that you can capture the human activity of learning in a list of statements.*
>
> (Brown et al. 1993)

Expressing then their belief that however elaborate a description of the learning phenomena is, the sum is always greater than its parts, they claim that what is lacking is 'something that cannot be expressed in language or symbols or mathematical images'. This is 'the elusive reality of mathematics'. They then quote Zizek who interprets Lacan and his motif of symbolisation...

> *...as a process which mortifies, drains, empties, craves the fullness of the real of the living body. But the Real is at the same time the product, the remainder, left over, scraps of this process of symbolisation, the remnants, the excess which escapes symbolisation and is as such produced by the symbolisation itself.*
>
> (ibid.)

In Chapter 1 mathematical activity was acknowledged as both personal and social. Rephrasing appropriately for this study Brown et al's quoting of Gattegno's 'only awareness is educable' to 'only awareness is communicable', I agree with them that 'the move [from the personal] into the social is always rooted in an attempt to label the personal'. But then, the same authors wonder, 'how do I articulate my awareness, my shifting from conscious to unconscious, into something usable' in communication?

Here is where symbolism and language come in as illustrated in the Lacanian strong metaphor below. I note that Lacan (1977) believed in the interpretation of the subconscious as a language:

> *The world I see becomes captured in language and as I seek to be even more refined in my describing ... I enter the world of mathematics.*
> *'Each new word is a step away from the Mother'*
> *Mathematics, the fantasy of maximising distance from the Mother, is the story that suppresses the Mother. But Mother is not so easily dismissed! I experience the world through my senses, those I had before I learnt to speak. In making sense I describe this world through retroactive naming. To talk mathematics I need to use the language of the tribe if I am to communicate, in a quest to be accepted by the Father. Mother, whilst still there is never*

> *quite the same as I capture more of her in the language of the Father. The Mother becomes the dumping ground, the other not captured within the structuring, the ignored left over after the stressing. Rumbling away under the surface of a regulated and described vision, the unseen and the unsayable, the subconscious biological murmurings.*
>
> *Not so much then 'I think therefore I am' but rather 'I speak therefore I construct'. The 'unified Cartesian subject' has become the main casualty of post-modernism being replaced by a subject analysable as a process, inextricably linked to a context which is itself a process. This subject, held in the successive stories told about him, can never be fully constituted since closure is always in the future; any descriptor is part of a chain that is never finished, the subject himself, acting as if he is the one in the mirror, is forever disappointed by the world resisting his actions in a slightly unexpected way. Nevertheless the stories he tells give structure to that which he describes and reflexively give structure and position to he who speaks, bringing self and world into union in a common inherited language that of the Father.*
>
> (ibid.)

In this study the observed phenomena, the novices' verbal expressions of their conscious thinking are seen from the point of view of a departure from the primitive language of pre-university mathematics to the 'language of the tribe', the formalism of advanced mathematical thinking. Within the observed learning context, this departure is a constructive process which is far from smooth and unproblematic. The separation from the Mother, in other words, and the entrance in the world of the Father is a process of initiation into the culture of the tribe through the communication of language, a process of enculturation. The signs of this process are the students' expressions of 'retroactive naming' of the world they are trying to understand, give structure to and take position in.

Sexton (1988), who studied problem solving thought processes through the students' expressions of their thinking, notes about the reliability of these expressions:

> *An often asked question is ' how can we be sure that what a student says is really what s/he is thinking?'. In other words, are the processes which a student uses to solve a problem altered by verbalisation? We cannot be sure but the work of Ericsson and Simon (1984) suggests that talking aloud and thinking aloud does not alter the problem solving processes. These researchers contend that the student only reports the information that s/he heeds as s/he works a problem, that vocalising does not change what is being heeded. However, it is possible that some thoughts occur so rapidly (even automatically) that they are not heeded and hence not reported. Such accounts [...] would result perhaps in incomplete reports of thought processes but not in inaccurate reports. [my underlining]*
>
> (Sexton 1988)

This 'elusive reality' of mathematical cognition, this approximation of thought processes achieved through linguistic expression, will be discussed further in Part Ib, the section on the links of the study with Cognitive Psychology.

Ib. The Cognitive Nature of the Phenomena to Be Explored — Links with the Psychology of Cognition

As far as research techniques are concerned, this study espouses some of the approaches used in Cognitive Psychology. In 1981 Greer was somewhat concerned with the absence of strong links between Cognitive Psychology and Mathematics Education since he thought that work was done in parallel in the two fields where interaction and exchange of ideas would be mutually helpful. This was 15 years ago but his outline of the reasons why a Cognitive Psyhology approach might apply to studies of mathematical thinking still holds.

> (1)Mathematical processes, by nature, are amenable to representation by information-processing models, since they break down into sequences of operations, transformations, logical steps etc.. There is an insidious danger here of assuming that the formal expressions of these sequences necessarily mirror cognitive processes.
>
> (2) The role of imagery in thinking is a current focus of interest for cognitive psychologists...
>
> (3) The general notion of different representations of a given problem, and translations between them, is a shared interest...

In the same vein Vergnaud (1990) notes that 'cognitive and developmental psychology are certainly essential in that they really question what a concept is; what an operational behaviour is; how they develop; what part is played by action, perception, and language in concept formation; and what part is played by social interaction.'

Greer goes on to stress that cognitive psychologists, unlike behaviourists, are essentially interested in the mental processes that intervene between stimuli and responses. 'Theories', he contends, 'about these mental processes have to be developed on the basis of indirect inferences from observed behaviour (which may include what the subject says about their own thinking)'. One of the methods he illustrates, evolved for this purpose, is the use of verbal protocols. In his discussion he raises the issue of the validity of an account given by a cognising subject of their own thought processes. He continues: '...asking subjects to describe their own thought processes... is likely to yield much more insight, though the findings must be treated with circumspection because of subjectivity in interpreting them, ...and because of the lack of control and standardisation. There is also the unavoidable problem that the act of verbalising cognitive processes itself affects those processes, and there is the whole question of validity, i.e. whether subjects can report accurately on their own thinking ... '. He then gives examples of sceptics, like Evans, and of keen users of the approach, like Dominowski who studied concept learning.

Despite the traces of strong positivism, which is not compatible with the perspective of this study, I have used Greer's words to highlight some of the methodological constraints underlying an analysis of the novice mathematician's expressions of their thinking processes. While I am at more ease than Greer in recognising the situational character of these verbal expressions and the context-embeddedness of the interpretation, I am also aware of the risk that cognising subjects cannot necessarily report accurately on their thought processes and also that, while reporting them, these processes may be altered.

As illustrated in the next sections of this Chapter and in Chapters 3, 4 and 5, the methodology of this study is designed with the intention to capture evidence of the students' thought processes — while the students are engaged in cognitive activities — in its least artificial form: naturalistic observation and semi-structured interviewing are the tools of data collection and I note that the students are rarely asked by the tutors directly in the tutorials to explain their thought processes. They are mostly asked to explain their actions. In the interviews I always asked them to talk in detail about a number of mathematical concepts or theorems. Therefore the study is designed to avoid the artificiality and the risk of distortion that questions the validity of the cognising subjects' accounts of their own thought processes.

The approach taken in this study is a combination of elements from the phenomenological and cognitive psychology approaches outlined above: this is an exploration of the novice mathematician's thought processes as reflected in the novice's indexical expressions and in their accounts of their own thought processes. It is because, as Balacheff points out in (1990a), 'borrowing the theoretical framework' from psychological theories on thinking is not a way to cope sufficiently with issues of knowing in a teaching/learning situation, that a multi-disciplinary approach is required.

In the above what seems to be implicitly assumed is the accessibility of cognitive structures. This however should be done with modesty. As Balacheff poignantly stresses

> *...we must give up the notion that what we observe is in some sense isomorphic to what the observed individual experiences. We must even give up the notion that what an individual expresses is necessarily a reflection of his or her beliefs — some beliefs may be so deeply embedded in an implicit world view as to be inarticulable. The collection, interpretation, and analysis of data must become interactive processes informed by both researchers and individual informants. Such research will require us to examine and re-examine what it means to understand one another's conceptions, to be scientific in our research and, finally, to engage ourselves in the reflective process of examining our own research agenda and paradigms.*
>
> (Balacheff 1990a)

On a more methodologically specific note he suggests that

> *...the nature of learners' conceptions can be traced not only from what learners state explicitly but also from the way they use them and from the class of problems these conceptions allow them to solve.*
>
> (ibid.)

Clearly observing learners while in the process of knowing in the 'informed' way Balacheff illustrates is a potentially illuminating approach as far as this knowing is concerned. Similarly Dubinsky and Lewin, referring to the act of thought, note that,

> *...the act itself remains inaccessible and idiosyncratic, dependent on the particular way in which a given subject notices and organises his/her experience. It would seem one never has direct access to cognitive processes — thought is an unconscious activity of mind — but, at best, only to what an individual can articulate or demonstrate at the moment of insight itself. Precisely what occurs at that moment seems as inaccessible as it is essential.*
>
> (Dubinsky & Lewin 1986)

In this study it has been assumed that access to cognitive structures can be relatively achieved by means of an extensive and close observation of the learner in action. As the last quotation points out, encouraging metacognition on the part of the observer and of the learner enhances the chances of success. In this study opportunity to encourage metacognition on the part of the learner was given during interviewing. However the main bulk of data for this study has its origins in a learning environment on which interventions were neither possible nor intended. In the following I present the Oxford tutorial as the selected natural learning environment in which the cognitive phenomena elaborated upon in this study were explored.

Ic. The Necessity to Study the Novice Mathematician's Thought Processes in a Natural Learning Environment. Tutorials as the Natural Learning Environment of This Study

Theory building is, as Davis says (1989), the *'trademark of science'*. However, as illustrated in the introduction to this chapter, PME-AMT is at a pre-paradigm phase. Therefore unlike in mature disciplines, such as the natural sciences, where theory generating tools are sharp and accurate enough to yield theory and subsequently engage in a process of verification, in disciplines at a pre-paradigm phase a hypothesis testing approach is not feasible due to the absence of testable hypotheses. Hence the theory generated from this study is data-grounded. Data has been obtained by turning to potentially rich sources of evidence: in other words it is suggested that in an inquiry that aims at the study of the novice mathematician's problematic encounter with mathematical abstraction, a learning context within which this encounter takes place must be identified and subsequently provide the raw evidence of the searched-for phenomena. It is assumed that this approach that is at the same time conscious of preconceptions and informed by previous developments can be an efficient one. In Part Ia this approach was described in more detail.

For the reasons outlined above one of the primary concerns of this study was to identify contexts where the novice mathematician's learning takes place. Initial considerations included the prospects of studying mathematicians' written accounts of their own thinking, triggered by Jacques Hadamard's work (1954). Given that

- mathematicians tend to focus on the outcome of their thinking,
- generally their accounts are superficial or even metaphysical and finally,
- these are accounts of learners at another level — after all this is a study of the idiosyncrasies in the thinking process of a learner at the brink of concrete mathematical thinking who is about to face mathematical abstraction —,

the search for a learning context continued elsewhere.

Next, the study of the novices' written work was considered. The inadequacies of this approach are also clear: written work is a form of monologue on the part of the learner and can only give an in-depth account of the learner's first unnegotiated (with peers, with tutors) response to a mathematical idea. To achieve access to the learner's evolving train of thought, one must trigger its expression, its visible manifestation. Given the prerequisite for the naturality of this process, as explained earlier in this Part, it was deemed that the first year of mathematical studies at university level, provides a conveniently organised context. In particular the individual or pair tutorials, given to students on a weekly basis constitute a habitat in which it is most likely that the students are offered a forum for expressing themselves mathematically.

In Chapters 3, 4 and 5 a detailed account is given of why a tutorial is a rich source of raw data that satisfies the purposes of this study, but here I outline briefly its basic features that are essential in comprehending the process of selecting and using observation and interviewing which will be elaborated upon in Part II: a tutorial, as semi-officially defined by the mathematics tutors participating in this study, is usually a 30-60 minute session the main focus of which is resolving the learner's queries as well as a series of other activities, such as presentation of new extra-curricular material (new definitions, theorems and proofs that were either omitted or not quite emphasised in the lectures or the problem sheets) and mainly exposition on the solutions for the questions in the problem sheets. Given that this material is the same for all first year mathematics undergraduates, one can thus guarantee the relatively uniform basis of the observation sessions. Depending on the degree of permissiveness on the part of the tutor — as to replacing the traditional monologue with a more conversational teaching style — and of openness on the part of the students — as to how exposed they allow themselves to be with regard to their mathematical understanding — observing tutorials can be variably productive of incidents on mathematical thinking. In this sense tutorials can be ideal for the purposes of this inquiry.

In sum this is a piece of qualitative research. As such, in order to justify its theorisation process, the definition of scientific validity needs to be widened so that vertical studies (ones that probe into the depths of the explored phenomena with regard to a small number of participants) share the traditional prestige of horizontal ones (mostly statistically valid studies that explore a large number of variables over large samples of participants). Sierpinska et al (1993), Abbott-Chapman (1993) and the authors of *Research Issues in Undergraduate Mathematics Learning* (Kaput & Dubinsky 1994), in their texts on what is research in mathematics education and what are its results, also make the point: in mathematics education the debate between qualitative and quantitative research is an unnecessary and false dichotomy. Each methodology is underlain by a number of theoretical assumptions and governed by its own technical rules: as long as researchers make these rules and assumptions competently overt and explicit, and as long as they accept that the validity of their findings is also subject to/ yielded from these rules, the two traditions can develop in parallel. Ideally a merge of quantitative and qualitative approaches is achieved in cases where the hypotheses tested in a quantitative piece of research have been generated via a qualitative approach.

PART II Data Collection Methodology. Unsystematic Observation and Semi-Structured Interviewing

In the following I present a concise theoretical profile of minimally participant, unsystematic observation and semi-structured interviewing, the methodological techniques employed in this study.

IIa. The rationale of the Data Collection Methodology

The rationale behind the selection of minimally participant, unsystematic observation and semi-structured interviewing as the main methodological techniques of this study, lies in the conceptualisation of the study as a data-grounded theory generating project. The theoretical origins of this concept lie in the work of Glaser and Strauss (1967; Strauss 1990) whose ideas I present below, and in Part III, as they have been filtered through the data collection and analysis processes.

In brief (see Part III for more details) Data Grounded Theory is a cumulative plan for progressive building up from facts: from data to substantive theory and then to more conceptual forms and formal theory. First however it is necessary to specify the features and qualities of the data that constitute the body of substantial evidence on which the theory generation process is grounded. To do so I outline below how Data Collection for this study is reconciled with the principles of theoretical sampling in the Glaser and Strauss theory.

Theoretical sampling — or *naturalistic sampling* as Ball calls it in (Hammersley 1993) — is the process of data collection for generating theory whereby the analyst jointly collects, codes and analyses data, decides which data to collect and where to find them in order to develop emerging theory. Initial decisions are not based on a firmly preconceived theoretical framework. A partial framework of local concepts is the basis and the analyst is open and sensitive to influences from a variety of sources. This sensitivity is valuable since it secures data collection from the exertion of power from one specific preconceived theoretical frame.

Data collection in this study did not take place within a firm theoretical frame. The focus of the study is conceived as a general address of an issue that seems to be significant in the field — the novice mathematician's problematic encounter with mathematical abstraction. The local concepts that are enrolled into the study come from the research literature on difficulties that the novice has in particular mathematical topics and in mathematical reasoning (see Chapter 1 Part III). The psychological background of the study is also diverse and locally valid (see Chapter 1 Part II). In other words this remarkable lack of unified theory in the field, in a way, forces an open-mindedness on the analyst who addresses the relevant issues. This open-mindedness is liberating and excruciating at the same time. Why this is so will become manifest in the course of the theory emerging process.

Previously in this chapter the nature of the sought-for evidence has been outlined (the indexical expressions of the novice's difficulties in the encounter with mathematical abstraction) as well as the context within which the inquiry takes place (tutorials). Some characteristics of the *Main Study* (such as the number of participants) were determined on the basis of the *Pilot Study's* experience. In the *Pilot*, as explained in Chapter 3, it became evident that the relatively uniform mathematical content of the tutorials as well as their regularity (once a week) provided a fairly uniform learning environment for observing a group of students in cognitive action. To use Glaser and Strauss terminology, this uniformity makes theoretical prediction possible.

The nature of the sought-for evidence and the context of inquiry coupled with the openness imposed by the theoretical espousals of the study led to determining the intended technical features of the observation and interviewing techniques. In subsequent chapters the implementation of these intentions is discussed.

Here it is necessary to stress that the observation techniques used in the two stages of the study *Pilot* and *Main*: details follow in chapters 3-5) were substantially different. In the latter the use of audio recording guaranteed an accurate account of the events and thus gave me, as the observer, the latitude to make less descriptive and potentially more insightful observations. These observations operated as pivots for the primary stages of data analysis. The differences between the two phases of observation are elaborated in Chapters 3 and 4. Here I outline the intended use of observation in the *Main* Study.

The use of the term 'unsystematic observation' contains a potential misunderstanding (Medley, Mitzel & Gage 1963): entering a very rich learning environment, such as a mathematics tutorial, with the intentions to 'observe unsystematically' does not imply that the observer has not clarified, at least generally, the dimensions of the observed events that are of particular interest to the inquiry (Cohen & Manion

1989). The existence of a primary focus is crucial (Van Dalen 1966); the degree of specificity may vary (Anderson 1995). Here the focus was the psychological aspects of the novice's learning as expressed in the context of the tutorial and the degree of specificity was low. The former made the inquiry relatively tight; the latter theoretically allows the emergence of some cognitive phenomena with regard to the novice's transition to abstract mathematical thinking that under a tighter focus might have remained unnoticed. By the same token (Anderson 1995) the interviews, carried out during the *Main* Study, were chosen to be semi-structured and broadly based on some critical observations on the tutorials. Details on their content follow in Chapter 4. The scope of mathematical topics these interviews cover as well as the openness of the questions regarding these mathematical topics are again indications of the low specificity of focus selected for the data collection of this study.

IIb. The Features of Unsystematic Minimally-Participant Observation that Served the Purposes of this Study

A term which relates fittingly the aims of this study with its data collection is Naturalistic Observation (Fraenkel & Wallen 1990). Previously in this chapter the need to explore advanced mathematical cognition in a 'natural learning environment' was explained. An assumption underlying the selection of Unsystematic Minimally-Participant Observation is that it provides an account of mathematical learning of an unprecedented, strong internal validity (Anderson 1995). In Chapter 4, details are given of how participation of the observer was kept to minimal levels; also an evaluation of the technique as used in this study is provided.

As far as the external validity of an account based on Unsystematic Minimally-Participant Observation is concerned ((Van Dalen 1966) and (Merriam 1988)), as explained below, the aspirations of a qualitative study is to provide an accurate and profound account of intensely context-embedded phenomena. The intricacy of the task lies in giving an account whose detail and acuteness simultaneously highlight the idiosyncrasy as well as the generality of the situation. This requires that the observer is well aware of the generalities established in other pieces of research and that, at the time of observation (as well as at the subsequent stages of analysis) the observer is able to draw appropriate abstractions without stripping the situation of its specificity. Observation provides material of immense diversity, density and complexity which the observer and analyst is required to engage in disentangling. As repeatedly stressed in this chapter this is where the strengths and the weaknesses of this methodology lie (Mercer & Walford 1991). The details on the use of Unsystematic Minimally-Participant Observation given in Chapter 4 explain how some weaknesses were restrained.

IIc. The Features of Semi-Structured Clinical Interviewing that Served the Purposes of this Study

As explained in the presentation of the *Main* Study in subsequent chapters, the interviews with the participants play a supportive role in the analysis of the main bulk of material that was produced via observation. Their loose structure was based on a selection of some striking difficulties of the participants with certain mathematical topics identified during tutorial observation (details are given in Chapter 4). Further investigation of these difficulties via interviewing was decided on the grounds of the following characteristics of clinical interviewing as identified by Jean Piaget (Ginsburg 1981; Piaget & Inhelder 1963) : despite its artificial setting a clinical interview is designed in a way that allows the identification, exploration and evaluation of the interviewee's thought processes. The first two aims are relevant to this study:

- to identify the participant's thought processes the interviewer sets an open-ended task, asks questions in a contingent manner and requests a considerable amount of reflection on the part of the interviewee
- to explore thought processes, the interviewer intends to facilitate rich verbalisations on the part of the interviewee. Thinking is a complex process which is not revealed by simple responses; extensive verbal expression is more informative. At the same time discourse analysis demonstrates how verbalisations do not necessarily mirror identically thought processes. In this sense triangulation, standardisation of findings (Borg 1963), is necessary and can be achieved either within clinical interviewing or via other methods. For instance in this study the co-ordinated use of naturalistic observation as well as clinical interviewing aims at fulfilling triangulating purposes.

Moreover a clinical interview aims at clarifying the ambiguity of verbal statements as well as checking out alternative explanations of the interviewee's conceptions, images and reasoning. Contingency and open-endedness are definitive characteristics of the clinical approach. In Chapter 4 details are given on how the clinical interviews made in this study secured the fulfilment of the above characteristics.

IId. Theoretical Sampling in the Study and a Deviation From the Glaser and Strauss Plan Regarding Theoretical Saturation

In this study there has been a substantial deviation from the Glaser and Strauss plan for data collection: the only forms of data processing that took place during data collection were *making of critical/evaluative notes during observation* and constructing descriptive summaries of the events in each tutorial on the same day, that is *construction of Scripts*. Therefore the decision when to stop sampling, which would normally, in the Glaser and Strauss plan, be indicated by theoretical saturation — that happens when no additional data develops any further the properties of the category — was informed by the experience of the *Pilot* study and external factors such as time constraints of the study and recommendations by more experienced colleagues in the field. What was continually corrected however was the note and script making process. Details on that are given in Chapter 4 but briefly I would describe these corrections as geared towards giving a less judgmental and gradually more focused account.

It must also be made clear that the sense in which the term *sampling* is used here is quite distinct from the statistical use of the term: the primary aim here is not to obtain accurate evidence on a wide distribution of people but to discover categories and properties. Working on a small sample of Oxford undergraduates, that have been chosen through procedures of mutual agreement and volunteering is a process of theoretical sampling. The aim of the selection of participants here is not statistical representativity but gaining access to individuals who are willing to share their mathematical thinking in a way that will help the emergence of rich theoretical categories regarding the novice's mathematical cognition. In this sense the sampling of this study has been systematic (Hammersley 1990).

Finally, in Part III I outline briefly the principles underlying the data-analysis of the study.

PART III Data Analysis Methodology. Data-Grounded Theory

In this last part of the chapter an account is given of the principles underlying the methodological techniques espoused in the data analysis of the study. Details on how these principles were put into practice are given in Chapter 5. Here the data analysis methodology is presented as a data-grounded theory emerging process in which a blend of techniques has been used from a diversity of methodological traditions: most notably an inductive approach to analysis of qualitative data and discourse analysis. A number of texts on the methods of analysis in qualitative research have been consulted, most notably (Adelman 1981; Bliss, Monk, & Ogborn 1983; Hitchcock & Hughes 1991; Lofland & Lofland 1995; Merriam 1988 and Miles & Huberman 1984), and have influenced the formation of techniques presented in Chapter 5.

According to Glaser and Strauss effective theory should enable explanation, advance theory in the field, be usable in practical applications and provide a stance towards future data and style of research. Glaser and Strauss use the term 'comparative analysis' for a method of generating theory through drawing patterns or making deductions that underlie a set of analytical units. Typical uses of comparison include: fact replication with comparative evidence and spotting the indicators of the conceptual category the event in question belongs to, drawing empirical generalisations and specifying concepts.

The process of generating theory is closely linked with its form of presentation. A difference between data-grounded generated theories and logicodeductive theories is that the former is generally more prone to a discussional form of presentation as opposed to the propositional form. It must be noted though that the propositional form allows easier and possibly riskier leaps to deductions. As shown in subsequent chapters, the presentation of the study has been designed in a way that allows the inductive approach of its theory generation to become visible (details in Chapter 5).

Further characteristics of the study that have determined its analytical strategies originate in Discourse Analysis and also the theory of Epistemological Obstacles.

Links to Discourse Analysis. This is a study which aims at gaining access to thought processes through the participants' verbalisations (earlier called verbal or indexical expressions). This has led to the adoption of certain techniques often attached to the methodological tradition of discourse analysis (Dijk 1985 and Coulthard 1985). As demonstrated in Chapter 5 where particular qualitative techniques employed in the study are presented in detail, the main directive of the processing of the raw material has been towards the extraction of data relevant to the aims of the study. Successive filtering has resulted in a compilation of episodes which are all relevant to the novices' advanced mathematical cognition and are underlain by the iterated reappearance of a certain number of cognitive phenomena. These phenomena, or in Vergnaud's words these conceptual fields (1990), are the categories that Glaser and Strauss describe as expected to emerge from this inductive process.

Links to the Theory of Epistemological Obstacles. Mathematical learning as explicated in Chapter 1 is perceived in this study to evolve as a process of constant confrontation of epistemological obstacles. As a result the Theory of Epistemological Obstacles as conceived by Bachelard and revived and refined by Brousseau *'directs the thought'* of this piece of research. So, according to Sierpinska (1994) to whom these words belong, in that sense Epistemological Obstacles emerge as another dimension of the *'categories'* mentioned above. She describes the somewhat elusive nature of a category:

> *It is possible that the most characteristic feature of a category is that it is hard to grasp with a definition, difficult to enclose within a rigid theory. A category does not belong to the world of theories; if it functions the way it does — by directing the thought — it is because it works somewhere between and above the vernacular and the research field. It is better described by the use that was made of it in research, what questions did it lead to, what explanations did it provide, what kind of discourse has developed around it.*
>
> (Sierpinska 1994, p.134)

The categories that have 'directed the thought' and emerged from this study fulfil Sierpinska's description: the study is

- driven by issues relating to advanced mathematical cognition,
- directed by a frame of mind that describes learning as a process of conflict and confrontation of difficulties, and,
- leading to the formation of conceptual categories regarding advanced mathematical cognition.

In Chapters 1 and 2 the conceptualisation of the study has been presented. In the subsequent chapters the *Pilot* and the *Main* Study are presented as its realisation.

Chapter 3
The Pilot Study

Introduction

In this chapter I present an account of the *Pilot Study*. The presentation is in chronological order even though it is inevitable that particular emphasis is put retrospectively on certain aspects of the study that have turned out to be significant in the light of the *Main Study*. The observation period is described as well as the main phases of data processing. Findings are also briefly presented. The overall presentation is geared to sustaining the points made in the final section where the formative influence of the *Pilot* on the *Main Study* is reviewed.

Summary: From October 1992 to June 1993 ten first-year mathematics students were observed during tutorials given to pairs of them on Analysis, Topology and Linear Algebra. Fieldnotes were taken during observation. Coding and categorising the data followed. Subsequent focus was on the category constituted by the students' difficulties regarding mathematical understanding. Its contents were further processed and analysed. A triad of difficulties — topical, logical and symbolic — was used as a descriptive tool and provided the order of presentation in the analysis. The study exerted strong influence on determining the methodological and thematic structure of the *Main Study*.

PART I Data Collection And Data Processing

In the following an account is given of the data collection and the data processing period of the *Pilot Study.*

Ia. Data Collection

During the academic year 1992/3 ten first year mathematics students, all in the same college in Oxford and tutored by the same tutor, were observed during the weekly 30-minute tutorials that were given to pairs of them. Fieldnotes were taken during observation of the tutorials. My presence in the sessions was first negotiated with the tutor, in whose office the sessions were taking place, and then with the students. In a short introductory discussion that I had with each pair, the non-evaluative nature of my inquiry was emphasised as well as the right of the volunteers to withdraw from the study at any stage or to ask me to pause note-making or leave the office on specific occasions. I orally guaranteed confidentiality and promised to keep the identity of the volunteers concealed in all types of reference to instances from the tutorials. Moreover some information was given to the volunteers on my research interests. Apart from a few volunteers who expressed a stronger interest in my work and requested further details — after the conclusion of the observation period — the tutor and the students were told generally that I work on *'how people understand mathematics and the problems they have'*. From informal conversations with the volunteers it can be deduced that the impression remained unaltered until the end of the observation period. After fieldwork was completed, the tutor and some students asked me what would '*become of all these notes*'. My reply was that I would *'scrutinise the notes in order to find out about the students' learning difficulties with some concepts'*. Then I would try and *'see whether there are any patterns in these difficulties that repeat themselves in various mathematical topics'*.

The observation technique employed in these tutorials was, as stated in Chapter 2, minimally-participant. This 'fly on the wall' approach, the preference for invisibility, on the part of the observer was also explained to the volunteers. It is virtually impossible to measure the degree of influence of the observer's presence. In the pilot-study tutor's words, however, the students' behaviour did not change. Later, during the *Main Study*, other tutors commented on the students' behaviour in words such as *'more aware of their mistakes'*, *'keener to talk during the tutorial'* etc. whereas students refer to the tutors' behaviour as *'more polite'*, *'keeping on time'* etc.. No other evidence apart from these informal accounts exists of the observer's influence of presence. In any case these accounts either indicate insignificant influence or influence towards a direction beneficial for the purposes of the study, such as *'both sides more willing to talk about mathematics'*. Finally it must be noted that nobody dropped out of the study and there has been no incident of protest for disturbance either by the tutor or the students.

The *Pilot Study* as a whole signifies a methodological and a thematic shift, or a gradual selective process that is reflected in the form and content of the *Main Study*. The two processes took place in parallel but not independently.

Quite early in the observation sessions my intention was to put the emphasis of observation more on the cognitive than the affective aspects of mathematical learning. Of course it is impossible to discuss issues on mathematical thinking without considering, for instance, the awe or fear with which certain mathematical topics are confronted by the learners; or ignoring the vast role played by motivation; or separating instances of the learner's cognitive behaviour from the tutor's stimulation of these instances. In general the context in which learning is explored, and here this context is the Oxford tutorial, is part of the study too; it becomes in a sense part of the research question. In resonance with this idea in Chapters 4 and 5, special reference will be made to characteristic features of this context. At this point I must stress that the tutor is seen as part of the context too, in fact the tutor's role is a major factor in conditioning the learning context, and so the issues arising from the tutor's influence also merit considerable attention.

The focus however of this inquiry is on the learner's cognitive processes with regard to mathematical thinking. In the term 'cognitive' here two streams of thought converge: the epistemological — concerning the mathematics discussed in the tutorial — and the psychological — concerning the strictly personal ways in which tutor and students construct mathematical knowledge.

While observation was still going on, preliminary analysis of the data started. Data consisted of fieldnotes produced during the tutorials. These fieldnotes were accounts of the events taking place during observation, that is they were rather concise, snapshot-like reproductions of the mathematical as well as the didactical content of the tutorial. Five pairs of students were observed in the same afternoon and the intention of the tutor was to keep the contents of the tutorials more or less uniform. Whatever deviations from her plan took place can be attributed to the students' interventions, that were either digressing completely from the topic the tutor had in mind, or took the conversation to a different direction within the same topic. In any case, the uniformity of content and tutor attitude encountered in the *Pilot Study* was unique and never truly repeated in the tutorials observed in the *Main*

Study, where apart from the relatively pivotal role played by the Mathematical Institute's weekly problem sheets, tutorial content and tutor attitudes were diverse and not following any immediately visible pattern — even though, in retrospect, it would be possible to trace each one of the participating tutor's preferences for some particular tutoring style.

However the *Pilot Study* data, as well as the *Main Study* data from this college and tutor, have an austere, consistently repeated internal structure. The reason for this is that this tutor does not strictly follow the lectures' or the problem sheets' content. For each tutorial the tutor has pre-determined what is to be discussed. Usually it is a theorem, a problem or a concept that has struck her as mathematically important or didactically problematic. For example: the proof of Lagrange's theorem, or a question on finding the Taylor series of a function, or the concept of spanning set. This scheme was occasionally altered in the *Main Study* but in ways that its essence remained the same. More elaboration on this will follow in Chapters 4 and 5.

This firm adherence to the tutorial's internal structure and content — sometimes regardless of particular student needs arising during the sessions — can be attributed to the fact that in a single afternoon five sessions were to take place, four of them consecutively and uninterruptedly. This intense schedule would inevitably have effects on the tutor's quality of teaching. To maintain the quality of her input and her strength, so that, in her words, 'everybody gets the same', she developed throughout years of tutoring this allegiance to uniformity.

The reason I make a special reference to this uniformity is that it has an almost immediate effect on the data. The fieldnotes made during one afternoon, where the observer witnesses the proof of the same theorem for five times in a row, are progressively changing. To illustrate this change I would say that in a continuum, one end of which is 'mathematical content' and the other is 'didactical content', the content of the fieldnotes taken in the five sessions in the same afternoon tends to move from one end to the other. By the end of the fifth session I would be so familiar with the mathematical content of the session that I could give my undivided attention to observing and registering the students' responses and interaction with the tutor. As a result, in the fieldnotes, progressively from the first to the fifth session, fewer and fewer details of the mathematics are given while the didactical account — directly quoting the participants or indirectly describing and commenting on the occurrence — becomes richer. This would not be so but for the uniformity described above.

This variation in the content of the fieldnotes from the five sessions in the same afternoon — gradually less mathematical and more didactical — implies that all data cannot hold the same status when analysis begins. To extract comparable analytical units the fieldnotes were rid of their mathematical elements, which is a relatively easy task, and the purely didactical account was kept. I stress that this necessity to *purify* the fieldnotes arose mostly in the *Pilot Study*, since note-taking was the only source of information for reconstruction of the events in the tutorials. Tape recording replaced note-taking in the *Main Study* as far as the registration of the mathematical content of the tutorials is concerned. This complete reliance upon the fieldnotes for a mathematical as well as a didactical reconstruction of the tutorial events made note-taking during the *Pilot Study* a more versatile task and the execution of this task is itself an interesting research methodological issue.

The subsequent task was to separate exaggerated empathy from the narrative. I note that the doses of uncritical empathy in the data were drastically reduced as experience in observation was mounting up. Personal comments or hasty, on-the-spot interpretations of the events progressively disappear. Words such as 'successfully', 'late', 'upset', 'happy', 'fail', 'bad', 'insensitive', 'smart' etc. that are frequently used in the earlier fieldnotes also disappear in the course of observation.

However, as I explain in the following, I did not make the very most of the uniform structure of the *Pilot*'s tutorials. The reason is that even though the tutorials from this college, as it later became apparent in the *Main Study*, provide complete and autonomous learning episodes, it is impossible to reconstruct these episodes merely from fieldnotes. Gestures, responses to questions, even short pieces of dialogue can be registered and support a reconstructed account. But the missing pieces from the puzzle of the whole picture/episode were counted as too many and too important to make the reconstruction of long episodes reliable. This is one of the reasons that a technique that would enhance the reliability of these reconstructions had to be selected for the *Main Study*. In this case this technique was tape recording and the reasons for this selection are given in Chapters 4 and 5.

The analytical units, therefore, for the *Pilot Study* are mostly short incidents that were deemed significant for the students' learning process. It must be noted that in the transformation of the material into analytical units, only 20 hours of material were used. These were the fieldnotes from Trinity Term (April - June 1993). The reason for working only on Trinity Term material is a practical and a psychological one. Since the ideas of potential data processing approaches started emerging when more than half of the *Pilot Study*'s data collection was done, it was seen as reasonable to implement these ideas on more recent fieldnotes whose freshness in my memory would perhaps make a reconstruction of the events more feasible. Moreover the fieldnotes from Trinity Term are significantly better considering the cumulative experience of the very messy, unfocused and unconsciously judgmental observations of the first months. So the fieldnotes eventually used as data for processing in the *Pilot Study* already bear the lessons of experience from several hours of observation. It is hoped that one of the major benefits from these hours of observation was the cultivation of the skill to control concentration on gradually more specific foci and go beyond a mere mental wandering amongst the several, multi-layered occurrences of the tutorial. So, in a sense, choosing to work on the fieldnotes made in the latest parts of the observation period implies that I have chosen to work on the part of the material that I think is of better quality: more didactically focused, mathematically eloquent and creatively empathetic.

Ib. Data Processing

The major parts of data processing took place immediately after data collection was completed. During observation, nevertheless, and preferably on the same day or at the latest the following day, the fieldnotes were typed, edited and printed. This procedure constituted the first stage of reflection on the data since typing evoked the still fresh memory of the tutorials and made the typed reconstruction richer. The material, on which the analysis given subsequently was done, consists of approximately 30 pages of typed Trinity Term fieldnotes (see for example Appendix A for Chapter 3). The mathematical content has been kept separately for reference.

Reading the typed fieldnotes and comparing with the fieldnotes and the registers of the mathematical content followed. Soon it became evident that observations were revolving around two axes:

- the tutor-theme designated the T-theme: teaching style, philosophy of mathematics and of mathematics teaching etc., and
- the students-theme designated the S-theme: difficulties, intuition, metamathematical worries, originality etc..

Having the advantage of overview (observation was completed by then) I could reasonably claim that in the fieldnotes no particular emphasis could be spotted on either of the themes.

An immediate effect of this is that the collected material could be analysed either from a T or an S-perspective. A more critical concern arising though is that of whether this liberty with the focus of observation was taken at the expense of the tightness and comprehensiveness of the data: the intention of this inquiry is to study the novice mathematician's thought processes and consideration of other issues, such as the ones addressed under the umbrella of the T-theme, is only justifiable if it contributes to the realisation of this intention. This is one of the major methodological outcomes/lessons from the *Pilot Study* that helped in improving the plan and execution of data collection in the *Main Study*. As mentioned above the role of the tutor in the formation of the structure and content of the tutorial is tremendous. A tutorial is the most intimate learning encounter that the novice has in Oxford. In it, there is a bilateral contribution to learning and apparently the tutor is one of the two sources for this contribution. As explained later, the most prominent difference between the note-taking processes of the *Pilot Study* and the *Main Study* is the tightening of the focus of observation in the *Main Study* along the lines of, to use the *Pilot's* jargon, the S-theme.

However the process of categorising followed without any special emphasis on either of the themes. It was intended that categories emerged naturally from thorough scrutiny of the data and that preconceptions would be kept minimal (see Chapter 2 for the theoretical backing of the course of action of the data-processing described here).

Each incident was allocated to a category and all categories were tabulated in a *Chart of Incidents* (see Appendix B for Chapter 3). I note here that despite the naturalness that this categorising process claims to have maintained, a learning incident is apparently a sum much greater than its parts, namely the affective, the cognitive, the epistemological, and so on. The label attached to each incident represents not an attempt for its holistic understanding but addresses what the researcher saw as its dominant feature.

I also note that a closer look at the categories reveals overlappings: there are incidents that can be allocated to more than one of these categories. As suggested in the relevant literature reviewed in Chapter 2, in this case ruthless determination is the cure for uncertainty. Still the issue of how fine the line is between constructing a descriptive tool for ordering the data and an explanatory/interpretive one remains. Looking back to the *Pilot's* categories I find it hard to distinguish between description and explanation/interpretation. The intention at the time was to construct a descriptive tool and the reason why no further categorising took place was that, in any attempt to do so, explanation/ interpretation was gradually taking over and in a rather crude and unsystematic way.

Subsequently, given the intention of the study to explore the novice mathematician's learning difficulties, the categories indexed S.DIFF were subjected to further analysis. Moreover the categories T.EX and S.EX attracted my attention to the role played by tacit or explicit conventions — institutional, linguistic and other — in the students' emotional response, performance and cognitive behaviour before and during exams. This evolved into a focal point for the observation in the *Main Study*.

Elaboration of the S.DIFF categories allowed a series of refinements with regard to the focus of the study to take place. In this preliminary stage of exploring the instances of pathological learning, I began to see the use of the term *pathological* as compatible with my strong conviction that constructing mathematical knowledge is a distinctly personal mental process. My approach was now taking shape towards the study of instances of commonly-recognised-as-flawed mathematical behaviour of the learner, while acknowledging the intersubjective nature of mathematical knowledge and with the intention to establish associations with other learning theories such as the theory of Epistemological Obstacles. Chronologically this conceptualisation of the study was forming in parallel with the theoretical background of the study presented in Chapter 1.

In a sense this inquiry, as is evident in terms like 'difficulties', 'pathological understanding', 'obstacles', 'flawed' etc. is about exploring the common denominator of trouble in the novices' mathematical cognition, including the rare instances in which the novice has a conception which deviates from the commonly expected but denies classification in the above mentioned 'negative' terms. So alternative student approaches or conceptions also became part of the focus of the study. This issue of inclusion was one more way in which the *Pilot* contributed in the refined problematique of the data collection and processing in the *Main Study*.

In Part II of this Chapter I present some examples of the data and data processing that took place within the S.DIFF categories. The triad of difficulties that were identified and further processed consisted of:

- topical difficulties (S.DIFF.TOP), that is difficulties with the learning of particular concepts (function, limit, integral, sup and inf, remainder of the Taylor series and so on),
- logical difficulties (S.DIFF.LOG), that is difficulties related to mathematical reasoning (modus ponens, mathematical induction and so on), and

- symbolic difficulties (S.DIFF.SYMB), that is difficulties related to the interpretation and manipulation of mathematical notation ||, ΣΣ, *ij*-notation, change of variables and so on).

This triad is not free of methodological constraints: by allocating an instance to one of these categories one attaches a preliminary interpretation to it. As explained above that is why categorising did not go further than that. The possibility of further categorising was dismissed as prone to the error of artificial clustering.

Subsequently every instance in the three categories was analysed individually. The analysis was linked with current research in Advanced Mathematical Thinking. In general data processing revealed that the instances were characterised by a balance between

- highlighting elements of the strictly personal way a learner constructs mathematical knowledge and
- recognising some patterns in mathematical cognition, previously identified by other researchers.

Before proceeding to the presentation of some findings I wish to remind the reader of the context in which these incidents took place: the tutorials were given between the 27th of April and the 8th of June 1993 to first year mathematics undergraduates. The ten students are all members of the same college in Oxford and vary in background, age and nationality (age: 18-19 and one 25, public/state schools, 7 British, 1 Italian, 1 Pakistani, 1 Greek). The choice of tutor relied on volunteering — her participation had been negotiated in the first weeks of the Michaelmas Term 1992. The tutorials were held in the tutor's office in college from 14:00 to 17:00 on Tuesday afternoons.

PART II A Sample of Findings

Before presenting a sample of the findings from the *Pilot Study* I would like to stress that the presentation is concise and selective and it aims at giving

- a naturalistic flavour of the kind of data obtained in the tutorials,
- a demonstration of the range of topics and psychological processes encountered in the tutorial context, and finally,

- an account of how the *Pilot's* experience became a bridge between the theoretical conceptualisation of the study and the realisation of the *Main Study*.

The presentation follows the structure indicated by the S.DIFF triad and the allocation of the instances to the various categories served the purpose of primarily ordering the data. I note that the presented instances are diverse and of fragmentary character: these features of the material are discussed in Part III.

In the following, I present a sample of observations from the *Pilot Study* relating to the students' difficulties with some mathematical concepts (Topical Difficulties) discussed in the observed tutorials. Logical and Symbolic Difficulties follow as well as a brief reference to some unclassified but significant other instances.

IIa. Students' Topical Difficulties

The transition from familiar, concrete to abstract mathematical contexts generates a series of difficulties for the novice. Some of these difficulties were observed in these tutorials. The following examples relate to the use of the concept of *function* in the context of mappings between vector spaces in *Vectorial Analysis* and to the introductory concepts of *Topology*, most notably the notion of *compactness*.

> *Example 1: The notion of function in the unfamiliar context of mappings between vector spaces. The mathematical question in the tutorial is about finding the matrix of a mapping between two vector spaces. During the discussion the tutor asks:* 'what if the matrix we found was the zero matrix O? What would that mean for the mapping between the two vector spaces?'. *The student answers that the mapping would* 'map everything to zero'. *A little later the question is repeated for another pair of vector spaces and another mapping. Only this time instead of* **O**, *the tutor asks about* **I**, *the identity matrix. The same student replies* 'everything would go to one'.

Both the tutor's questions here are supposed to be simple. They merely seem to require the student to generalise from

- the notion of the constant real function that maps every real number to zero ($f(x)=0$) and
- the notion of the identity function, the real function that leaves every real number the same ($f(x)=x$),

to the notion of mapping a vector of one vector space to another via a matrix that in the first case it is **O** and in the second it is I. Whereas the first transfer seems to be carried out smoothly, the second is not. A vector that is mapped via the identity matrix remains the same; the student's claim *'everything would go to 1'* is a perplexed interpretation possibly of the kind *'since 0 means everything mapped to 0 then 1 means everything would go to 1'*. It seems that it cannot be taken for granted that the transfer

> from thinking in terms of one variable correspondences between numbers
> to thinking in terms of mappings between vector spaces

— vectors after all are multidimensional and non-numerical mathematical objects — comes natural to the novice learner.

> *Example 2. The notion of compactness.* Topology seems to be one of the areas in which the difficulties of the transition from the concrete to the abstract are mostly apparent. Most evidently the students seem to make no sense of the definition of *compactness*. As a flavour of the basic difficulties the students have with their introduction to Topology, I cite two instances relating to the notion of *set of sets*.
>
> *Instance 1:*
> *X is an open set.*
> *Tutor:* What is an open cover of *X*?...first what is an open cover?
> *Student:* ...is it the union of lots of open sets?
> *Tutor:* ...well, it's a family of open sets.
>
> *Instance 2:*
> *$\{X_i\}$ is a cover for X. This means that X is a subset of UX_i*
> *Student:* Why then X is not in UX_i?

Since the issues raised by the students' questions are elaborated upon in the *Main Study* here I merely cite the instances noting their significance with regard to the kind of evidence of the students' learning difficulties available during tutorial observation. Knowing what a *union of sets* is and what a *subset* is does not imply that the notions of a *cover $\{X_i\}$* and of UX_i will also come naturally.

One of the consequences of the novices' difficulties with the transition from concrete to formal mathematical contexts seems to be that they tend to adhere persistently to the contexts with which they are familiar from earlier mathematical experiences. The following examples relate to the use of the concept of *function* as a mapping in *Vectorial Analysis* and to the definition of a *vector space*.

> *Example 3. The concept of Function in Vectorial Analysis. The same student as in Example 1, but in another instance, suggests, when asked to find the value of c in $f(x) = c$ = constant:* 'let's try a couple of values for x'.

This is a standard practice used when looking for the values of a and b in $f(x) = ax+b$, where f is a linear function. A familiar practice is activated and employed inappropriately. Incidentally this instance is evidence of the students' persistent image of linearity: linear functions and their properties seem to dominate a large part of the novices' concept image of function. Similar observations were raised in these tutorials in cases where the students demonstrated difficulty with replacing variable x with $x+1$ in $f(x)$ and often replied that $f(x+1) = f(x)+1$ or $f(x+1) = f(x)+f(1)$ which are properties held only by some linear functions.

> *Example 4. The definition of Vector Space. The students seem constantly to assume that a vector space is always defined over $\Re$. When the tutor brings their attention to other possibilities and asks them what the scalars are, responses are like the one below:*
> *Student:* ...constants.
> *Tutor:* Constants?
> *Student:* Numbers.
> *Tutor:* How do you know?
> *Student:* Are they real?

The students are usually reluctant to pay special attention to 'details' like this because the majority of vector spaces they have to deal with at this stage is defined over $\Re$. So in this case this avoidance of the general case of a field in favour of the specific and familiar $\Re$ seems to be convenient.

The students occasionally express a preference for notation at the expense of conciseness and comprehensiveness. For example, and in relation to the use of the concept of absolute value, as well as notation ||, in a number of cases the students appear as if they *do* know that $|a-b|<c$ translates into $-c<a-b<c$, but given an inequality of the latter form, they can rarely see it concisely represented by the former.

This becomes more explicit in the cases where in the course of a proof they devise the inequalities themselves. In these tutorials no student came up with a direct algebraic expression involving ||. Their explicit preference was for long inequality algebraic expressions which, despite their length, appear visually as much closer to their mental image of the inequality to be expressed.

'||' is a compressed expression whose interpretation and handling repeatedly seemed to be problematic. It is therefore expected that in a selective process amongst various alternatives the learners will not favour a mathematical tool that, so far, has caused them inconvenience and confusion. Its comprehensiveness and elegance are of little relevance: when a mathematical tool is not embraced smoothly as a personal construct, the learner's response to it is likely to be poor; however highly recommended is its use for epistemological reasons.

In the following I present a sample of observations from the *Pilot Study* relating to the students' difficulties with mathematical reasoning.

IIb. Students' Logical Difficulties

As a sample of the students' logical difficulties I refer to their common practice of *stripping of the theorems of their conditions* ; also to some *idiosyncratic uses of Mathematical Induction*.

Theorems stripped of their conditions. One part of the novices' occasional tendency towards logical inconsistency in these tutorials has been the students' tendency *to use the implied premises of theorems without securing the validity of the prerequisite conditions or neglecting some of them.* The theorems in connection to which the above tendency has been observed are:

- a form of the Mean Value theorem (Differential Calculus),
- the test for the convergence of alternating series (Sequences and Series),
- the ratio test for the convergence of positive series (Sequences and Series) and
- theorems on the existence of an integral of a function (Integral Calculus).

I note here that *these theorems have as the implied premise not a worded statement, but a formula* (e.g. in the Mean Value theorem: the formula for the value of the derivative of f on a particular point). Therefore it is possible, even though no further relevant evidence is supplied by these data, that the novices treat these theorems as follows: a formula is a condensed, succinct-hence-handy mathematical expression which may have a strong appeal for most students. Formulas are seen as immediate solutions to problems — and often they are. Thus at the sight of a promising formula, the students operate in a state of *cognitive excitement* and ignore the conditions.

Moreover I note that occasionally the tutors treat the students' negligence with considerable lightness: a likely reason for that might be that exam questions are *designed to fit* the undergraduates' up-to-date knowledge — and vice versa, the undergraduates are trained in a particular kind of knowledge designed to fit exam questions. It is thus more likely than not that a mathematical problem, being part (ii) of an exam question with, for instance, the Mean Value theorem as part (i), is this theorem's application. Therefore very little is at stake when the student neglects checking the conditions. What in this case seems to be totally forgotten is the *power of habit*. A student who cultivates this attitude of negligence may escape without any harm this particular or similar situations. Soon however she will find herself in mathematically more open, exploratory situations in which checking out the conditions of a theorem is essential. It is therefore possible that light treatment of this sort of learning behaviour might turn out to be a carrier of potential *epistemological bugs* for the near future.

It also seems to be the case that students fail to see *the necessity of the conditions* in a theorem: they are merely concerned with (and thus remember) only the *useful* part of the theorem. Its conditions are seen as sheer paraphernalia. *Integral calculus* appears to be particularly conducive to this. With its self-contained integral formulas, students are *tempted* in a cognitive sense to neglect the conditions under which the formulas hold. The issue brought up here is of a more general significance. Conditions in theorems play the role of progressive restrictions to the generality of a statement. They are the outcome of an *exclusion process* during which mathematicians verify for which family of mathematical objects this formula holds. One might wonder whether the students' behaviour regarding the handling of theorems partly reflects this developmental order, that is whether they behave in such a way because they are somehow at an earlier stage of this exclusion process, that is at the stage of deceptively believing in the generality of the formula. Evidence from these tutorials that would illuminate this further is not available but, as explained in Part III, it is useful to know that the students' verbalisations in the tutorials can raise such significant issues.

Idiosyncratic uses of Mathematical Induction. In these tutorials the students were not at ease with the use and meaning of Mathematical Induction. Here is a sample of their unease:

> *Example 5. In proving that under certain conditions $a^n \rightarrow 1$ a student suggests:* 'I tried induction'.

The student seems to confuse proving propositions involving an integer variable with proving limits of arithmetical sequences.

> *Example 6. The students are not at ease with the idea that mathematical induction can be used to prove a statement $P(n)$ $\forall n$ as well as for proving a proposition $P(n)$ $\forall n>k$, for some fixed k. Similarly in the convergence of series they are not comfortable with the idea that excluding a few terms in the beginning of the series (up to a fixed number k) is a valid practice and*

> *does not influence the value of the* 'infinite sum', *as they often call the series. In one occasion the tutor's explanation* 'we don't worry about the terms before k because this is a finite part of the series and doesn't change anything' *was followed by a student's surprised:* 'do we just forget about them?'.

Now what disturbs these students is probably how it is possible for a considerable, perfectly measurable decrease in quantity to leave things unchanged. In finite terms this doesn't make sense. The incident adds evidence to the *persistence of the finite mode of thinking*, a quite common phenomenon in novice mathematical thinking.

In the following I present a sample of observations from the *Pilot Study* that relate to the students' difficulties with mathematical notation.

IIc. Students' Symbolic Difficulties

Notation is apparently one of the most frequently raised issues in tutorials since novices seem prone to attribute a great deal of their occasionally dysfunctional writing to its novelties. In particular, I refer here to two cases where notation seems to impede understanding: *ij-notation* and *$\sum\sum$-notation*.

> *Example 7. Consider V a vector space such that $V \subseteq \{f / f: \Re^n \rightarrow \Re\}$. During the tutorial the set S, defined as $S=\{f_i / f_i\ \Re^n \rightarrow \Re$ such that $f_i(x_j)=\delta_{ij}$ where $i,j = 1,\ldots,n\}$, has been identified as a potential basis for V (δ_{ij} is Kronecker's Delta). To prove that S is linearly independent, one of the two students in the tutorial suggests writing a linear combination $af_1+bf_2+\ldots$ of f_i and proving that $a=b=\ldots=0$. The tutor asks the student to use 'better notation' and the student changes a, b,...to a_i. Then the other student protests:* 'Now I'm lost in this'.

The novices' attitude of avoidance towards the notation of indices is intriguing: it seems that this notation's epistemological virtue (namely its succinctness and flexibility to represent multidimensional situations) is its didactical vice (namely its rejection by the novices as too complicated). Evidence of a similar attitude towards $\sum$-notation is available in the same data. It may be of relevance here to mention that for instance $\sum\sum$-notation, even $\sum$-notation, was introduced as late as 1830 by Jarrett, whereas sums of quantities of course were studied much earlier (Cajori 1993). It is therefore a possibility that the learner — in parallel with the 19th century mathematician before Jarrett's intervention — understands and works with sums but not necessarily with the notation commonly used. The instance cited above exemplifies this possibility. In the *Main Study* the implications of the novices' problematic use of mathematical notation is embedded more globally in the issues related to their enculturation into advanced mathematical thinking. Moreover their difficulty in using the new notation competently is juxtaposed to the occasions where they demonstrate facility in grasping intuitively a mathematical argument but cannot express it formally.

I close Part II with a sample of instances from the *Pilot's* data that do not necessarily fall into one of the three S.DIFF categories but, as discussed in Part III, seem to illuminate aspects of the novices' thinking processes and therefore demonstrate the potential of tutorials as a source of evidence for the novices' learning.

IId. Miscellaneous but Significant Instances

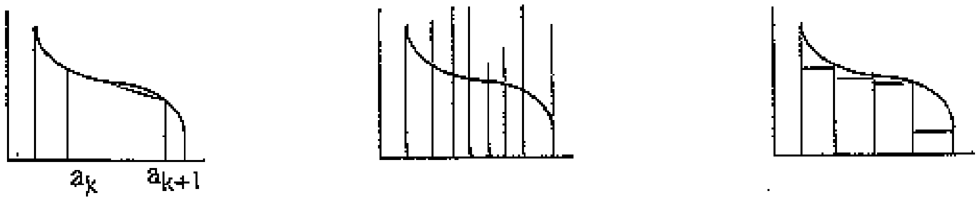
The instances from the *Pilot's* data presented in this section are examples of

- how novice mathematical behaviour often deviates from common mathematical practices (example from Integral Calculus),
- how students often find it hard to accommodate intuitive practices in their newly acquired formal practices (example from Differential Calculus), and,
- how the students make tacit assumptions about exam questions.

Deviations from common mathematical practices: calculating the area under a graph. Integrals/integration is usually presented to students as either *'area under a graph'* or as *'the reverse of differentiation'*. In the tutorials, where the following example comes from, the tutor suggests a link between these two perceptions and explains the idea of approximating the area under the graph of a function f with calculating the areas of *'simpler'* figures. The story continues as follows (I note that the first student mentioned in the example is Greek and is probably under the influence of a strongly Euclidean school-geometry past) :

> *Example 8. In one of the tutorials a student suggests using* 'triangles' *as the simpler figures mentioned by the tutor. When the student's attempt to express the area of triangles in* 'simple' *terms fails, the tutor points at rectangles as convenient simpler figures.*
> *In the same fashion, in the tutorial given to another pair of students A and B on the same day, student A suggests ['...' means : a short hesitant pause]:*
> *Student A:* ...trapezia.
> *The tutor and Student B laugh. Then:*
> *Student B*: ...strips maybe.
> *Tutor*: Yes, what about the top then?
> *Student A*: ...well ignore it...and make it a rectangle.

One interpretation is that the student's image of the situation was similar to the one illustrated in the figures below:



That is Student A approximates the area under the graph as the sum of the areas of the trapezia such as the one defined by points $(a_k, 0)$, $(a_{k+1}, 0)$, $(a_{k+1}, f(a_{k+1}))$ and $(a_k, f(a_k))$. To the others' laughing reaction, Student B changes to strips (see the figure above). The tutor then picks Student B's suggestion and turns the two students' attention to the 'tops'. Student A, possibly seeing the possibility of acquiring rectangles by dispensing with the 'tops', suggests 'ignore it then...and make it a rectangle'.

It is not possible to see here whether in the above instance the route from trapezia to strips and finally to rectangles reflects a tentative reconstruction of the memory of a definition, a definition associated with the calculation of strip-like rectangles; or whether this is a reflection of the tutor's orchestrated effort to lead the students to the realisation of the necessity of rectangles in the definition of an integral. It is also possible that the word 'approximation' evoked in the student's mind the Trapezium Rule for the approximation to an integral as the limit of a sum of trapezia, exact only for linear functions, used in the first years' course on Analytical and Numerical Methods; hence the suggestion to use trapezia. It is due to the limitations of exploring knowledge constructs through observation that the above explanations must remain tentative. On occasions like this one what became evident during the *Pilot Study* was not only some aspects of the novice mathematician's thought processes, but also the strong phenomenological character of the study as outlined in Chapter 2.

The ambiguous propriety of guessing limits. Towards the end of their first year novices seem to be convinced of the necessity to formalise their way of mathematical thinking, even though they have only vague ideas about what this formalisation entails and how it is carried out. A tendency they seem to adopt and which will be further elaborated in the *Main Study* is caution against intuition. Their confidence in its validity is shaken especially as compared to their considerable relying upon intuition at school. What they expect from mathematical arguments is a chain of implications the links of which are established with transparent clarity. As a result they are disappointed when this expectation is let down:

Example 9. In a session on Limits and Continuity the tutor asks the students for the limit of d^n under certain conditions. One student says: 'I know it is 1 but I do not know how...'. *In another tutorial, the exchange between student and tutor is as follows:*
Student: Can we have different cases for $c <,= \text{or} > 1$?
Tutor: No. But this helps you to see the answer really.
Student: Oh...it will be one actually!

This practice of guessing the limit, and then proving it, is very common in mathematics but it seems to raise some suspicion of impropriety in some students. In one occasion one of them asks: 'How about when you have somehow found a limit for yourself and then you prove it? Is it OK?'. This demand for transparency in the procedures of arguing mathematically may well originate in the students' semi-unconscious nostalgia for the clear-cut straightness of school-mathematics. Novices in the beginning of an advanced mathematics course are in a constant process of learning about what mathematical thinking is and this is a fluid position full of insecurities. In the cases presented above the metamathematical lesson to be learned is that the process of inventing in mathematics is far from governed by clear-cut rules but the standards regarding the product of invention are. So one is allowed to feel or see a limit of a sequence in any which way they can. The rules of the game are firm when this limit has to be proved. Invention can use an ample range of resources, some of them inarticulable; this is not so for conviction. The students' insecurity may lie in the difficulty to make this distinction and the opportunity to learn how to make this distinction should be part of their novel advanced mathematical experience.

The students tacit assumptions about exam questions: depth, length, complexity and linguistic conventions. During Trinity Term the students appeared increasingly concerned with exams. Their concern was expressed in a variety of queries about the nature of exam questions, some of which I cite here as indications of the existence of a body of assumptions permeating the students' perceptions of examinations. The quotations are from the *Pilot's* tutorials on a variety of mathematical topics:

Example 10. From a discussion of a Linear Algebra exam question a part of which is about 'what is a basis of a vector space'.
Student: Is it enough to say that this set of vectors spans and is linearly independent?
Tutor: I think they need a bit more.
The tutor then explains what she means by 'more'.

In another occasion a student is surprised when she realises that 'the answer is only three lines' *and in another tutorial a student wonders about an answer to an exam question whether* 'is it supposed to be long?...because the

> question is longer...considering how long they usually are...'. *Another common perception among these students appeared to be that a short answer is* 'trickier'.
>
> *In these tutorials evidence was also given of the impact language has on the students' perception of an exam question. For example, often* 'Give a basis for vector space V...' *was interpreted as* 'Provide a basis for vector space V but not necessarily prove it'.

The above is evidence to a slight re-focusing of the students' attention towards the end of the academic year which seems to alter not simply the flavour of their questions but also to a certain degree the nature of their mathematical inquiry. Their questions are more overtly about communication, mathematical writing and interpretation of the mathematical language used in the exams. I note that behind the technical character of these questions lies perhaps a certain amount of metamathematical concern further elaborated in the *Main Study*.

Finally, in the following I illustrate how the findings from the *Pilot Study* have, mostly methodologically but also thematically, contributed to the formation of the *Main Study*.

PART III The Influence of the Pilot Study on the Main Study

In this part an account is given of how the *Pilot Study* operated as a learning experience that contributed substantially to the formation of some features of the *Main Study*. The points that are raised here aim to illustrate

- how the *Pilot Study* confirmed that tutorials are a rich source of evidence for an investigation of mathematical cognition,
- how the *Pilot Study* drew attention to potentially interesting foci of analysis for the *Main Study* and,
- how in the Pilot study it became evident that the note-taking observation technique did not guarantee as detailed as necessary an access to the students' thought processes.

I note here that some technical features of the *Main Study* (such as the amount of data to be collected; or the adherence to collecting data from a diversity of mathematical topics) that were determined partly on the basis of the *Pilot* experience, will be accounted for in Chapter 4.

IIIa. Confirming that tutorials are a rich source of evidence on mathematical cognition

As explained earlier in this chapter observing tutorials was chosen as a way to gain access to the novices' expressions of mathematical cognition on the grounds that, in the intimacy of a one-to-one or two-to-one session, the novices would express freely and in detail the problematic aspects of their encounter with mathematical abstraction and that the individuality of their learning process would be highlighted. Even though the initial impression of what a tutorial contains was altered in the course of the *Pilot Study*, the above expectation was encouraged. The data collected during the *Pilot Study* illustrate that the potential of tutorials, as a source of material, is high.

In the above, it is contended that my initial impression of what a tutorial contains was altered during the Pilot Study. My preconception of a tutorial as a forum in which novices articulate their ideas, and in particular the problematic aspects of their mathematical thinking, was proved partly false because, as it turned out, the tutorials I have been observing are highly tutor-centred sessions. In brief the tutor's influence in these tutorials seemed to rely upon the following:

- her decision to focus on new material, on problem sheets or on resolving student queries;
- her openness to student participation and student suggestions for the structure and content of the session;
- her raising of controversial issues that generated discussion or preference for a tidy exposition.

Student influence on the other hand depended

- on how passive or participant the student is;
- on what the attitude towards tutorials is (such as: tutorial as a course obligation reluctantly attended. Or: tutorial as a helpful session that one must make the most of);
- on the effort on the part of the student to adjust the tutorial to their own needs.

Thus by the end of the *Pilot Study* the expectations about the contents of a tutorial had been reformed. In the meantime a thematic quandary had risen from the *Pilot* data analysis: whether the focus of the inquiry for the *Main Study* will be topical/

conceptual — that is focusing on the mathematical topics and concepts dealt with in the tutorials — or psychological/ procedural — that is focusing globally, regardless of concepts and topics, on the learning processes of the students. Given that quandary, the type of sought-for incidents in the tutorials took the form of:

- student monologues in which the student articulates a discourse on problematic aspects of their mathematical thinking (such as a specific topical query, a difficulty with reasoning, a metamathematical concern) either explicitly or implicitly via a presentation of their response to a mathematical problem,
- tutor/student dialogues of a similar to the above content, and
- student/student dialogues of similar content.

I note here that the latter is rare because tutorials are strongly tutor-centred and the tutor in most cases leaves meagre space for peer interaction. In the above list of sought-for incidents I have not included instances of the tutor's monologues since these can only indirectly be of use: either in cases where the tutor propagates ways of mathematical thinking that seem to have a formative influence on the students; or when the tutor explicitly points at the students' difficulties.

Thus while observation in the *Main Study* would still be open and relatively unsystematic, the experience from the *Pilot Study* rendered the process considerably more focused.

As far as the data analysis is concerned, the form that the findings (a sample of which was given in Part II) have taken indicates the potential of the data based theory emerging process as a method that facilitates the identification of trends regarding the novices' thinking processes. In its earlier stages the analysis of the *Pilot Study* material turned out to be an important *exercise in locating and maintaining a focus for the analysis of the incidents*. Some of the outcomes of this exercise have been mentioned in Parts I and II. What signified however the analysis of the incidents from the *Pilot Study* as an immense learning experience for the data analysis in the *Main Study*, was the following realisation: since the nature of the data allows an equipotent approach either from a teaching or from a learning perspective, it is an imperative necessity to keep in mind that what is sought for in the analysis of the material is the access to the novice's learning processes. A concern about the teaching dimension can only be incidental and is valid only as far as it illuminates the learning process. Digressions from the above aim are beyond the scope of this study, are costly time-wise and only weaken the discourse on the novice's mathematical thinking.

IIIb. The emergence of potentially interesting foci for the Main Study

In Part II a sample of findings has been presented in order to exhibit the potential of the material collected for the *Pilot Study*. This sampling is not random: it is done retrospectively, after a large part of the *Main Study* has been completed. This implies that the selection of the material from the *Pilot* that is presented in Part II and the way of presentation aim at highlighting the ways in which the *Pilot* provided potential foci for the data analysis in the *Main Study*. Strictly speaking the *Pilot Study* material shares very little with the *Main Study* mathematical material: the material analysed in the *Pilot Study* is a slice of the Trinity Term 1993 data whereas data collection for the *Main Study* took place in Michaelmas Term 1993 and Hilary Term 1994. Nevertheless despite the differences regarding their mathematical content, there is evidence in the *Pilot Study* foreshadowing the raising of some issues in the *Main Study*.

IIIc. The insufficiency of note-making; the need to obtain a detailed record of the tutorials

There is a remarkable number of instances from the *Pilot Study* material that could not be submitted to analysis. For a variety of reasons, the account of the events given in the fieldnotes was not detailed enough to allow interpretation. Note-taking is a selective recording of the events and in these tutorials the account of the events was interspersed with references to the mathematical content of the tutorials and with comments on the events from a didactical point of view. Moreover it was not technically possible to keep up with the conversation and reproduce accurately on paper more than a few consecutive lines of dialogue. This is the major reason why the size of incidents examined in the *Pilot* is substantially smaller than the size of incidents in the *Main Study*. The following two examples illustrate the impact that the lack of detail and the selective character of note-taking had on the nature and quality of the data.

> *Example 1. In the fieldnotes it is repeatedly stressed that the novices have serious problems with handling δ–ε definitions. It seems that the requirement to work with this kind of formalism is a stumbling block for the novices in their first encounters with Calculus. However in the Pilot Study it has been impossible to record the details of this difficulty since this would involve reproducing on paper the minute details of very fast and technical pieces of conversation.*

> *Example 2. In some cases a detailed account of the events in a tutorial is given in the fieldnotes but not on the aspects of the conversation that turned*

> *out to be significant for the analysis. In this case the conversation was about the students' responses to a problem involving solving three simultaneous equations in which the coefficients of the three unknowns x, y and z are parameters a, b and c. One student was explaining her solution. At the time what struck me as significant and therefore worthy of note-taking was the student's use of mathematical logic in her solving strategy. Later, when I revisited the instance and looked closer at her solution, I realised that her logic was rather trivial and unproblematic but her use of the notions of variable, unknown and parameter was not. Unfortunately I had no record of these details of the episode.*

What these two examples illustrate is the high hit-and-miss risk of relying upon a spontaneous selection of what is noteworthy in an observed situation; also the technical impossibility of giving a complete account. The latter was made even more obvious in the cases where the mathematical content of the tutorials was on topics in which my background is not particularly strong, that is applied mathematics. In these cases even following the events closely was difficult and so the account is not reliable enough to allow analysis and interpretation.

The above led to the realisation that in the *Main Study* note-taking should be replaced or supplemented by another data collection technique. Tape recording was selected as the approach that most appropriately resolves the technical problems mentioned above. A justification of this selection as well as the reasons that led to supplementing fieldnotes and tapes with interviews is given in Chapter 4.

Chapter 4
The Main Study: Procedures of *Data Collection*

Introduction

In this chapter I present an account of the *Data Collection* period. The presentation is in chronological order and my intention is to highlight those aspects of the learning context in which this study was carried out, that is the Oxford tutorial, and of the methodology which was employed, that have turned out to be significant in the light of the subsequent stages of the study. Hence I present the search for volunteers, the tutorial observation and the interviews. The intention here is to establish the links between the *conceptualisation* of the data collection methodology of the study, elaborated in Chapter 2, and its *realisation* as it took place during observation and interviewing.

Summary: From October 1993 to March 1994 twenty first-year mathematics students were observed for 14 weeks during tutorials on Analysis, Topology and Linear Algebra given to them individually or in pairs. The tutorials were tape-recorded and notes were taken during observation. The students were interviewed twice, in the middle and in the end of the observation period.

PART I The Participants and the Learning Context of Tutorials

In this part I present my search for participants as it took place in the first two weeks of Michaelmas Term 1993. The way the volunteering tutors and students were found is relevant to their profile. The findings of this study are derived from observations made during tutorials given by a small number of selected tutors to a small number of selected students in Oxford. This selection was based, not on the principles of statistical representativity, but on willingness to participate; also on timetable concerns such as clashes between tutorials.

Willingness to participate in an educational study possibly reflects some features in the psychological profile of the volunteering tutors and students which I, as the person in charge of the final selection of participants, hoped for and pursued. Evidently friendly and conversational tutors and open, expressive students who would not hesitate about vocalising their mathematical difficulties were sought for. The quest for these features was not explicit in the search campaign but the procedures that led to the formation of the body of participants constituted a reconciliation process between what was sought for and what was available.

At first a number of tutors from a variety of Oxford Colleges was approached. The experience of the *Pilot Study* had produced an estimate of the number of students that could fit reasonably in a week's schedule of observation (approximately twenty) so, given that most colleges had between 4 and 10 first year mathematics undergraduates that year, that implied that I should aim at an estimated number of three to five colleges. Negotiations with college tutors led to six positive responses, and three negative ones (the tutor was too busy to dedicate time to helping the researcher contacting and convincing the students; the tutor was concerned about the interference of the researcher's presence with the undergraduates' learning during the tutorials but was willing to allow one interview per student; the tutor thought students were too shy on the first term of their studies but was willing to participate in the study in Hilary Term). In the cases of negative responses the *Pilot Study* tutor's reassuring recommendations — at my request the *Pilot Study* tutor had informally discussed her experience with some of the potential participating tutors — had not succeeded in battling out these tutors' concerns. Moreover one of the positive responses had to be put aside because it turned out that the applied mathematical content of this tutor's tutorials did not coincide with the pure mathematical content of the others.

Subsequently the students (that were given tutorials by the tutors who had responded positively) were contacted. The introductory meetings were arranged by the tutors and were held in the college rooms where the tutorials were given. A short presentation of the researcher — on her work in general and on the nature of the requested participation in particular — was followed by questions from the students. The meetings ended with the researcher offering reassurances on the unintrusive, non-assessing character of her presence, on confidentiality and on the participants' right to withdraw from the study at any moment. It is worth noting that explanations about the nature of the investigation were not extensively given and particularly geared towards preventing the possibility that the study was perceived by the students as an assessment of their mathematical ability. It was clearly stated that this is an investigation of the novice mathematician's understanding and in particular of the problematic aspects of this understanding. Students then were given some time to decide and were asked to communicate their decisions to the tutor who would in turn inform the researcher.

By the end of the week an overwhelming 24 positive responses and 4 negative (these students were concerned about having the privacy of their individual tutorials disturbed but voted for the presence of the researcher at the group tutorial) accelerated the completion of the search for participants which had lasted the first two weeks of Michaelmas Term. Twenty one students were selected from four colleges. Final selection was based mostly on timetable concerns; also on maintaining a male-female balance among tutors and students.

The search for participants lasted the first two weeks of Michaelmas Term which is a crucial time, given that the purpose of this investigation is to capture aspects of the novel experience of advanced mathematics. Sparing these first few days of the novices' experiences was inevitable since attempts to contact tutors and students before the beginning of term had repeatedly failed. Moreover even if contact had been possible, the likelihood of tutors accepting collaboration before meeting their students, or of students accepting the researcher's presence in a teaching session they had never experienced before is small.

The background of the students participating in this study is diverse but their A-level experience is taken as a homogenising factor. Also despite the absence of any sign of linguistic difficulty, it must be noted that one of the students was Kenyan, one was of Turkish origin and another of Swiss-French. In the presentation of the findings, I shall refer to some special cases where significant differences in the students' background have been identified.

In retrospect it can be said that the overall majority of students who responded positively were friendly and open even though in some cases I felt that their positive response may have been the result of their reluctance to come to conflict with their tutor's willingness to participate in the study. However, no explicit evidence of that is available. As for the tutors, the willingness with which they responded positively to the invitation, constitutes evidence of a genuine interest in aiding a didactical investigation of tutorials. The role played by personality factors will become more evident throughout the study but also in the subsequent parts of this chapter.

Having finalised the body of participants, I then realised that, given my limited knowledge of the course structure and protocol, I needed to explore the learning context of the investigation, that is the tutorials. My mathematical background guaranteed a relatively immediate access to the mathematical content of the sessions. Gathering the relevant information regarding the course in general was the next step before moving on to the actual observation. Before and during observation the following documents constituted the body of information that helped me familiarise with the course and in particular with the essentially private learning context that tutorials are:

- The Lecture Lists for Michaelmas Term 1993 and Hilary Term 1994,
- The Synopsis of Moderations (First Year) Lectures in Mathematics for 1993-94,
- The Reading List for Moderations in Mathematics,
- The weekly Problem Sheets on the 11 courses taken by the first year students during these terms,
- The *Guide to Studying Mathematics at Oxford University: How Do Undergraduates Do Mathematics?* by C.J.K. Batty.

I note here that the latter, with its chapters on University Study, University Mathematics, The Formulation of Mathematical Statements and on Proofs, has the enormous ambition of initiating the novice into the rigorous mathematical thinking that is to succeed school mathematical thinking. Expectations regarding its influence were high but informal conversations with tutors led to a more sceptical approach: its thorough written style was deemed to be difficult by novices. Therefore the document had been practically marginalised. The few scattered indirect references to the Guide made by the students during the six months of observation will be mentioned when the novice's difficulties with the shift to a more rigorous way of thinking mathematically is presented in the subsequent chapters.

The documents mentioned above were circulated in the Mathematical Institute. There is however a bulk of documents circulated at college level and these will be looked at separately when necessary. These include documents whose content ranges from:

- Revision Suggestions,
- Warm-Up Exercises for the Institute's Problem Sheets,
- Suggested Solutions to some Questions from the Institute's Problem Sheets,
- Collections (the January college-based examination that follows Christmas revision, the structure of which is usually suggested by the tutor on the eighth week of Michaelmas Term),

to

- A Few Hints on How to Read a Mathematical Textbook and,
- Briefings on the Greek Alphabet.

There was no intention to include in the above a study of the students' written work. Occasionally however I was allowed by the students to survey briefly their weekly drafts.

PART II Tutorial Observation and Interviewing

In the following I present an account and an evaluation of the tutorial observation and the interviews with the participants.

IIa. Observation of Tutorials

I first present a factual and then an evaluative account of tutorial observation.

IIa.i Observation of Tutorials: a Factual Account

The body of participating tutors and students was formed in the first two weeks of Michaelmas Term 1993. Observation started on the 3rd week and lasted for the remaining six weeks of the term. In Hilary Term observation lasted eight weeks. As mentioned in Part I, missing the first two weeks of Michaelmas Term was inevitable. Given that

- the mathematical content of the observed tutorials is mostly determined by the content of the weekly problem sheets of the Mathematical Institute and that,
- when observation started on the third week, the tutorial was on the second weekly sheet,

it is fair to assume that the tutorials of Weeks 1 and 2 were, respectively, a general introductory session and a tutorial on the first weekly problem sheets. For the sake of continuity I informally discussed the contents of the first two tutorials with the tutors. From a methodological point of view my presence in the sessions from the very beginning might have created tensions, since the tutorial on Week 1 is the novice's first entering of this learning context. From the point of view of the thematic focus of the study (the novice's encounter with mathematical abstraction), however, the experience might have been valuable even though, according to most tutors, students do not participate extensively in the first sessions. In sum I think that giving the students the opportunity to experience tutorials for a couple of sessions and then asking them to decide whether they wish to participate in the study was fair and possibly made my suggestion to attend the sessions more easily accepted.

In Michaelmas Term observation of tutorials was 10 hours per week for individual or pair sessions plus 3 hours of group sessions in two of the four colleges. The allocation of time per college, hence per tutor, for Michaelmas Term is as follows:

8 students	in 4×30 minute pair sessions
5 students	in 2×60 minute pair sessions and 2×60 individual sessions
5 students	in 2×60 minute (one pair and one three-student) sessions
4 students	in 4×30 minute individual sessions

I note that in the second college of the above table the pairing of the five students alternated so that by the end of term all of them had an equal number of pair and individual sessions. From the 22 students participating in Michaelmas Term, one left Oxford in Hilary Term for personal reasons and another one had to be excluded from the observation because of timetable clashes and because the observation and recording of material in one of the four colleges was drastically reduced in Hilary Term. Also one of the tutors, due to sabbatical leave, was replaced by two tutors of the same college and also changed the pairing of students. This implied that one of this college's Michaelmas student volunteers had to be replaced by another student from the same college. Therefore the picture for Hilary Term is as follows for the five tutors in the four colleges:

8 students	in 4×30 minute pair sessions
4 students	in 2×60 minute pair sessions
4 students	in 2×60 minute pair sessions (two tutors)
4 students	in 4×30 minute individual sessions.

Group sessions were also observed in Hilary Term.

The mathematical content of the tutorials can be rather loosely defined as relevant to the content of the lectures and problem sheets of the Mathematical Institute for Michaelmas Term on

Linear Algebra
Continuity and Differentiability
Geometry

and for Hilary Term on

Groups, Rings and Fields
Sequences and Series
Topology

Occasionally other topics are addressed but the main bulk of discussions were on the above.

The tutorials were held in Oxford college rooms. In two of the four colleges the observer shared the same table with the tutor and the students. In the other two the observer was seated in a corner of the room having auditory and visual access in one of the colleges and only auditory to the other.

The individual and pair sessions were tape-recorded and the observer made notes during observation. The need to obtain a record of the sessions beyond fieldnotes was an aftermath of the *Pilot Study*, the data of which were of a strongly fragmentary and derivative character (see Chapter 3, Part Ib). Data originating in fieldnotes are instant reconstructions of the events and thus lack the accuracy of non-selective recording. Initially video-recording was considered as a viable option. At the very early stages of the negotiations with the tutors it became clear that video was too intimidating for most prospective participants and a too-visible interference in the intimate environment of college rooms where the tutorials were held. So, with the risk of missing non-audible communications, I eventually decided to proceed with suggesting audio recording to the participants.

Tape recording was more easily welcomed. It also provides a non-selective account of the events in the tutorials and allows the observer to keep a complementary record of the aspects of the tutorial that an audio tape cannot capture, such as writing on paper or on the blackboard and gestures. Depending on the situation, the observer's comments and observations might also fit in the fieldnotes. As mentioned above the intention was that through the available documents, the tape recordings and the fieldnotes, as comprehensive as possible an understanding of the learning occurrences in a tutorial would be achieved.

IIa.ii Observation of Tutorials: an Evaluative Account

By the end of data collection my general impression was that the expectations set up during the *Pilot Study*, regarding the didactical content of tutorials, had been fulfilled. Tutorials are indeed a rich source of learning incidents and the primary stages of Data Analysis (see Chapter 5, Part I) indicated that the technical strategies employed provided a reliable record of these learning instances. However I also felt strongly that processing these data would be a more complex task than previously because of the diversity of teaching style, mathematical content and quantity of the data. Briefly I would describe the above impression as a realisation of the increased number of variables determining the nature of the data. As a sample of these new variables I note the shift from one tutor to four and the shift from the uniform content and structure of the tutorials observed in the *Pilot Study* to the nearly chaotic diversity of the mathematical content, tutoring styles and student personalities in the *Main Study*.

Another notable difference concerns the content of the fieldnotes: in the Pilot *Study* these fieldnotes were the data; in the *Main Study* they were a supplement consisting mostly of things not captured in the recording and some primary indicators of where analysis could possibly focus.

By the end of the second term an amount of approximately 120 hours of recordings had been gathered. This is less than the sum of hours of observation mainly because of the instant editing that took place in one of the colleges. The hours of observation in this college decreased from 4 in Michaelmas Term to 2 in Hilary Term, because I decided that the tutor's long monologues that contain exposition and technical advice are not particularly useful to the purposes of this study. During the two hours of observation in this college in Hilary Term, recording was paused on such occasions.

Other losses are due to the absence of students but these were miniscule. Also there have been minor losses of material due to faults of the recording equipment. Audibility of the recordings varies but is generally good. From the beginning this has been a factor that merited special attention given that the study is in my second language.

The effect of the researcher's presence can be described here, on the basis of my judgement and the participants' comments, as minimally significant. In the beginning however the degree of self-consciousness, mostly on the part of the tutors, is slightly higher and so was the number of instances in which the tutor would break the tacit rule of the researcher's invisibility and address her during a session. A slight increase in these instances has been observed towards the end of Hilary Term where all participants and the researcher felt the reasons for a strict attachment to the rules of minimally participant observation were no longer required. My invisibility was imperative in the beginning of observation. This was the time that a relation of trust and openness had to be established. Towards the end of the sixth month the relation had been evidently well established: I had been accommodated and my presence in the tutorials raised no questions; it was deemed natural.

Some students admit that they had noticed that at some instances my note-taking appeared keener but that they could not see a pattern in this behaviour. When they were convinced of the non-threatening character of my presence, they stopped paying attention. Notably, after the first interview, most of them say that they had a more solid understanding of the purposes of my investigation. According to tutors and students, no visible change in their attitude or behaviour can be linked to my presence.

A few students refer to their tutors as 'a little more polite and patient' during recordings and some tutors thought their students were 'more attentive'. Even if these changes occurred extensively — and they did not — they would promote a better realisation of the purposes of data collection: a more polite and patient tutor and a more attentive student is likely to enhance the chances for a more mathematically and didactically creative, and revealing, session.

On my part, instant understanding of the occurrences on the tutorials during observation depended upon the strength of my mathematical background in the relevant topic and the clarity of expression on the part of the tutor and the students. For example the instances where the tutor was correcting the students' drafts during the tutorials were usually unhelpful because I generally had no access to the drafts. Also the tutor's and the students' comments on the drafts, despite their relevance, most of the times, to the difficulties the students had with the questions in the problem sheets, could not always be helpful because of their fragmentary and unclear referencing. For example in quite a few occasions the tutor and the students point at various parts of the students' drafts and refer to what they point at as 'this', 'it', 'here' or 'there'. Since seating arrangements did not allow me to have any visual access to the drafts, these references were hard to understand. I used to work on the Institute's problem sheets beforehand and this proved a good strategy for enhancing understanding in most cases.

As opposed to these fragmentary and unclear references to previous problems, complete presentations of questions from the problem sheets, or new questions/theorems, by the tutor or the student, were usually understandable and transparent. They thus constitute the material that, as it turned out in the analysis, produced data of the highest quality.

IIb. Interviews With the Participating Students

In the following I present first a factual and then an evaluative account of the interviews with the participating students.

IIb.i Interviews With the Participating Students: A Factual Account

During observation a few generally problematic areas of the mathematical content discussed in the tutorials were identified and thus became the focus of the loosely structured interviews conducted on the eighth week of each term. Eleven students were interviewed in Michaelmas and nineteen in Hilary Term. The scenarios of the two interviews are as follows:

Michaelmas Interview

Area 1. The student is asked if she remembers the Bolzano-Weierstrass Theorem. What is a Bounded Set? What is an Accumulation Point? What is her picture of an Accumulation Point? Can she talk about it? If she is given the set A={ 1/n: n is a positive integer} can she talk about the content and the Accumulation Points of A? Does A have any Isolated Points? What is an Isolated Point? Are these two terms — Acc.P and Isol.P — opposite words?

Would she say the same about Open and Closed Set? What is an Open Set? What is a Closed Set? Can she give examples?

Area 2. What does the Squeeze (or Sandwich) Principle say? What is a limit? Pictures? Examples?

Area 3. The student is given the words 'to span', 'to be spanned by', 'the span', 'the spanning' and asked to talk about them.

Hilary Interview

Area 1. What is a sequence? What is a series? What does it mean if 'a sequence converges'? What does it mean if 'a series converges'? How do we check out convergence? What do the various tests say? Why is the Ratio Test true? Any other tests?

Area 2. What is a Compact Set? Definitions? Pictures? Examples?

Area 3. What does the First Isomorphism Theorem for Groups say? What is an isomorphism? What is a Kernel? What is the Quotient Group? What is a Coset?

Interviews were 25-30 minutes long and were tape recorded. A compilation of the students writing and drawing during the interviews has been kept. None of the students who were invited to an interview declined and all except one kept the appointment. In Michaelmas Term only a selection of students was interviewed. Selection was based on the impression that these students made to the researcher as to their willingness to vocalise their mathematical thinking. Audibility varies depending on the surroundings of the interview but is generally very good.

IIb.ii Interviews With the Participating Students: An Evaluative Account

The interviews from Michaelmas Term are relatively flawed because there is an eagerness on my part to cover all the topics in the scenario. As a consequence the interviewees are sometimes not given enough time to reflect on a question. Moreover in some occasions valuable time is lost when I, carried away by the persistent request of the interviewees for correction, indulged in exposition. These occurrences are rare and offered a remarkable learning experience of things to be avoided in Hilary Term interviews. These are more disciplined and better prepared. Eventually, nevertheless, I felt that both sets of interviews have substantially achieved their aims. Given the lack of time for a thorough preparation (the interviews had to be done in term time because of most students' absence during vacation) the identification of problematic topics to be touched upon in the interviews was generally successful.

PART III Allowing for Feedback and Maintaining Contact with the Participants. The beginning of Data Analysis

Data collection was completed in March 1994. No withdrawals from the study occurred and at no instance was the observer asked to leave or pause recording. The general ambience between volunteers and observer can be described as growingly amicable. Amicability on the part of the participants is probably evidence that the researcher's presence is not seen as threatening or disrupting. Therefore it is reasonable to assume that the students' behaviour has been fairly natural during observation.

When participants expressed their interest in my findings during observation I generally referred to the course and to areas of the mathematical content with which the students appeared to have difficulty. My reluctance to be more specific was due to a methodological concern related to the presupposition that, in a minimally participant observation, the participants should not be influenced towards behaviour that interferes with the naturalistic character of the observation: in other words, I clearly avoided any statements that might have deterred the participants from openly discussing the problematic aspects of their mathematical learning. In this vein I was consistently repeating during observation that my presence was not of an evaluative nature and that I was interested more in their difficulties than in an impeccable demonstration of mathematical competence. The degree of openness that the students and tutors demonstrated during data collection as well as a remarkable pile of comments according to which the experience of participating in the study is seen as very positive — the comments on my presence in the tutorials and the interviews vary from 'definitely not obtrusive' to cathartic and relieving 'confession' — indicate that the participants were convinced and that they welcomed such an investigation.

I was very willing to reveal the details of my work after data collection was completed. This resulted in a number of informal discussions with the participants and also in an invitation by one of the participating colleges to a seminar in which I presented some data and analysis. I also accepted a similar invitation by the Mathematical Institute at Oxford. All participants declared their willingness to provide me with any feedback that would turn out to be necessary in the course of data analysis.

In sum the above outlined ambience had a positive influence on the quality of the data in the sense that the students did not appear intimidated and at times were willing to elaborate further on their thinking knowing that I found it interesting — especially towards the end of the observation period.

Chapter 5
The Main Study: Procedures of *Data Analysis*

Introduction

In this chapter I present an account of the *Data Analysis* period. The presentation is in chronological order and my intention is to highlight the evolving stages of data processing as far as the focus of the research as well as the employed techniques are concerned. For this purpose I present an account of

- the data processing during data collection and
- the data processing after data collection

for both the material collected in the tutorial recordings and the interviews. As in Chapter 4 the intention is to establish the links between the *conceptualisation* of the data analysis methodology of the study, elaborated upon in Chapter 2, and its *realisation* as it took place during and after observation and interviewing.

Summary: Since March 1994 the recordings of the observed tutorials and the interviews have been selectively transcribed and broken into thematic parts. Episodes related to the novices' learning difficulties have been extracted. The number of Episodes that appear in the final version of Data Analysis was gradually reduced as the focus of the research became tighter. The theorising of the study is grounded on the analysis of the Episodes in each thematic part.

PART I Data Processing During Data Collection

In the following I give an account of the initial stages of ordering the collected material as well as of the first attempt to explore its contents which led to the construction of the *Scripts* and to the first steps towards the tightening of the research focus.

Ia. Initial ordering of the Data

As explained in Chapter 4, data collection, from tutorial observation and interviews, took place during the first two terms of the participants' first year in Oxford. During the 14 weeks of data collection — interviews were taken on the 6th and on the 14th week of observation — it was imperative that a neat record of the procedures was kept. The vast amount of collected information (recordings, relevant documentation, fieldnotes, personal experience of immersing into the field) as well as the short time during which this information was collected made it necessary to order the data in a way that would facilitate the re-evoking of the research experience in the subsequent months and would allow quick access to the material.

At a practical level this implied that

- all recordings were immediately labelled (date, time, college, participants/ interviewees),
- all documents were also labelled with dates and in most cases with a brief note on the document's relevance and reason to keep it in file and,
- all recordings were matched with problem sheets and fieldnotes; interview recordings were matched with the interviewees' writings while interviewed.

At a more reflective level and in an attempt to capture and preserve the atmosphere of the data collection period, as well as register the off-the-record input to the study that sprang from the informal contact with the participants, a Researcher's Log was kept from the days of the search for volunteers up to and including the earlier stages of data processing. Unfortunately my consistency in keeping the Log flagged as data processing carried on, possibly because most of the notes and observations I deemed important became part of the data analysis. As a paradoxical result, while reporting the procedures of Data Collection and Data Analysis in Chapters 4 and 5, I found it increasingly difficult to reconstruct an account for the latter despite its being considerably more recent. This is the major reason why the account of the procedures of Data Analysis consists mainly of the final techniques used in the study and not of the various attempts at other techniques that were abandoned in the course of analysis.

The Log contains a large amount of information that turned out to be variably useful in the course of the study. As far as tutorials are concerned the log was used as a register of immediate off-the-record (namely off the recordings and off the fieldnotes) observations made during informal conversation with the participants from the time when volunteers to participate in the study were sought, until the completion of the data collection. So for instance the final selection of participants, tutors and students, was largely based on the impressions registered in the Log (for example with regard to the ambience of my first meeting with the tutor or of the introductory meetings with the students); similarly the selection of interviewees for the first interviews — only half of the students were interviewed on the 6th week — was largely based on it (for example I chose to interview approximately half of the students in each college and I based my choice on the expectation, formed during the first five weeks of observation, that some of the students would be more open).

Moreover, since during interviewing I made no notes but observed and listened/ responded carefully to the interviewees, the Log served as a memoir of my off-the-record impressions of the interviews with regard to:

- the interviewees' tone of voice and body language,
- their amicability, openness, fatigue/willingness to participate and elaborate,
- the interview environment: for instance noise, privacy.
- self evaluation: formative comments that allowed me to improve the quality of subsequent interviews (for instance the need to allow more time for the interviewees to respond; the need to be more flexible and not insist on covering all the issues by all means and at the expense of the quality of the coverage).

The part of the Log devoted to the observations made during the tutorials mostly contains the students' evaluative comments of their day-to-day experience of the lectures, the tutoring, the problem sheets, the textbooks, their peers. Time constraints and difficulty of lectures and problem sheets dominate these comments. Some are the students' queries about the study and the goals of my observation and interviewing. As for the tutors, in their comments registered in the Log, they address general issues of the difficulty of some courses and their expectations of the students. Occasionally they inquire about my findings or explain some of the approaches they took during teaching. In the presentation of the data, Chapters 6 to 10, references to the Log are made whenever it is deemed necessary or illuminating.

In conclusion it seems that good management of the collected material is crucial for

- practical reasons, such as not wasting time on looking for items like tapes or documents, and,
- for reasons directly related to the quality of the research, such as the degree to which the ordering of the material allows easier references and making of connections.

Apart from keeping a reasonably neat and comprehensive record of the collected material, during Data Collection, the construction of *Scripts* took place.

Ib. The construction of *Scripts*

A *Script* is a summary of the tutorial events that highlights the structure of the tutorial. It constitutes a primary compartmentalisation of the tutorial to its component incidents. It contains no transcribed parts but indicates where transcribing might eventually prove necessary. It also contains my reflective comments made in the *freshness of the moment*. Especially because of the value of the latter there was a general attempt towards constructing the *Scripts* as close to the actual time of the tutorial as possible. So it was intended that the *Scripts* were constructed within the same day, or at least within the first couple of days, of the time when the recording had been made. The text of the *Scripts* was based upon

- one non-stop listening to the tape
- fieldnotes
- general consulting of problem sheets and other documents.

The aim of employing the freshness of the moment to provide a concise and accurate account of the events in the tutorial was achieved with variable degrees of success during the two terms of observation. Success was considerably higher during the second term of observation since more convenient time allocation of the tutorial recordings during the day allowed a day-by-day construction of the *Scripts* . Accumulation of research experience also contributed to a more efficient exploitation of time.

In conclusion the aim of maintaining a neat record and *keeping up with the data* during Data Collection was generally achieved. What did not so evidently succeed was progressive focusing of the thematic aims of the research. In the following I give an account of the primary attempts at tightening the focus of the study.

Ic. A step of Progressive Focusing: refining the focus of the study during data collection

Due to the density of the data collection, there was scarcely any time to reflect on the collected material. The only record of reflection during data collection are the impressionistic, telegraphic comments in the Log, in the fieldnotes or partly in the *Scripts* . So striking impressions of the students' overall learning behaviour (for example *'their difficulty to express formally'*, or *'the intimidating role of new mathematical notation'*) or primary interpretations of their learning behaviour (for example *'the conflict between the school-mathematics approach and the university-mathematics approach'*) were registered but only in a general and haphazardous way.

However a first attempt for a global reflection on the contents of the collected material became necessary immediately after the completion of the data collection on the occasion of the Transfer to DPhil Status, the administrative process during which the Oxford University Department of Educational Studies assesses the progress of Probationary Research Students. Due to the requirements of the Transfer process, a first scanning through the fieldnotes and the *Scripts* led to a first, primary collection of themes that drove the study. I quote from the Transfer Paper in order to illustrate the status of the Research Focus at the time:

> ...
> *A few directions regarding the handling of the material, each of which approaches the material from a different angle, are listed below. These reflect themes that it is intended to address in the analysis and are by no means exhaustive.*
>
> *COMMENT: The process of data analysis is expected to be cyclic. That is, as themes emerge in the course of the analytical process, the material has to be revisited in the light of the new themes. Ceasing this seemingly never-ending process depends on the degree to which the aims of the study will have been achieved.*
>
> *The First-Principles/Theorem -Deduced Proofs Debate. In their encounter with mathematical problems novices are very little aware of what knowledge is taken for granted, what is the continuity between school knowledge and university knowledge, in brief what is the status of the knowledge they are expected to handle. In most occasions the above is crystallised in the question of what a novice is allowed to assume in a deductive procedure, for instance whether in this procedure all deductions must be based on First Principles or on already known Propositions. This question is of a*

foundational nature and the absence or insufficiency of an answer given to novices poses severe hindrances in their route towards mathematical abstraction.

The Different Levels of Mathematical Knowledge. The mathematics that a novice encounters varies in content from arithmetical and algebraic calculations applied to numbers or other elements such as matrices, vectors, functions, sets, sets of sets etc., to calculational techniques, proofs, methods of proof, items on the content and use of mathematical objects, and also a variety of metamathematical information. This diversity in content implies the necessity for a diversity-aware approach to teaching. Tutoring lacking in this awareness leads to the novice's problematic handling of mathematical abstraction.

The portrait of a problematic question. An issue to be addressed is what renders a question in a problem sheet problematic. Gathering the queries of the students on the problem sheet questions might illuminate that.

Patterns in the Tutor's Advice to the Novices. As mentioned before, the tutor usually propagates ways of mathematical thinking — that range from practical advice to fostering a whole philosophical perspective. Studying the influence of these pieces of advice on the formation of the novice's mathematical behaviour might provide insight into aspects of the cognitive processes studied here.

Mathematics as a Debate. A large part of the novice's mathematical behaviour is determined by the way mathematics has been presented to them. Given the flexibility of the tutorial — technically it is almost entirely up to the tutor to determine its content and structure — these styles of presentation vary. Some general patterns are the following: tutoring as a proof-and-refutation process (optimisation of solutions/proofs is one example; tutor as a facilitator between two students who fail to communicate is another); the abundant use of the Socratic Method. Especially on the latter, it is worth noting that the novice is a naive believer and that the traces of an inquisitive/critical spirit are practically extinguished by this method which is highly tutor-driven and allows no exploratory deviations. It is worth investigating how the tutoring process of this kind hinders the novice's encounter with mathematical abstraction.

The (Un)Naturality of Concepts. Introducing new concepts to the novices is a process that most novices characterise as unnatural, that is lacking in providing with sufficient justification for the existence and the utility of the introduced concepts. This unnaturality appears as a major obstacle to abstraction. A cross-topic investigation might illuminate further details.

Other themes of a lesser range have also been primarily identified.

NOTE: The above approaches are tentative. The formation of the subsequent levels of analysis whilst analysis is taking place is inherent in qualitative research.

I note that the themes cited in the Transfer paper addressed a wide variety of issues (epistemological, contextual, psychological) related to the learning difficulties of the novice mathematician's encounter with mathematical abstraction. As a step towards the progressive focusing of the study, the subsequent stages of the data analysis more emphatically concentrated on the psychological issues but remained intensively — even though more implicitly — informed by epistemological and contextual issues.

PART II Data Processing After Data Collection

As explained in Part I data collection was completed in March '94 and my Transfer to DPhil status took place in April '94. By then Data Analysis was at the state of having constructed *Scripts* for the tutorials. During May and June and from September to mid-December '94 interviews and tutorials were transcribed, the former fully and the latter selectively. In the following I account for the analytical procedures during and after transcribing.

IIa. The Interviews

During data collection the interview material appeared to me as less intellectually intimidating and hence more manageable than the tutorial material. The interviews

- a had a pre-determined topical (mathematical) structure — as the Interview Scenarios in Chapter 4 illustrate,
- b consisted of Q&A discussions during which the floor was given almost exclusively to the students. That implied that, compared with the tutorials in which student discourse was diluted in a jam of tutor exposition, irrelevant/general conversation and other interferences, no purification process was necessary.
- c were potentially providing foci for the further stages of analysis. The mathematical topics touched upon in the interviews were selected on the basis of my preliminary reading of the fieldnotes in terms of mathematical areas that the students had demonstrated exceptional difficulty with. Therefore the interviews could confirm or question the correctness of this first reading.

On the grounds of criteria a, b and c, I decided to commence analysis from the interviews despite my initial intention to use them as supportive material to the tutorial material.

Full transcribing of the interview recordings was based upon

- repeated listening of the audio-recording of the interview
- the interviewees' writing during the interview which had been kept for reference.

As expected, criteria a and b contributed to an almost immediate ordering of the interview data. The expectation however expressed in criterion c was not fully satisfied as I explain in the following.

In May '94 the interviews were fully transcribed. In the subsequent months seven texts, related to

- Accumulation Points/Isolation Points,
- Openness/Closedness, Boundedness
- Limits
- Spanning Sets and Bases
- Convergence of Series and Sequences
- Compactness
- The First Isomorphism Theorem for Groups and Related Concepts

were constructed and two papers (on Accumulation Points (Nardi 1995), on Spanning Sets (Nardi, to appear)) have come out of further analysis. These texts are compilations of the students' conceptions as expressed in their definitions, examples and images of the concepts. In the interviews the students also discussed what they find elusive about some of these concepts. The interview material is used as supportive reinforcement of some findings from the tutorial material.

The interviews did indeed illustrate and elaborate upon the students' difficulties with the mathematical areas I had touched upon; they even incidentally illuminated some other areas. However the interview data were strongly topically-centred and this was in a slight opposition to the evolving psychological directions of the study. This is a study of some general cognitive phenomena that characterise the novice's experience of advanced mathematics and the interviews explored the students' perceptions of a number of specific mathematical concepts. Taken out of the context of cognition-in-action — as this action was observed in the tutorials — the students' discourse on these mathematical concepts can be seen as a supplement of the rich evidence of the tutorials. Therefore the interview material did not provide the foci of further analysis, as expected in criterion c, but potentially helpful evidence for the findings from the tutorial material.

IIb. The Tutorials

In the following I present an account of the construction of *Selective Transcripts*, the *Episode Extracting* and the subsequent stages of analysis that the extracted Episodes were submitted.

IIb.i The construction of Selective Transcripts

Selective transcribing of the tutorial recordings followed the transcribing of the interviews. Transcribing was based upon

- the *Scripts*
- a second non-stop listening of the tape
- the fieldnotes
- related documents
- doing the mathematics where it was deemed necessary after the second tape listening.

Immediately after the concerted study of the above, an extended and iterated listening of the tape led to the construction of *Selective Transcripts.* These protocols constitute the central bulk of data used in the subsequent stages of data analysis. In the following I explain the process of construction as well as the criteria on which it was based.

Unlike the interviews, which were fully transcribed, the tutorial material which was strongly characterised by the complexity of phenomena unfolding in a very natural learning environment, had to be submitted to an ordering and filtering process. The criteria of the selective transcribing process are crucial and I have tried to maintain them consistently throughout the construction of the *Selective Transcripts.* So, I excluded from transcribing but for the sake of continuity summarised:

- the tutors' exposition, in the form of long monologues, long technical calculations especially ones I had no visual access to,
- material on applied mathematical topics that my mathematical background did not allow me a good understanding of,
- general comments or informal chat.

As a result, the final *Selective Transcripts* have the benefit of continuity and are more focused than the *Scripts* which were non emphatic telegraphic texts. To allow quick reference to the audio recordings *Selective Transcripts* are numbered according to the recording equipment's counter.

So far the criteria that I have described as the basis for the selection of the material to be transcribed are largely structural, that is they are criteria relating to the preferred structure of the learning interactions that were transcribed in detail. From these criteria it is evident that my intention was to filter out of the complex tutorial material the instances of higher student participation; hence the parts of the tutorial recordings that are fully transcribed are mostly the dialogues between tutor and students and student monologues. In the following I explain how the selection of the material to be fully transcribed related to my reflection on the focus of the study.

Before addressing further the issue of selecting material for transcription that relates to the objectives of the study, I digress in order to refer briefly to the technical difficulties of transcribing. Having very little experience in transcribing, not knowing with precision which of the material would eventually turn out to be useful, and having pressing time constraints, I decided that a minimum of one (towards the end of the transcribing period: two) tape every day should be dealt with — so that by the end of 1994 the material would have been fully covered. This implied that where there were substantial difficulties (inaudibility, restricted understanding of the mathematical discussion) transcribing was left for later and these parts of the tutorial were summarised. Still in the cases where transcribing of some significant parts of the recording was postponed for the future, I concentrated on pointing out the necessity for further listening so that in subsequent readings of the *Selective Transcripts* these parts were revisited with extra care.

Considering both the material that was excluded from transcription and the material that was included, it is evident that the parts of the tutorials that mostly attracted my attention were instances in which the students participated in a protagonistic way (hence the preference for the students' monologues or dialogues with peers or tutor). Progressively it was becoming clear that, as far as the structure of the sought-for incidents was concerned, I was seeking learning events with a beginning, a middle and an end, *full stories on learning* (I note however that at that stage any evidence of the students' cognition, however fragmentary and incomplete, was also attracting my attention). True to the phenomenological agenda of the study, the aspiration here was to extract information about the problematic aspects of the learners' cognition from their expressions of this cognition. As far as this material is concerned I searched for these expressions in the students' interaction with their tutors and their peers during the tutorials. At this stage I realised that the objective recordings, juxtaposed with the less objective (but still avoiding decisive selection) *Scripts* and fieldnotes, were transforming into subjective, researcher-processed data. In other words the construction of *Selective Transcripts* was the first genuine attempt to impose my loose but substantial theoretical framework on the data. This personal imposition on the material increased through the various stages of Data Analysis and it is the researcher's task to guarantee that this process preserves and highlights the essence of the phenomena to be studied.

The subsequent stage of analysis was to determine which parts of the transcribed material would become the object of detailed analysis. The analysis would be informed by the theoretical tools developed in the field of Advanced Mathematical Thinking (elaborated upon in Chapter 1) but before reaching the stage of being able to use these tools, the material which would be submitted to this analysis had to be further ordered and purified. I quote from the Transfer Paper my first declaration of analytical intentions as tentatively made at that very early stage:

> *The current suggestion for this ordering has its origins in the nature of the tutorials that have been observed: as it turned out the majority of the tutorials can be structurally described in terms of the weekly problem sheets, that is most of the tutorials consist of the tutor's and students' elaboration on the questions in the problem sheets. Introducing new material, referring to lecture content and resolving queries, that constitute the main body of other activities carrying on during tutorials, are mostly within the frame of reference to question solving from the problem sheets. Thus an initial ordering of the data could take this somewhat* natural structure *into consideration.* A question-by-question rearrangement of the data so that they form a cross-tutor/cross-student presentation of answers *could be a way to carry this out. The diversity of approaches can consequently be reflected.*
>
> *The following question might then, tentatively, be addressed:* what are the components of the students' problematic encounter with mathematical abstraction? Concept Formation *and* Reasoning *are two areas that emerge from a primary scanning through the [data].*
>
> *Following up the analysis during the* Pilot Study, *a section of the* Chart of Incidents *— the taxonomy of the* Pilot Study *incidents — which consists of the student's difficulties and other aspects of their mathematical behaviour can be expanded and refined in order to incorporate the* Main Study *material.*
>
> ...
>
> *The triplet of* topical - logical - symbolic *difficulties has been kept. This categorisation is expected to facilitate a concept-centred perspective on the analysis of the incidents.*

The above intentions are heavily influenced by the then-recent experience of the *Pilot Study* and are also largely mathematically-centred in the proposed restructuring of the data. The concern for a global psychodidactical approach to the novice's cognition is also evident. In the following section I explain how the heavily mathematically-centred approach to the structuring of the data was reconciled with the global cognitive approach of the study.

IIb.ii The extraction of Episodes

Following the construction of the *Selective Transcripts* and the reflection upon the nature of the sought-for learning incidents, came a critical scanning of the transcribed material during which the tutorials were broken into their constituent learning incidents. A crucial step of that time was the decision to concentrate on the extraction of complete learning episodes. Already during transcribing this preference had begun to show. Episode extracting was generally based on

- the degree of participation of the students in the instance
- the potential of the instance to reveal aspects of the students' learning processes
- the completeness with which a specific piece of mathematics was dealt with.

The outcome of this selection was a list of 110 Episodes, a large number of which was double, triple, quadruple or fivefold incidents of the same piece of mathematics dealt with different students. The tabulated form of the list was arranged in three-columns:

College, Week - Course	Mathematical Content	Didactical Content

So an entry like

M3-LA	LA2.7	assumed unproved theorem, trouble with matrix properties, 'is it enough to show it for 3x3 matrices?', b/b dialectics

would mean that this Episode took place in College M on the third week of term, relates to the Linear Algebra course and in particular question 7 from Problem Sheet

No 2. There was no need to specify the term because this was implied in the course (Linear Algebra was only taught during the first term). The name of the participant student was occasionally mentioned in the third column when it seemed necessary. Also in the third column I cite a sample-reminder of the things that attracted my attention and usually were representative of the occurrence: so, for example, in the third column of the above table, 'b/b dialectics' means that one of the students was asked to present her solution of LA2.7 on the blackboard and that a debate on her solution followed; 'trouble with matrix properties' indicates the types of conceptual difficulties the student appeared to have in the tutorial; 'assumed unproved theorem' and 'is it enough to show it for 3×3 matrices?' indicates the student's reasoning difficulties. The last quote, which is a reproduction of the student's exact words is an example of a reminder, a strong impression of something which seemed to encapsulate the didactical substance of the Episode. I note that it is not true that the substance of all the extracted Episodes can be condensed in a quotation but these tentative condensations contributed significantly to a first global view of the Episodic material. I also stress that I made the concise commentary in the third column with caution because I wanted to avoid premature interpretations of the events in the Episodes.

Subsequently the tighter focusing of the criteria for the sought-for Episodes led to a further drastic reduction to their number. To illustrate how the tightening of the focus occurred I note that this first list of Episodes was compiled between January and May '95. In the meantime the drafting of the first chapters of the Thesis as well as processing of the interview material led to a reflection on the data that made further filtering of the Episodes necessary and more obvious. The main criterion that drove the formation of the first Episode List was a general pursuit of whole learning incidents initiated by a declared learning difficulty of the students, either announced by them or identified by the tutor. This criterion was quite general and as a result the extracted Episodes were of a large number and of mixed quality. So the next step was to decide how many of these Episodes were in fact giving evidence that allowed me to go beyond the identification of learning difficulties and access potential sources of the students' troubled cognition. For instance, in most of the extracted Episodes the cognitive conflicts that initiated the discussions appeared to be resolved by the end of the incident. Further exploration of the incidents could lead to discovering whether cognitive conflicts had actually been resolved or not.

The new filtering process was carried out after another listening to the recordings; also bearing in mind that the focus was now on Episodes that not only reveal learning difficulties, but appear as potentially permitting some access to the sources of the novice's cognitive conflict as well as possibly providing answers with regard to whether and how the conflict has been resolved. The 70 Episodes that were selected were tabulated in a similar three-column table to the one described earlier. The excluded ones were summarised and incorporated in the rest of the material that was beginning to take the form of auxiliary, supportive data. At the same time a few mathematical areas appeared as dominating the mathematical content of the material which until then was arranged chronologically and in accordance with the course and the Problem Sheet number: the foundations of analysis, calculus, topology, linear algebra and abstract algebra. The observations that determined the filtering also began to become more stable. A tendency to reflect on the incidents, while trying to filter further, into the conceptual, reasoning and notational/linguistic components of the students' difficulties, was forming (possibly reminiscent of the S.DIFF triad of the *Pilot Study*). Also I felt it was increasingly necessary that I am very comfortable with the mathematics of the Episodes. A pile of notes on the solutions of the problems began to accumulate. My reading of the literature was becoming increasingly triggered by the things I could see in the Episodes. Reading the *Selective Transcripts* and listening to the tapes was becoming increasingly loaded with expectations. Predictions of what the outcome of the Episode or the next line in the mouth of a particular tutor or a student would be were becoming gradually more successful; and I knew it was not simply because I had listened to the tapes again and again. It was because there was a clear emergence of patterned behaviours which was fermented by my constant familiarisation with the data.

IIb.iii The construction of Analytical Texts

While Episode extracting was approaching completion, I began to see that the analytical intuitions described in Part IIb.ii were in need of discipline and order. The 70 Episodes were physically extracted from the *Selective Transcripts*. They formed the main body of data, denominated *Episodic Material* and were filed chronologically. The rest of the material, the *Non-Episodic Material* was also filed chronologically, summarised/ tightened up further and kept for reference and as supporting material for the Episode Analysis. In the meantime further processing of the interview material had produced the two publications mentioned in Part IIa. I quote from (Nardi 1995) to illustrate the form that the processing of the 70 Episodes took in the subsequent months:

> *Each [Episode] is...reviewed individually and its analysis takes the form of a text which consists of the following sections:*
>
> > *Section (S): a summary that highlights the focal mathematical and didactical aspects of the piece,*
> > *Section (M): a presentation of the mathematics in the piece,*

> *Section (PSY): a psychological presentation of the highlighted points in the first section,*
> *Section (INTER): an interpretation of the psychological subtext identified in the previous section and*
> *Section (PED): a consideration of the pedagogical implications of the above analysis.*

Analysis is now at the stage of producing series of the texts described above with regard to particular mathematical topics. It is intended that the completion of this process will be followed by a period of overviewing and reflecting on the findings for each topic. The outcome of this reflective overview is expected to be the emergence of patterns in the observed learners' cognition. These patterns will constitute the themes along the lines of which a psychological, cross-topical profile of the novice mathematician's difficulties in the encounter with mathematical abstraction will be drawn.

Therefore each Episode was treated as an analytical unit. A text for each Episode was prepared consisting of the five sections cited above. The Non-Episodic Material helped form an introduction to each text regarding the context of the Episode. The fieldnotes and relevant mathematical documents helped form Section (M). The Episodes were rearranged in terms of the five topical areas (the foundations of analysis, calculus, topology, linear algebra and abstract algebra) that had emerged as dominating the mathematical content of the *Episodic Material*. In the process the 70 episodes were further reduced to 50 on the simple basis of excluding a few more mathematical areas (for instance Permutations in Abstract Algebra is the topic of discussion for a considerable number of Episodes and have been kept separately as material for future analysis and reference). From June to mid-July, September, October, and mid-November to mid-December the filtering of Episodes, their allocation to one of the five thematic/topical units and the construction of the five-section *Analytical Texts* was completed. In the meantime the *Analytical Texts* had taken the following form:

> Section (M) was appendixed or footnoted
> Section (S) was a heavily summarised factual account; only crucial dialogues were kept in
> Section (PSY) and Section (INTER) merged into a section called *An Interpretive Account of the Episode*.
> Section (PED) was almost made redundant since pedagogical, and particularly teaching, issues are marginally addressed in the *Interpretative Account* as part of the attempt to reconstruct the context within the learning incident occurred.

In the construction of the *Interpretive Accounts* a variety of tools, developed within the field of Advanced Mathematical Thinking, are used. The links between the theoretical underpinnings of the analysis and the actual analysis as presented in the subsequent chapters will be made individually in the *Interpretive Account* of each Episode. The interpretation of the incidents is also largely reinforced by the *Non-Episodic Material* and the *Interview Material*.

In the above I have attempted to illustrate the long and painstaking process of how 120 hours of recording transformed into 1500 pages of transcripts and distilled into the 32 learning incidents that constitute the main body of evidence for the theoretical cases of the study. In brief the presentation of the *Analytical Texts* follows mostly the structure: fact - interpretation - conclusions. Conclusions for each Episode inform the construction of a Conclusion Part for each of the four thematic/topical units featured in Chapters 6 to 9 (the reduction of the number of Episodes used as evidence to 32 and of the thematic areas from 5 to 4 was dictated by the space limitations of the thesis). Chapter 10 constitutes the cross-topical theorising part of the study and brings together the findings reported in Chapters 6 to 9.

Interlude

An Introduction to Chapters 6-9

Throughout the study the focus was balanced between the *topical* — that is, related to specific mathematical areas or concepts — and the *cross-topical* — that is, running through several mathematical areas — aspects of the novices' cognition. The arrangement of the analysis in five topical areas (Foundational Analysis, Calculus, Topology, Linear Algebra, Group Theory) is the overt evidence of the topical focus. Less overtly, this arrangement illustrates metaphorically the novices' journey from formal arithmetic (Foundational Analysis) to the advanced arithmetic of Calculus (enriched with the manipulation of the infinitesimally small or large quantities of the limiting process), then the transition from the numerical to the generalised set-theoretical contexts of Topology and Linear Algebra and, finally, to the total abstraction of Group Theory. This journey was also approximately chronological: the material on the Foundational Analysis comes from the first weeks of observation, whereas the material from Group Theory comes from the last.

The general intention of the study was to carry on with this balanced view between the topical and the cross-topical aspects to the end. As a result, within each area, emerged the, what I call, *paradigmatically problematic concepts*, that is the concepts towards which most of the students' cognitive concern and difficulty seemed to converge during observation. These were:

- *supremum and infimum of a set*
- *limit*
- *compactness*
- *spanning set*
- *cosets*

and were further explored in the interviews which were structured around reinforcing the evidence provided about them in the tutorials.

However, due to the limitations of space in the thesis this balanced view between the topical and the cross-topical aspects of the novices' cognition had to be abandoned. As a result, the selection of Episodes in Chapters 6-9, while keeping the topical structure of the initial data analysis, is geared towards the presentation of the cross-topical themes that emerged from — and directed — the various stages of analysis, rather than the specific learning difficulties within each topic. I note that some topical aspects of the analysis feature in (Nardi, 1994) and (Nardi, in press). Moreover, again due to the limitations of space, the topical area of Topology was left out of the presentation in the thesis: in the above described continuum (from Foundational Analysis to Group Theory) Topology – along with Linear Algebra – are the two middle areas in terms of the transition from the numerical to the abstract. Having to choose between Topology and Linear Algebra, the extent and strength of the analysed material on the latter and the possible elimination of the former from the first year undergraduate mathematics curriculum in subsequent years, resulted in leaving Topology out.

The Episodes presented in Chapters 6-9 should be read as paradigmatic cases of the cross-topical themes of which a synthesis is presented in Chapter 10 – that is, unless otherwise specified, the Episodes represent trends in the analysed material. Even the description of idiosyncratic cases serves the purpose of accentuating these trends. It is due to the richness and complexity of the naturalistic data collected in this study that the topical and the cross-topical elements are so intertwined that their distinction at times collapses – for instance, within the discourse on the novices' image constructing of new concepts. The topical analytical discourse has been toned down and kept only to the levels in which it serves the development of the cross-topical analytical discourse.

The themes, exemplified in Chapters 6-9 and elaborated upon in Chapter 10, can be concisely described here as features of the novice's encounter with mathematical abstraction. This encounter is seen in this study as an enculturation/cognitive process. The new culture is Advanced Mathematics and it is introduced by an expert mathematician, the tutor. The themes around which the analysis revolves relate to

- the novices' *concept-image construction* which is seen as
  a problematic interaction with the concept definitions and
  an attempt for the construction of meaningful metaphors and *raisons-d'-être* of the new concepts and the new reasoning,

- the tension between the informal-intuitive-and-verbal and formal-abstract-and-symbolic modes of thinking reflecting
  the *tension between verbal and formal/symbolic language* and
  the *tension between informal and formal modes of reasoning*.
  The difficulties in formalising have been identified to be
  difficulties with the mechanics of formal mathematical reasoning
  as well as
  difficulties of applying the mechanics of formal mathematical reasoning in a well-integrated and contextualised manner.

The above outlined enculturation/mental process into the reasoning mode of Advanced Mathematical Thinking was moreover strongly determined by the teaching environment within which it was observed, the tutorial. Finding that

- a didactical style dominated by exposition seemed to influence this *enculturation* — an interactively formative process — to the degree that it converted it into *acculturation* — an authoritative enforcement of mathematical expertise — process,

some of the analysis focused on elaborating the effect of this style on the novices' cognition.

Each Chapter contains 8 Episodes. In Part I of each Chapter a table is provided with a summary of relevant information on the Episodes. In Part II, each Episode is presented in one Section. Each Section consists of

- a title which expresses the main point made in the section
- a brief account of the context of the particular incident
- a factual account of the Episode containing transcribed parts
- an interpretive account of the Episode
- conclusion.

In Part III, I present a concise synthesis of the focal points that were highlighted in the analysis of each Episode. The syntheses in Parts III of Chapters 6-9 constitute the intermediate theorising stage between the presentation of the paradigmatical cases, the Episodes, and the global theorising of Chapter 10. I note that the didactical (relating to teaching) and methodological observations made in the Episodes are synthesised directly in Chapter 10 and not in Parts III of Chapters 6-9.

Throughout the analysis the relevant mathematical material — that is the mathematical questions discussed in the Episodes — has been coded and abbreviated as follows. For example:

CD2.1

means

Problem Sheet Number 2- Course: Continuity and Differentiability - Question 1

B7

means

Problem Sheet B - Course: General - Question 7

SS7.1

means

Problem Sheet Number 7 Course: Series and Sequences- Question 1

LA6.29
means
Problem Sheet Number 6- Course: Linear Algebra- Question 29

GRF5.8
means
Problem Sheet 5- Course: Groups, Rings and Fields- Question 8

In the following table I list the questions from the Michaelmas and Hilary Term problem sheets referred to in these Episodes:

Chapter 6	Chapter 7	Chapter 8	Chapter 9
CD2.1 - CD2.6 CD3.3	CD4.1 CD5.1 CD7.1 CD7.2 B6 SS4.1 SS4.3 SS4.4 B7 B10 SS7.1	LA5.23 LA5.24 LA6.26 LA6.29 LA7.35 B3	GRF5.1 GRF5.6 GRF5.8 GRF7.3 GRF8.5

These questions — and their answers — can be found in Appendices For Chapter 6, 7, 8 and 9. The answers to the mathematical problems provided in the Appendices are the solutions — that I have reconstructed either from the fieldnotes or the recordings — discussed in the tutorials.

Recommendation to the Reader: A familiarity with the mathematical content of the incidents is crucial in the understanding of the Factual as well as the Interpretive Account of the Episodes. A recommended reading technique is the following:

- **read the title, the context and the structure of the section**
- **read the Factual Account in parallel with the mathematical problems and their answers provided in the Appendices,**
- **read the Interpretive Account while consulting the Factual Account.**

Chapter 6

The Novices' Encounter With Mathematical Abstraction: Cases from *Foundational Analysis*

PART I A Guide to the Paradigmatical Cases (Episodes) Presented in this Chapter

The following table contains contextual information with regard to the 8 Episodes presented in this Chapter.

Episode Number	**Time of Incident Term - Week**	**Participants**	**Mathematical Content**
1	Michaelmas 3	Jack and Andrew	*Archimedean Property*
2	Michaelmas 3	Cathy/George and Ben	CD2.1
3	Michaelmas 3	Jack and Andrew	CD2.2
4	Michaelmas 3	Kelle	CD2.3
5	Michaelmas 3	Jack/Andrew, Kelle	CD2.4
6	Michaelmas 3	Alan, Connie	CD2.5
7	Michaelmas 3	Cornelia	CD2.6
8	Michaelmas 4	Jack and Andrew	CD3.3

PART II Data and Analysis

In the following I present the factual and interpretive accounts and conclusions for the 8 Episodes of the table in the previous page. In Part III then I synthesise the findings of Part II related to Foundational Analysis and briefly discuss the wider cognitive issues that are presented in the overall synthesis of the data analysis in Chapter 10.

Section (i) First Steps of Initiation Into Mathematical Formalism: Meaning and Proof of the *Archimedean Property*

Context: See Extract 6.1

Structure: After several attempts the students and the tutor agree on a formulation of the *ArchPr*. With the help of Andrew, Jack completes a proof of the property on the b/b. The tutor critically reviews his writing style and they discuss alternative formulations of the property.

The Episode: A Factual Account. See Extract 6.1

An Interpretive Account: The Analysis

Formulating the ArchPr. The *ArchPr* says that $\forall x \in \Re\ \exists n \in N\ x < n$. A1-J1-A2 feature the students' attempts to reconstruct a formulation of the *ArchPr* and T1-T2 are the tutor's refutations of these attempts. A1 is a partial reconstruction of the *ArchPr* that maintains its logical structure (the two quantifiers and the inequality) but not its meaning. J1 is an attempt to defend A1 by adding more information on y. A2 is a response to the tutor's objection to J1, that however close to x Jack chooses y, he can choose y even closer: what Andrew says is that in this case he can get even closer to x than the tutor with his new y. The rather meaningless circularity of the students' arguments is dismissed by the tutor who halts the interactive process by asking them to look up an acceptable formulation of the *ArchPr*. I note here that that same week other students have had difficulty with reconstructing the *ArchPr*. An illustrative attempt is student Connie's in Extract 6.6: 'The *Archimedes' Principle* isn't saying that there exists this number >1, bigger than any number?'.

The mathematical meaninglessness of the students' attempts does not however deprive the interaction of its didactical interest as a dynamic interactive process of conjectures and refutations. The dynamics of this process are intriguing because the students are both concentrating on reconstructing the *ArchPr* as well as to defend themselves and each other to the tutor. J1 initiates the mutually supportive cognitive action of the students and A2 accentuates their complementarity. So in a sense their exchange of words highlights both cognitive (attempt to reconstruct the *ArchPr*) and socioaffective (support each other and self in order to confront the tutor's objections) aspects of the students' thinking process.

Subsequently the tutor's question about what would be the implication if the *ArchPr* did not hold is an attempt to highlight its meaning through its negation ($\exists a \in \Re$ $\forall n \in N\ n \leq a$). Possibly Andrew tries to negate the *ArchPr* in his mind and, carried away by $n \leq a$ and ignoring the quantifiers determining n and a, claims that the negation means that the real numbers are bounded above. In this and subsequent chapters the novices' difficulty with negation and quantified statements will be repeatedly mentioned.

Proving the ArchPr: an Interactive Procedure. Jack suggests reaching contradiction by assuming the *ArchPr* is false. Then N will be bounded above and, since it is non-empty, by the *Completeness Axiom* it will have a supremum. He then seems unable to continue. Andrew suggests then applying the definition of supremum for $\varepsilon=1$. Jack, preoccupied with what he has written on the b/b so far, ignores the suggestion and repeats the definition. The tutor intervenes in order to recommend the use of Andrew's suggestion. In the next few seconds Jack overcomes the impasse he appeared to have reached and completes the proof. Jack accomplishes at first to construct accurately the supposition that the *ArchPr* is false. This supposition implies that N has a sup. He also reconstructs the definition of supremum but he cannot co-ordinate the two elements (falsity of *ArchPr* and definition of sup) so that the contradiction emerges. The co-ordination finally takes place in the few seconds of him pausing to think after the tutor insists that he uses Andrew's suggestion. So, in a sense, Andrew and the tutor have discreetly scaffolded Jack's final overcoming of the impasse without heavy-handedly imposing it upon him. With their interventions they have created what appears to be suitable for Jack conditions that helped him find his way out.

Proving the ArchPr: on Jack's writing. The complementarity and mutual support of the two students is also illustrated in Andrew's defending Jack's writing with 'this is what the lecturer wrote'. The problem though is that Jack has been reproducing on his draft and then on the b/b what, according to the tutor, is the lecturer's b/b technique. The tutor encourages the students to modify this abbreviated, semi-colloquial writing style to a more flexible, narrative and descriptively richer writing style. In fact the tutor's critique reflects a contemporary trend in the didactics of advanced mathematics according to which the learners are encouraged to express their ideas in words and not simply in quantified and categorical notation. The reason I present the tutor's critique in detail is that it reflects the use of formalism by the novices which at this stage can be seen as awkward. It also illustrates how the students are urged with lessons like this to a wiser interpretation and manipulation of mathematical formalism.

For instance Jack's use of quantifiers in $\exists\, sup N \in \Re$. is rather problematic: it seems to be a literal translation of 'the supremum of N exists'. As far as the use of n_ε is concerned, the tutor's comments have a more controversial effect. This is a notation, he claims, that might induce the impression that there exists a function between ε and n. I think that it is equally likely to induce the impression simply that n depends on ε. Jack seems to be disturbed by this ambiguity as well as Andrew who reckons whether this is a good reason to avoid this notation.

The whole Episode stands as a metaphor for the maturation process which the novices must go through in their first encounters with mathematical formalism. In the above, the tutor is not simply teaching a proof of the *ArchPr*: he is setting up an initiation into the idiosyncratic precision and rigour of mathematics as

- his critique of Jack's writing,
- his speech on the essence of the *ArchPr* and how it affects the system within which the Archimedean mathematics works, and,
- his cautionary remark towards the end on the naturality of defining ϕ to be $1/\varepsilon$ (and not ε to be $1/\phi$)

illustrate. The novices are naturally overwhelmed by the dazzling caution for detail that this new approach to mathematical thinking (and expressing) entails.

Conclusion: In the above, the formulation and proof of the *ArchPr* stimulated a student-tutor interaction process mostly characterised by the tutor's enculturating interventions with regard to the students' reasoning (handling quantified propositional statements: negation and co-ordination to reach contradiction) and expression. From a didactical point of view, a dialectically successful instance was examined where the tutor and a student discreetly scaffolded another student's overcoming of an obstacle in a proof.

Section (ii) The Problem of Clarifying What Knowledge Can Be Assumed in a Proof and the Role of Quantifiers in Establishing the Generality of a Proof

Context: See Extract 6.2

Structure: In the following Ben presents his solution. The tutor refutes his proof. George then outlines his proof. The tutor refutes his formal writing and presents a proof. A discussion follows on the role of quantifiers in establishing the generality of a proof.

The Episode: A Factual Account. See Extract 6.2.

An Interpretive Account: The Analysis

What is problematic with Ben's approach. As the tutor observes Ben has attempted a direct application of part i in his proof for part ii. In this, he has ignored that the pair of arbitrary numbers used in part i are rational, whereas in part ii they are real. He interprets 'if x is irrational' as 'there exists x which is irrational' (fig.2a), assumes then that a/x and b/x will then be irrational, with a and b being rational, and, via part i, he deduces the existence of s, a rational number between them. So, by the moment he is interrupted by the tutor, Ben

- has ignored the requirement for choosing arbitrarily the two numbers to start with (his a/x and b/x are not arbitrarily chosen real numbers) and,
- he is using the statement that between two irrational numbers there is always a rational number.

Therefore Ben is committing two serious formalistic errors:

A he ignores the requirement for a universal quantification of his proof (for <u>any</u> two real numbers).

B he is using a previously unproved statement either as a misinterpretation of part i or in effect of lack of rigour (that is, in the absence of the realisation that previously unproved statements cannot be assumed in a mathematical proof).

A Note on Cathy's Questions. Cathy remains non-participant during the Episode, possibly because she has not answered the question in her drafts and thus feels that she has no contribution to make. However towards the end of the tutor's presentation she sounds preoccupied with how his proof preserves the universality of a and b. Like Ben (A) Cathy is not comfortable yet with universal quantification.

Moreover if Ben (B) under-reacts to the requirements for rigour (by assuming more than allowed), Cathy seems to over-react (by questioning the assumption of factual knowledge, such as the existence of irrational numbers). In both cases, over- and under- reaction constitute evidence of an unease with issues regarding the knowledge one is allowed to assume.

What is problematic with George's approach. George's initial idea to base the proof for real numbers a and b, in part ii, on a generalisation of the available information on rational and irrational numbers is a demonstration of more confident reasoning than Ben. His is a problem of expressing mathematically a general idea (G1) to which he was led by intuition. So,

In his mind he possibly wanted to prove:	*between any two real numbers there is a rational number (1)*
In his draft, as the tutor says, he started doing so by supposing:	*there is a rational number between two real numbers (2)*
In the tutorial he reads his draft:	*as if he had written (1). But he has written (2).*

(2) is a fact, which the tutor emphasises by pointing at 0 and 1 as two real numbers between which there is an irrational number. George then realises (G2 and G3) that there is a substantial difference between his intentions and the reflection of those in his writing. Again, the problematic aspect of the student's expression, here in writing, has been one regarding universal quantification of mathematical statements.

Conclusion: In the above, the students are not at ease with the assumptions they are allowed to make when engaged in proving fundamental statements (e.g. CD2.1): they have found it hard to distinguish, formulate and prove a universally quantified statement (they do not appear ready to choose arbitrary numbers, establishing thus the universality of their proof, and maintain this arbitratiness through the proof with consistency). Ample similar evidence on CD2.1 is available from other tutorials too. CD2.1 seems to be of a particularly problematic nature mostly because it is not clarified to the students what statements regarding the real numbers can be assumed. As seen in Cathy's and Ben's over- and under- reaction to the requirements for rigour, students seem to be vulnerable to issues related to assumed knowledge: in other words they have been sensitised to the increased requirements of rigour in the new course but then abandoned to clarify these requirements on their own. This raises the didactical question as to how, in this state of uncertainty about the rules, are the novices expected to play successfully the mathematical game of foundational Analysis.

Section (iii) Mathematical Induction and the Triangle Inequality: Cultivating More Fruitful Uses of Intuition and Hindsight as Features of the Shift to More Expert Mathematical Practices

Context: See Extract 6.3

Structure: In the following the tutor presents a 'back-to-basics' proof of the Base Statement of the Mathematical Induction for the triangle inequality as opposed to the geometric proof given in the lectures. The students then discuss the role of the ≤ sign and the necessity to start this inductive proof from n=2 instead of n=1.

The Episode: A Factual Account. See Extract 6.3

An Interpretive Account: The Analysis

A Cognitive Shift Towards More Rigorous Mathematical Practices. This Episode, mostly through the tutor's exposition on

- the need to reduce the interference of intuition with the construction of mathematical arguments, and,
- the foundational requirement for economy of principles and conciseness in mathematical reasoning,

exemplifies the change of mathematical culture that the novices are required to go through. The moral for the novices here is captured in the tutor's words on intuition: 'it's the right way to understand it and remember it, but it's not the right way to prove it'. The substantially different role of intuition in the expert and the novice mathematician's mind is illustrated in this Episode if one juxtaposes the left and right columns of the following table:

Jack's and Andrew's reference to the lecturer's use of a geometrical picture in his proof for the triangle inequality for two numbers	(a) the tutor's use of the same picture — as a tool for comprehension — in his response to Andrew's question on the meaning of ≤ in the triangle inequality and (b) the tutor's pragmatic response to where the proof by mathematical induction of the triangle inequality should start from (the triangle inequality 'is a basic assertion for two. It then has some information in it. For one it doesn't')

For Jack and Andrew the lecturer's picture is a convincing argument for the case n=2 whereas for the tutor it is a vivid illustration of how the triangle inequality works and only a starting point for the mathematisation of the argument (the tutor looks at the proof for $n=2$ in view of the prospective proof for $n=k$).

A Mathematical Mind in Genesis. Like in Extract 6.1, Andrew's evolving mathematical persona emerges as that of

- a competent executioner of mathematical operations. The competence demonstrated in Andrew's elaboration on the eight possibilities in the proof of the triangle inequality for two numbers is remarkable: not only does he point out to the tutor the possibility of reducing the number of cases to be written down but also helps the tutor realise what this number is.
- a practitioner. When the tutor criticises the students' use of squaring and of a geometric picture for the proof of the triangle inequality, Andrew's first reaction is to wonder whether this 'is not the right thing to do'. The strong practical action-centred semantics of words such as 'thing' and 'to do' and in particular the use of the word 'right' is an indication that Andrew is in search of rigorous recipes.
- an observant formalist. These are the two observations that Andrew makes in this Episode that initiate dynamic interaction among the tutor and the two students:

Observation 1. In A1-A3 Andrew seems to think that the $\leq$ symbol in the triangle inequality for $n=2$ is used redundantly because what actually is intended is to say that, depending on the sign of x and y, $|x+y|$ is equal to one of $\pm(x\pm y)$, whereas $\leq$ gives the impression that the whole range of values less than $|x|+|y|$ is covered.

Observation 2. In A6 he wonders whether starting the proof by Mathematical Induction from $n=1$ invalidates the proof which the tutor suggests they start from $n=2$.

As far as O1 is concerned, as Jack and the tutor observe, $\leq$ might be 'discarding information' but 'it is true' as well. It is also a representation that can be uniformly transferred to the context of complex numbers and vectors. Therefore by hindsight it is a functional and economical symbol to use. Similar lack of mathematical experience seems to have led Andrew to O2: the general form of Mathematical Induction allows the proof of statements for all $n\in N$ where $n>k$ and k is a natural number. Both students in this Episode seem to think that Mathematical Induction compulsorily covers all natural numbers: hence Jack started his proof from $n=1$.

The idea seems to linger in Andrew's mind in A7: he still seems to believe that to use the statement for n=2, one has to prove the statement for n=1.

In any case Andrew's observations are acute and insightful and, even though they are not flawless, they convey an impression that Andrew is being efficiently enculturated into the new status-quo of rigour and precision of university mathematics. In other Episodes, nevertheless, evidence is provided of how a similar attitude, when taken to pedantic extents, occasionally has hampered his efficiency and his mathematical growth.

Conclusion: In the above, the novices are in need of, and engaged in, a cognitive shift from unrigorous to rigorous forms of mathematical thinking. In particular a discussion of the meaning of $\leq$ in the triangle inequality and the Base Clause of Mathematical Induction gave rise to the students' metamathematical queries and highlighted how mathematical experience empowers hindsight as well reinforces a more fruitful use of intuition.

Section (iv) The Problem of Clarifying What Knowledge Can Be Assumed in the Context of an Application of the *Completeness Axiom*

Context:	See Extract 6.4
Structure:	In the following the tutor criticises Kelle's solution for CD2.3 and by closed questioning she elicits from him a proof for CD2.3i. The tutor then proves part ii by exposition and comments upon Kelle's solution: his idea was 'right' but the proof was not formally correct. Also he has misunderstood what exactly he was allowed to assume. In his final comment Kelle seems to have followed the tutor's exposition.
The Episode:	A Factual Account. See Extract 6.4

An Interpretive Account: The Analysis

The Impact of the Tutor's Attitude Towards Kelle's Written or Oral Suggestions. The tutor's approach here has such a tremendous impact on the student's approach to the tutorial that the shift from his intended to the eventually demonstrated style is visible. The tutor sets out with criticisms on his written proof for CD2.3 and elicits from him, with closed questioning, the definition and properties of the infimum of a set. Kelle complies but soon returns to his proof which he seems to wish to discuss. The tutor ignores his suggestion and outlines hers. During her exposition he interrupts her in order to highlight what he thinks went wrong with his proof, namely that he didn't know he could assume *Completeness*. The tutor then continues her exposition, this time with a tighter focus on his misunderstanding on assumed knowledge. Again in the end of her exposition he takes the initiative to recapitulate the syllogism for the proof in part ii (K5).

In the above, I think there is an underlying conflict of preferred dynamics. The student implicitly expresses a need for a stronger dialectic tutoring style (that is a tutorial in which the discussion of his own individual oral and written suggestions is given priority). The tutor on the other hand only slightly adapts to the student's implicitly expressed needs and her exposition indicates that she prefers a straightforward presentation of *correct* answers instead of discussing problematic ones.

Kelle's Misunderstood Approach. Kelle did not know that the *Completeness* of the Real Numbers (and in particular the version he knew: every bounded above set has a supremum) can be assumed and consequently re-formulated for sets that are bounded below. Hence he decided to re-write the statement-to-be-proved in a way that involves the supremum of a set which reminded him of *Completeness*. He then set out to prove *Completeness*. So he did not exactly, as the tutor contends in the beginning of the Episode, 'state what the question asked him to prove'.

So Kelle, despite having 'the right idea' of the proof,

- first, ignores one of the rules of the foundational game in Analysis, namely that he can assume the *Completeness* of the Real Numbers
- secondly, tacitly assumes that the infimum of a set is contained in the set. When asked about it, he doesn't give the impression that he has been preoccupied at all with whether a set, that has an infimum, contains it or not. Unconsciously however he seems to have adopted the belief that a set contains its inf. Similar evidence is available in the literature (see Chapter 1, IIIc.iii) in the context of limits of sequences.

As a result Kelle's proof has not taken off from its primary, intuitive grounds and has remained incomplete.

Kelle in general (see Context) seems to be a very reluctant formalist: he prefers the formulaic and concrete tasks of the Probability course and he is achieving a much better performance there than the 'mind-blowingly' rigorous course on Continuity and Differentiability. His handling of proof also suffers: as the tutor comments (see Context): he has proved the triangle inequality for $n=2$ and assumed that by simply saying 'by Mathematical Induction it is true for $n=k$' his proof was complete. Similarly to Extract 6.3, the proof that the lecturer used (via a geometric representation of the triangle inequality for $n=2$) acted as a strongly convincing argument for the novice who did not feel the need to strengthen his belief. As a result, Kelle added 'by Mathematical Induction it is true for $n=k$' as a formality only because he knew he was expected to cover all natural numbers. In sum Kelle does not behave as if he is convinced of the necessity for proof: once he is convinced, for instance by an illustrative picture, he seems to want to abstain from further action related to proof.

Conclusion: In the above, an application of the *Completeness Axiom* revealed conceptual difficulties related to the infimum and the supremum of a set and illuminated a student's perplexity with the knowledge one is allowed to assume (the novice wondered whether he should prove the *Completeness Axiom* and whether he could turn the Axiom around for infima). The student also appeared reluctant to formalise. He also relied on concrete, intuitive arguments. From a didactical point of view the student also seemed to incline towards discussing his flawed strategies but the tutor in most cases seemed to prefer an exemplary presentation of flawless answers.

Section (v) Preliminary Conceptions of Limit and Infinite Largeness. The Two-Step Battle Between Intuition and Formalisation: Conceptualising and Materialising the Necessity for Formal Proof

Context:	See Extract 6.5i and ii
Structure:	In Kelle's Extract the tutor criticises the student's written approach to CD2.4 and then they discuss the proof of parts i and ii. In Jack's and Andrew's Extract the tutor is disappointed with their treatment of CD2.4ii. The students subsequently attempt unsuccessfully to defend Andrew's approach and the tutor presents two alternatives for the proof. The latter sparks off a metamathematical discussion.
The Episode:	A Factual Account. See Extract 6.5i (Kelle) and Extract 6.5ii (Jack and Andrew)

An Interpretive Account: The Analysis

Kelle's Problematic Handling of the ε-Definition of $Lims_n$. (On Extract 6.5i) Kelle in the beginning of the Episode is rather uncomfortable with the tutor's rephrasing of CD2.4i in terms of limits, sequences and convergence. Despite his implicit request to stay within a familiar lexical territory, during the incident the tutor occasionally returns to this terminology. However Kelle's major difficulty seems to be the manipulation of the ε-definition of limit.

K1 and K2 are hesitant and confused attempts to reproduce the proposition in CD2.4i and Kelle sounds baffled with the quantifiers for n and ε. K1, K2 are followed by the tutor's objection (T1) and her pictorial exposition on what CD2.4i is actually about. K3 is then received by the tutor as progress (T3) and she proceeds with elaborating on the nature of ε. K4 is evidence of how undecided Kelle is about the nature of ε: ε can be any number, ε is a number we can choose, ε is a number we must find. T4 then is an attempt to establish a connection between Kelle's intuitive knowledge that $\theta^n \rightarrow 0$ and ε as a tool to formally express this knowledge. K5 and K6 indicate that this connection has not been established. Kelle is convinced that $\theta^n \rightarrow 0$ and he does not hesitate to use this to-him-established fact (K6) in his vague attempt to answer the tutor's formal question about proving it. When the tutor intervenes in order to interrupt the vicious circle of his thinking and completes the proof, he listens and concludes in a dramatically expressive frustrated tone 'It's so obvious that it goes to nought'. In part ii, prompted by the tutor about $\eta>1$ implying $1/\eta<1$, Kelle vaguely suggests 'inverting' which the tutor interprets as applying CD2.4i on CD2.4ii putting $\theta=1/\eta$.

As an observer I am not convinced that Kelle has been persuaded of the necessity or has learned how to present his intuitive ideas formally, namely via the ε-definition of limit. In Chapter 7 (Calculus), ample evidence is provided of the hardship such learning entails. The purpose for presenting this Extract in this Chapter (on the Foundations of Analysis) is to begin drawing a picture of how Analysis as a tool for quantifying and manipulating relations appears as a very difficult language for the learner from the very start. Moreover the difficulty of the language seems to be worsened by the novices' reluctance to accept its necessity and utility.

The Success of Interactive Dynamics and a Noble Failure. (On Extract 6.5ii) Andrew's intention is to prove that S has no upper bound by assuming it has and then reaching contradiction. Unfortunately this intention is not clearly stated by Andrew and it takes several interventions by the tutor to clarify it (up to and including Andrew's writing on the b/b). Then Andrew's thinking collapses: η has been given as a real number >1. Andrew thinks that this allows him to take η as a natural number n+1 and write η^n as $(n+1)^n$. In the latter not only η loses its generality (from an arbitrary real number it becomes a natural number) but it also acquires an illegitimate double meaning:

- first, (remember: S was defined as the set of all powers of η^n) n runs though N. And,
- second, n is defined by Andrew as a specific natural number used to substitute η with n+1.

Andrew's confusion is recognised by the tutor and Jack who laughs. Andrew resigns but, as his next question to the tutor indicates, in his mind the idea of contradiction still lingers. The tutor also is still open about investigating the idea of approaching the proof by contradiction and Jack also tries to defend it. Jack actually tries to pin down the contradiction by constructing a number, η^m, which he suggests then they try and prove it's the largest number in S. So, in a sense, Jack tries to develop Andrew's idea a step further, or in the tutor's earlier words 'to justify his belief by showing some strategy': Andrew has suggested assuming that S has an upper bound and Jack suggests a number which they can try and prove it's an upper bound. He then hopes that on the way they shall reach contradiction. Jack's suggestion is dubious in two ways:

- first, even if η^m is proved not to be the largest number of S, it is not implied that any other element of S is not its largest number, and,

- second, as the tutor notes, if m is an upper bound of S, it is not then necessary that η^m is the largest number of S.

So in the first, Jack appears to be unaware of the necessity to prove that S has no largest number in general — and not to choose one number, for instance η^m, and then prove it is not the largest number of S; and in the second he has unjustifiably assumed that, if m is an upper bound of S, then η^m is the largest number in S.

Andrew's subsequent suggestion that η^m is accessible but not attainable sounds like a finitistic attempt to express the arbitrary largeness of η^n and is perhaps his first digression from the idea of contradiction. He thus rephrases the sought-for result to 'trying to prove that η^m is the largest number in S'. Andrew's 'you can actually... there' is an expression of his intuitive idea of the arbitrary largeness of η^m but this is all too vague for the tutor to accept as plausible arguments.

In the above (apart from the students' interesting conceptions of infinite largeness and apart from Andrew's confused semantics in his attempt to construct a contradiction for CD2.4i) what looms as intriguing is the dynamics of the didactic interaction. Andrew's strength of belief in his proof by contradiction gradually diminishes but is not annihilated. Jack laughs with Andrew's flawed argument but tries to take the idea of contradiction a step further. Despite the final failure, the novices here exhibit a tendency to act on refuted arguments and modify them until they become adequate. In contrast to other Episodes in this Chapter, where the inadequacy of the novices' arguments has been pointed out to them, here, given a forum to discuss these arguments, they reach this understanding in a mutually critical and argumentative procedure. Subsequently, when the tutor decides to pursue no further the idea of contradiction and suggests two different proofs for CD2.4ii, the students, possibly having the urgency of their dynamic action removed, resume their traditional, passive position and only participate when asked closed questions.

The students' interaction with the tutor becomes again more dynamic in the presentation of the tutor's second proof for CD2.4ii with Jack's suggestion to 'expand' $(1+\zeta)^n$. This suggestion comes to conflict with what the tutor had in mind (to use the inequality $(1+\zeta)^n > 1+n\zeta$ proved in the previous week) and he insists on trying to elicit this from the students. With Andrew's distracting suggestion ('factorisation') the tutor, possibly in fear of losing completely the attention achieved so far in the group, retreats and accepts Jack's suggestion. In the end he reminds them of the inequality they could have used instead of the binomial expansion. So in this case the dynamic interaction between the tutor and the students has resulted in a redirection of the proving procedure towards the students' more basic choice than the tutor's (that is, Jack's suggestion to use the Binomial Theorem — with which he is already quite familiar — as opposed to the more recently introduced inequality suggested by the tutor).

A Comprehensive Dialogue on the Novices' Major Metamathematical Trouble. The concluding part of Extract 6.5ii is a dialogue that consolidates in the most illustrative way the students' difficulty with distinguishing between assumed knowledge and knowledge to be proved. Because it captures what emerges as one of the themes underlying my discourse on the problematic aspects of the novices' encounter with mathematical abstraction at the beginning of their studies, it will not be further discussed here but it will be incorporated in a later, more theoretical discussion of the issue.

Conclusion: In the above, a discussion of a novice's problematic handling of the ε-definition of limit revealed his reluctance and difficulty in formalising what he thinks as an obviously true proposition. Moreover, and in the context of discussing the students' conceptions of infinite largeness, their ambivalence on what knowledge can be assumed and how assuming knowledge is compatible with the demands for axiomatic rigour made by the lecturers and tutors in the beginning of the course were touched upon. From a didactical point of view, dynamic interaction among the tutor and the students proved a fruitful way of refuting the students' flawed approaches. On the other hand, exposition seemed rather inevitable for the presentation of *correct* proofs.

Apart from the common mathematical content, what links the two Extracts is the contrast between the informal, intuitive and confused approach of the student in Extract 6.5i and the more willingly formal approaches of the students in Extract 6.5ii. In none of the Extracts do the students achieve fully a formal presentation, but what differentiates the first student from the other two is that the latter seem to have conceptualised the necessity to be formal and struggle with the materialisation of this conceptualisation, whereas the former is still engaged in the vicious circle of assuming in his proofs what is to him intuitively obvious and what he is actually being asked to prove. The different states of mind of the students in these Extracts imply their different cognitive needs. This diversity implies the necessity for a didactical flexibility with regard to the accommodation of cognitive needs.

Section (vi) The Unsettling Character of the Logical Conjunctions in the Definitions of $S\cup T$ and $S\cap T$ and the Complexity of the Notion of Supremum: the Varying Persuasion of Mathematical Arguments and the Importance of Semantic and Linguistic Clarity

Context:	See Extract 6.6i and ii
Structure:	In Extract 6.6i the tutor talks to Alan about what is wrong in his proof and then presents parts of CD2.5. In Extract 6.6ii, at Connie's request, the tutor presents parts of CD2.5.
The Episode:	A Factual Account. See Extract 6.6i (Alan) and Extract 6.6ii (Connie)

An Interpretive Account: The Analysis

Alan's Problematic Proof. In fig.6b I present a reconstruction of Alan's proof and note that I find the third line of his proof hard to understand (why does $\alpha\notin S\cup T$ imply that $\alpha<\sup S$?). Alan then chooses a particular x in $S\cup T$ and, taking advantage of the given options (that then x can be in either S or T), he assumes that x is in S. Then he claims he reached contradiction by observing that, by the definition of $\sup S$, x must be less than $\sup S$ and, by the definition of α, x must also be less than α. From the third line of his proof α is less than $\sup S$, therefore Alan claims that he found a number, α, which is less than $\sup S$ and for $x\in S$ $x<\alpha$, which contradicts the definition of $\sup S$ as the lowest upper bound for S. I think that, contrary to the tutor's contention that Alan is making a claim about a specific x and then illegitimately twists his claim to a general x, Alan's claim is general about $x\in S$. I also think that the contradiction works if one assumes that α is not the maximum of $\sup S$ and $\sup T$, that is if one assumes that $\sup S>\sup T$ and chooses α to be $\sup T$, that is the smaller of the two. What is not clear is Alan's assumptions on which he tries to build the contradiction; also the third line of his proof blurs the presentation of his assumptions even more. In sum, I think that the tutor, confused by Alan's presentation, may have given a precipitate dismissal of Alan's idea. It is also possible that the fourth line of Alan's proof evokes the impression to the reader or the listener that the student switches illegitimately from a general element of $S\cup T$ to those elements of $S\cup T$ that are also elements of S.

The possibility that Alan, the tutor and myself do not seem to agree on whether Alan's approach is legitimate, highlights the intricate character of foundational analysis and by implication of axiomatic mathematics. From a semantic point of view, the lack of clarity in Alan's presentation (third line) as well as his inclination towards potentially misunderstood statements (fourth line) emphasise the imperative need that the novices acquire the skill to manipulate the very powerful linguistic and notational means of expression that formal mathematics can offer. If Alan's idea for the contradiction in CD2.5i is correct, then it has been sadly let down by his presentation; if it is not, then the ambiguity in his presentation has prevented the tutor — and possibly me — from pinpointing successfully its problematic aspects.

Connie's Problematic Conceptualisation and Manipulation of $S \cup T$. One of the possible problems with Alan's proof above was the lack of clarity in selecting an element in $S \cup T$ and maintaining the arbitrariness of this selection throughout the proof. Connie in Extract 6.6ii sounds also concerned (C1-C3) with selecting an element in $S \cup T$. She also sounds as if she is concerned with losing the arbitrariness of this selection if x belongs either to S or T. As a result she seems to suggest that x must be chosen as an element of $S \cup T$ which belongs to both S and T. What Connie is missing then is that exactly by choosing x as an element of both S and T, she is definitely losing the arbitrariness of her choice.

In sum, in both occasions, the definition of $S \cup T$ seems to engender an unease in the students' expressions which relate to CD2.5i. The unease seems to originate in the bi-lateral form of the definition ($x \in S \cup T$ iff $x \in S$ <u>or</u> $x \in T$) and the students' uncertainty about where x belongs (S, T, both S and T?) when chosen as $x \in S \cup T$. This multiplicity of options in the definition is perhaps unsettling as especially Connie's responses in C1-C3 illustrate. This unsettling emotion possibly results in the students' unsure handling of the related logical conjunctions and interferes with the formality of their proof.

The Non-Equivalent Psychological Power of Two Counterexamples. To prove that a set has a supremum, one has to prove that the set is non-empty and that it is bounded above. If any of these two conditions does not hold then the set has no supremum. In CD2.5ii to prove that $\sup(S \cap T)$ is not necessarily equal to the $\min\{\sup S, \sup T\}$, one can either show that $S \cap T$ does not necessarily have a supremum or that even if it does, this does not have to be $\min\{\sup S, \sup T\}$. In the first case one can prove that $S \cap T$ has no supremum by illustrating a case in which for instance $S \cap T$ is the empty set; in the second case by pointing at two sets such as $\{1,2\}$ and $\{1,3\}$. Both are logically equivalent and satisfactory counterexamples. They nevertheless seem to differ in persuasiveness: in the tutor's words (end of Extract 6.6i) the latter is 'slightly more convincing'. Similarly, students in other tutorials, that I observed but are not reported in this chapter, claimed that they did not stop looking for a counterexample for CD2.5ii until they finally thought of a pair of sets, such as $\{1,2\}$

and $\{1,3\}$, even though they kept coming across more examples of the first case, namely pairs of sets with an empty intersection.

A possible interpretation for this reluctance towards counterexamples of the first kind, that is pairs of sets of an empty intersection, is

- first, that the counterexamples of the second kind refute $\sup(S\cap T)=\min\{\sup S, \sup T\}$ as opposed to the more formalistic refutation achieved by the counterexamples of the first kind. In other words, in answering a question such as 'is it true that $a=b$?', students and tutor here have expressed a psychological preference for
  showing a case where $a\neq b$ (second kind)
  instead of
  showing a case where a does not exist (first kind).

- Secondly, in the case of CD2.5ii the counterexamples of the first kind did not question the upper-boundedness of the intersection but its non-emptiness. When discussing however the existence of the supremum it seems that, even though the two conditions, upper-boundedness and non-emptiness, are logically equivalent, in the mathematician's mind the former carries more weight than the latter.

So, in the above, I have tried to illustrate the difference between the epistemological and the psychological grounds of the tutor's and the students' preference. Epistemologically the counterexamples of the first kind are equivalent to the counterexamples of the second kind; psychologically however they do not seem to be equivalent. The significance of this distinction lies in the fact that the students, even after finding counterexamples of the first kind, continued to search for a psychologically satisfactory answer, that is counterexamples of the second kind. That is the priority of their own sense of conviction overshadowed the execution of the strictly mathematical task (to find a formally acceptable counterexample). In other words, the personal and subjective took over the priority from the impersonal and objective.

Connie's 'that's it?': a Reduced Concept Image of the Supremum of a Set. In order to prove that a number is the supremum of a set one has to show that this number is an upper bound of the set and also that it is its least upper bound. Evidence from Extracts 6.4, 6.6 and 6.7 illustrates that the novices quite commonly ignore the second condition. In Extracts 6.6 some intrinsic characteristics of the concept of supremum have been identified (semantic and conceptual) which, along with a general tendency of the novices to be careless about checking out all the prerequisite conditions for an implication, seem to offer a possible interpretation for C4, Connie's reduced conception of what is involved in proving that a number is the supremum of a set.

Conclusion: In the above, discussions of a problem on sup$S \cap T$ and sup$S \cup T$ gave rise to some idiosyncratic perceptions related to the definition of $S \cap T$ and $S \cup T$ (the possibly unsettling effect of the bi-lateral form of the definition of $S \cup T$) and the definition of supremum (ignoring the second condition of the definition). At a metamathematical level the varying degree of conviction that different counterexamples seem to carry was related to the deeply subjective character of mathematical cognition. Finally the uncertainty generated by one of the students' presentation highlighted the necessity for the novices to acquire a clear and minimally-ambiguous semantic and linguistic command of their writing or presenting style.

Section (vii) The Overwhelming Linguistic and Conceptual Complexity of the Notions of Sup and Inf

Context: See Extract 6.7

Structure: In the following, the tutor and Cornelia complete the proof for CD2.6i that Cornelia could not finish. The discussion is characterised by very closed questioning.

The Episode: A Factual Account. See Extract 6.7

An Interpretive Account: The Analysis

The Tutor's Three No's to Cornelia. Cornelia has had difficulties and finally gave up proving the second property for the infimum of *kS*. During the tutor's closed questioning presentation of the proof, her three responses are successively met with his disapproval. The tutor's three No's to Cornelia signify three misunderstandings, on her part, that presumably have constituted part of the stumbling block that led to her giving up on the proof in the first place.

- First (C1), Cornelia ignores that k has been given as negative and, when dividing $b>ks$ by k, replies $b/k>s$, instead of $b/k<s$. This is a typical algebraic mistake, that at university level is usually attributed to carelessness, which Cornelia instantly corrects once her attention is drawn by the tutor to the sign of k.

- Secondly (C3), when asked what is s (that in the proof denotes sup$S$) she starts her response with 'the greatest'. She is interrupted by an impatient tutor who reminds her that s is the least upper bound for S. As in Extract 6.4 most students appear in a linguistic unease (in parallel with their difficulty with the new concept) with the terminology commonly used for supremum and infimum, that is least upper and greatest lower bound. Similarly to Kelle who at least three times in Extract 6.4 corrects himself interchanging least with greatest and vice versa, Cornelia, if given the chance, might as well have done so. Coming to terms with a new concept and its complex terminology is a demanding task that novices frequently find hard to carry out. I also note here that the alternation of the terms 'greatest' and 'least' in CD2.6 is even more cognitively demanding because of the reversal due to the negative sign of k.

- The two incidents cited above psychologically build up to Cornelia's resignation, signified by C4 and to her subsequent third and essential flawed response to the tutor's closed questions. When asked about what can be said about *b/k*, a number shown so far to be smaller than the supremum of the set, Cornelia deduces that it 'must be in the set'. In her response there is evidence of a conception regarding the supremum of a set according to which anything less than the supremum is necessarily contained in the set. This is true for intervals like (a,b) where b is the supremum and anything close to b but less than b belongs to the interval. The tutor notes that the set in their proof can be a 'dotty' one, breaking up thus her *continuous and dense* concept image.

Also her response is an illustration of a major difficulty of mathematical reasoning encountered by novices: confronted with an inequality such as $b/k<s$, the available next-steps (observations that will lead to the next step of the proof) are numerous. Whether the decision made by the student is a fruitful one relies on how well she has realised the existence of the various options and on how she will associate the data with the goal of the proof. In this case, Cornelia did not seem to have adequate control of her options even though her response seems to indicate that she follows the tutor's train of thought. In this sense, the tutor's suggestion 'following your nose through the definitions' is a recommendation the meaning of which is not as obvious to the novice as to the expert mathematician. The tutor's metaphor however is insightful in another sense: it captures the detective skills required in most mathematical activities, namely that given the clues one is able to make a choice that will turn out to be effective. Cornelia's response in Extract 6.7 exemplifies the cognitive behaviour of a novice who has not mastered these skills yet.

A Note on an Ostensibly Happy-Ending Story. Cornelia, after a succession of flawed responses that have been dismissed and modified by the tutor, completes his last sentence correctly, leaving thus a last impression of successful fulfilment of the tutor's task (which is to help Cornelia understand the proof for CD2.6 i). It is a methodological constraint of this study that no further evidence is provided with regard to whether Cornelia's last utterance, $ak<b$, is merely a correct calculation, an automatic algebraic reaction to the tutor's $a>b/k$; or whether it is a meaningful contribution to his proof and one which shows that her flawed initial conceptualisation of the proof has been reformed. As an observer, my impression is that, through the end of the Episode, Cornelia remains overwhelmed by the complexity of the definitions of supremum and infimum and uncertain of her own understanding.

Conclusion: In the above, evidence was given of difficulties embedded in the notion of supremum and infimum of a set, and in particular of the notion of infimum as the *greatest* lower bound of a set. From algebraic mistakes in the handling of inequalities to the confused use of the new terminology, additionally perplexed by the alternation of the terms 'greatest', 'least', 'upper' and 'lower', these concepts emerge as essentially problematic. A particular conception revealed in the Episode was the student's belief that a number smaller than the supremum of a set, must necessarily be in the set. At the metamathematical level, the novice seemed to be in hardship with confronting the multiplicity of options in the course of a proof and with co-ordinating a variety of information in order to pick an effective option.

Section (viii) The Difficulty of Realising and Justifying the Steps in a Proof and an Application of the *Archimedean Property*

Context: See Extract 6.8

Structure: In the following, Jack's proof is modified by the tutor. The tutor then presents a proof for CD3.3ii.

The Episode: A Factual Account. See Extract 6.8

An Interpretive Account: The Analysis

CD3.3 is a problem sheet question mostly relevant to Accumulation Points but its first two parts are still embedded in the context of Foundational Analysis. Therefore, in a sense, CD3.3 illuminates the rites of passage from a study of the foundations of Analysis to a more topological territory, such as the introduction of the concept of Accumulation Point. For this reason it provides a demonstration of how the learning behaviours of the novices are standardly carried on from fundamental Analysis to Topology. A significant part of this behaviour is the students' lack of clarity and precision when presenting otherwise deftly grasped ideas: Jack's presentation of his proof for CD3.3i offers some evidence for that.

The Inability to Justify a Correct but Complex Step in a Proof. Jack appears hesitant in justifying $x-1<n\leq x$, which is however approved by the tutor. He eventually achieves a justification of the inequality with the tutor's two prompts. I note here that the second prompt is more leading than the first: it leads Jack to think of n as one natural number among those greater than x-1. As a result of seeing n as one of the numbers $>x-1$, Jack is perhaps forced to see what has made him choose n and to realise that actually n is the least of these numbers. With Andrew's question about why they need to consider n as the least of these numbers, Jack is further reinforced to articulate the reasons for his choice ('you can't take any natural number bigger than that. You have to take the next one along'). So while in writing he did not feel the pressure for justification, in the dialectic process with Andrew and the tutor (as in Extract 6.1) he is forced to abandon his tendency for tacit-ness.

As discussed in earlier analysis, these tutorials seem to be considered by the tutor as an enculturation process for the novices into mathematical literacy and articulate expression. So his objections to Jack's proof and his demands for clarification are possibly more of a pedagogical character than strictly mathematical. In the same vein Jack's redundant introduction of ε, denoting the already given in the question by frac$x$, is also criticised by the tutor.

The tutor also seems to think of himself as a mediator between the complex presentation of the mathematical content in the problem sheets and the novices. Again as previously observed in this chapter, one of the sources of the novices' difficulties with Foundational Analysis seems to be the problematic role of quantifiers. As a result, and in view of the heavily phrased proposition in CD3.3ii, the tutor suggests a rephrasing. I note however that despite the students' difficulty with the heavily quantified expression, Jack sees through the statement that 'the whole thing is about number bases for natural numbers' and has applied Mathematical Induction correctly to prove it. So, in a sense, what the tutor assumed would be a difficult question has been adequately handled by at least one student, Jack. This raises the didactical issue of the accuracy of the tutors' prophecies. I note however that his general observation that quantifiers are mathematical tools of particular difficulty for the novices resonates with the findings of this study.

Incidentally I also note Jack's remark on the tutor's suggestion for a 'quicker way' than Jack's Mathematical Induction in CD3.3ii. Jack asks: 'Do you go straight to the general case?'. What Jack seems to be thinking is that the Base Clause of Mathematical Induction and the Inductive Hypothesis are perfunctory and that there may be ways that the crucial 'general case' of Mathematical Induction would stand on its own even in the absence of the previous two steps. This and other conceptions of Mathematical Induction in the context of Foundational Analysis will be discussed globally in Part III.

Conclusion: In the above, and in the context of a problem sheet question related to an application of the *Archimedean Property*, the discussion gives rise to a student's difficulty to articulate adequately a justification of some steps in his proof. Gradually he is led to the realisation of the choices he had made in a dialectic process with his tutor and fellow student. Moreover some evidence was provided of a student's conception of Mathematical Induction according to which the first two steps of the inductive proof may be seen as perfunctory.

PART III A Synthesis of the Findings in the Area of Foundational Analysis. Indications for the Cross-Topical Synthesis in Chapter 10

In this chapter the students' first experiences of mathematical formalism were explored in a series of Episodes from the first weeks of observation. As noted in the Interlude, the concept that emerged as *paradigmatically problematic* in these first weeks was the notion of *supremum/infimum* [1, 4, 6, 7][0]. Below I list a number of the novices' problematic perceptions mentioned in the Episodes:

- a set contains its inf,
- ignoring the second condition of the definition of supremum (*Approximation Lemma*),
- perplexity with the alternation of the terms 'greatest', 'least', 'upper' and 'lower' in the definitions of sup and inf,
- a number smaller than the supremum of a set, must necessarily be in the set.

Given the epistemological relevance of sup and inf to the notion of limit, these difficulties can be also seen as a prelude to their difficulties with limits and continuity explored in Chapter 7, whether this limit is a finite number or infinity [5]. Also, since the notions of sup and inf refer to sets, inevitably the students' set-theoretical perceptions were implicit in their finding suprema and infima as well as of $\cup$ and $\cap$ as operators between sets [6][1]. In the same context, and because a large part of the process of finding suprema and infima involves the manipulation of algebraic inequalities (and in general the arithmetical handling of algebraic quantities), the students' concept image of sup and inf, as well as their action related to these concepts, was also influenced by their algebraic skills as bequeathed by school mathematics. Finally in terms of the students' handling of the new mathematical formalism and logic, quantified logical propositions proved major obstacles; also their handling of the newly formalised proving technique of Mathematical Induction also revealed some of their attitudes towards mathematical proof.

In particular the notions of sup and inf were discussed in problems related to finding the sup and inf of various given sets [4, 6, 7], or to the *Archimedean Property* [1, 8] the proof of which involves the *Completeness Axiom* which in turn contains the notion of supremum. Formulating and proving the *Archimedean Property* revealed the

[0] The numbers in the brackets refer to the Episodes in this Chapter.

[1] where the possibly unsettling effect of the bi-lateral form of the definitions of $S\cap T$ and $S\cup T$ were discussed.

students' difficulty in the encounter with mathematical formalism both in terms of the *reasoning* used as well as the *new forms of expression* [also in 6]:

- the former (reasoning) was revealed in their weakness to co-ordinate a negated quantified logical proposition with the definition of sup in order to provide a proof by contradiction of the *Archimedean Property* [1, 5ii]; or in their difficulty [7] when confronted with the multiplicity of options in the course of a proof and with the need to co-ordinate a variety of information in order to pick an effective option; also [8] in their difficulty to articulate adequately a justification of some steps in a proof.

- the latter (semantics) was revealed in attempts to mimic the b/b writing technique of the lectures by avoiding ordinary language and introducing quantifiers and set-theoretical language [1].

The novices are not at ease with the assumptions they are allowed to make when engaged in proving fundamental statements [2]: that is to distinguish, formulate and prove a universally quantified statement, even though they demonstrate an adequate initial, intuitive grasp of the proof. Specifically, the students do not appear ready to choose arbitrary numbers, establishing thus the universality of their proof, and maintain this arbitrariness through the proof with consistency. So, in a sense, questions, such as CD2.1 in [2] in which it is not clarified to the students, for instance, what statements regarding the real numbers can be assumed, emerge as problematic. As seen [2] in the students' over- and under- reaction to the requirements for rigour, students seem to be vulnerable to issues related to assumed knowledge: in other words they have been sensitised to the increased requirements of rigour in the new course but then abandoned to clarify these requirements on their own.

This sensitisation reflects to a large extent that the novices are in need of, and engaged in, a cognitive shift from unrigorous to rigorous forms of mathematical thinking. Some evidence:

- their highly metamathematical discussions of
  - the meaning of $\leq$ in the triangle inequality (as a sign that is used in the triangle inequality not to denote inequality, but to denote a variety of options for equality) [3] and
  - the Base Clause of Mathematical Induction [3] (the debate of whether the Base Clause should be stated for $n=1$ or $n=2$)

- their ambivalence on what knowledge can be assumed and how assuming knowledge is compatible with the demands for axiomatic rigour made by the lecturers and tutors in the beginning of the course [5].

The juxtaposition of the above with the epistemologically founded responses of the tutor, highlight a strong characteristic of mathematical expertise that for the time the students are missing: how mathematical experience empowers hindsight, reinforces a more fruitful use of intuition and secures the embeddedness of mathematical knowledge. Incidentally I mention here the slightly more alarming evidence[8] of a perception of the Base Clause of Mathematical Induction and the n=k step as perfunctory (only the k+1 step of the proof was deemed important by the student).

Interestingly this juxtaposition between expert and novice approaches is not as clear-cut as possibly expected [6]: different — but mathematically equivalent — counterexamples seemed to carry different degrees of conviction both for novices and the expert. This was discussed in terms of the deeply subjective and vulnerable character of mathematical cognition.

Sensitisation to the requirements for rigour does not always imply that the novices are willing to attempt formalisation [1, 5ii]. On the contrary [4, 5i], in most cases difficulty in formalising leads to reluctance, avoidance and preference for concrete, intuitive arguments. Some evidence [4]: a student abstaining from using the *Completeness Axiom* from a proof because he knew the *Axiom* in terms of suprema and he did not realise that a symmetrical statement holds in terms of infima; or [5i] formalisation is not thought of as necessary when a proposition is perceived as obviously true. So, while some students [5] have conceptualised the necessity to be formal and struggle with the materialisation of this conceptualisation, others are still engaged in the vicious circle of assuming in their proofs what is to them intuitively obvious or what they are actually being asked to prove.

Chapter 7

The Novices' Encounter With Mathematical Abstraction: Cases from *Calculus*

PART I A Guide to the Paradigmatical Cases (Episodes) Presented in this Chapter

The following table contains contextual information with regard to the 8 Episodes presented in this Chapter.

Episode Number	Time of Incident Term - Week	Participants	Mathematical Content
1	Michaelmas 5	Cathy and George	CD4.1, limit and continuity
2	Michaelmas 6	Kelle	CD5.1, limit and continuity
3	Michaelmas 8	Jack and Andrew	CD7.1&2, limit, continuity, derivative. Double limits. The algebra of limits.
4	Hilary 2	Camille and Frances	The Fourier Series of a function
5	Hilary 4	Camille and Eleanor	B6, differentiation
6	Hilary 6	Abidul and Frances	B7, Intermediate Value Theorem
7	Hilary 7	Cary and Beth	B10, Taylor expansion of a function
8	Hilary 8	Cathy and Cliff	SS7.1, convergence of series

PART II: Data and Analysis

In the following I present the factual and interpretive accounts and conclusions for the 8 Episodes of the table in the previous page. In Part III then I synthesise the findings of Part II related to Calculus and briefly discuss the wider cognitive issues that are presented in the overall synthesis of the data analysis in Chapter 10.

Section (i) Constructing a Meaning of the Concept of Limit: Concept Definition and the Formalism of Mathematical Notation, Concept Image and Visualisation

Context: See Extract 7.1

Structure: In this Episode
- the students attempt a reconstruction of the formal definition of limit,
- the tutor and the students explore the students' intuitive conceptions of limit,
- the tutor reconstructs the formal definition of limit,
- the tutor and the students discuss the meaning of the formal definition of limit,
- the tutor and the students discuss the mechanism of the definition through examples
  - $f(x)=x$
  - CD4.1i
  - CD4.1ii
  - CD4.1v
- and, finally, they discuss the process of guessing limits, the A-level approach to graphing functions and the algebra of limits.

The Episode: A Factual Account. See Extract 7.1

An Interpretive Account: The Analysis

The Students' Inadequate 'Intuitive Concepts' of Limit. I note that the discussion between the tutor and the students on their conceptions of limit takes place soon after the students have been introduced to the concept and have worked on problem sheet CD4. Therefore this Episode captures in freshness the genesis of their conceptions.

Cathy's hesitation in the beginning signifies quite eloquently the paralysing effect that the request for formalism has on novices at this stage. George's reaction is to fully reproduce in (1) and (2) the definition given in the lectures giving however the impression that he does so in a mechanistic, uncritical way: he completely ignores the conditions under which $|f(x)-l|<\varepsilon$ holds in (1). Then, as a result of the tutor's objection, he restores the conditions impeccably in (2). That the formalistic integrity of G1 and G2 do not reflect the depth of his understanding, is illustrated in G1 which, on one hand, contains the main implication of the definition ($x \to a \Rightarrow f(x) \to l$) and, on the other hand, shows a very problematic perception of the nature of ε. As it will be repeatedly illustrated in this Chapter translating the $\to$ of the above implication into the quantification of δ and ε is probably the most cognitively problematic aspect of the definition of limit observed in these students.

Moreover C1-C3 are indications of Cathy's semantic misinterpretation of the notation used by the lecturer. In order to denote the dependence of δ on a and ε, the lecturer used $\delta(a,\varepsilon)$ which Cathy mistook, first, for δ being a function of a and ε (C1) and, then, for an interval (a,ε) of which δ is an element. In sum, so far Cathy appears to be completely at sea with the concept. The subsequent shift of Cathy's perception of the form and content of the definition of limit seems remarkable and is presented in the following along with further elaboration on the concept. The conversation starts immediately after the tutor has explained the definition of limit (first verbally, in colloquial terms such as 'approaches' and 'close', then formally in (3), introducing δ and ε, and graphically). While listening to him, Cathy sounds preoccupied with the definition of δ (C4).

In C4 Cathy appears as if she is trying to understand the definition of limit from its negation without realising that her words constitute parts of the negation. C5 is a verbalisation of the definition that seems to reflect a growing image of the limiting process as a machine (input: ε, output: δ). George also accentuates this with 'find a δ' in G2 as opposed to 'there would be a δ' in C5. C6 is more illuminating about Cathy's intentions: she seems to be struggling with the implication

if the limit exists then $\forall\varepsilon>0\ \exists\delta>0...$

and wants to find out if it is likely at all for the limit to exist and at the same time to be unable to find a δ. The tutor (T5) still thinks she is trying to construct the negation but she seems surprised at this interpretation (C7). Her surprise in C7 might also be attributed to the common belief that definitions are undeniable (by 'denying' the tutor possibly refers to negating the categorical proposition contained in the definition). Then in C8 it seems that, for the time, Cathy has resolved her perplexity at the definition of limit. Her 'close enough' verbalisation in C8 is an articulate one and is enthusiastically received by the tutor. However Cathy's enlightenment proves very short-lived.

Led by the tutor through applying the definition of limit on $f(x)=x$, Cathy appears capable of finding a δ on her own (I note that whereas ε is given by the tutor as equal to 10^{-472} she replies 'take $\delta=\varepsilon$' which shows that she has generalised the process of finding δ and doesn't need the specificity of 10^{-472}). Moreover both students appear to be beginning to comprehend the mechanism of the definition as seen in their observations on δ and its dependence on ε and the function.

Cathy's 'Prejudice' for Continuous Functions. C9 is an illustration of a very common 'prejudice' for continuous functions (see Chapter 1, IIIc.ii). Cathy dismisses the notion of taking a limit as redundant because, at this stage, in her concept image of function, all functions are continuous; therefore, when $x \to a$, the values of f tend to $f(a)$. The tutor's example and CD4.1i illustrate two types of discontinuity that render the process of limiting necessary ($\lim_{x\to a} f(x)$ exists but is $\neq f(a)$; $\lim_{x\to a+} f(x) \neq \lim_{x\to a-} f(x)$ respectively).

Subsequently the students remain silent even though at first Cathy simplifies the definition of CD4.1i and the tutor gives them a graph of the function in CD4.1i. I note here, even though the only evidence is the students' silence, that this shift from participation to passivity might be attributed to their unfamiliarity with piece-wise functions and to the tutor's use of the theorem that $\lim_{x\to a} f(x)$ exists if and only if $\lim_{x\to a+} f(x) = \lim_{x\to a-} f(x)$). Despite his tacit use of the theorem, the tutor then changes again to using the definition for a 'proper proof': to prove that the limit does not exist one has to show that, $\forall L \in \Re$, $L \neq \lim_{x\to a} f(x)$, i.e.,

$$\text{if } L \in \Re : \exists \varepsilon > 0 \text{ such that: } \forall \delta > 0,\ 0 < |x-a| < \delta \Rightarrow |f(x) - L| > \varepsilon$$

Cathy's 'in terms of epsilon' shows puzzlement and her 'does L have to be -1 and 1?' sounds reminiscent of her persistent image of all functions being continuous. She also sounds as if she has missed the argument that the definition must be negated $\forall L \in \Re$ and thus she is struggling with taking specific values of L. George (G3) in general sounds more at ease with the argument: he can see that for $L>1$, L then obviously cannot be the limit because there are no values of the function at all around it and he also contributes with a better offer for ε.

As the tutor comments immediately after G3, the cognitive puzzlement recorded above is not unanticipated given that the novices are hardly familiar with either the complicated nature of multi-quantified propositions or with the mechanisms of negating such propositions.

C10 and C11 are peculiar statements in which the student appears to have despatched her own statement 'as x gets close enough to 0, $f(x)$ gets close enough to 0' from the definition of limit. With the tutor's reassurance she restores the connection but, even though she declares her intention to find a δ, it is George who suggests $\delta=\varepsilon$.

In CD4.1v the tutor is unpleasantly surprised by the students' inability to picture $x\sin 1/x$ even though George's observation that $-1\leq \sin x\leq 1$ is a good start. Only when given the full picture, Cathy strikes a good guess for the limit of $x\sin 1/x$ and a fair, if not successful, attempt at choosing δ (even though her δ is not the optimal choice, the mechanisms of her choice appear to have improved remarkably: she seems to be solving $x\sin 1/x$ in terms of x and then suggesting $\delta=x$).

The inadequate mastering of some mathematical skills at school appears again in the tutor's comments upon CD4.6. What seems to underlie the students' resignation is

- for George, a weakness to proceed beyond intuitive 'right ideas' and
- for Cathy, a weakness in re-activating school mathematical knowledge, e.g. construction of graphs.

Cathy's comment on the algebra of limits as 'imprecise' might be evidence that the condemnation of school mathematics as unrigorous (which imbues introductory university courses: in these tutorials the students have been repeatedly asked to 'wipe out' their previous knowledge because they have to re-establish all their mathematics axiomatically) has been uncritically interpreted. What the lecturers and the tutors possibly mean by 'imprecise' is the use of the theorems from the algebra of limits without having proved them or without stating their use clearly. The algebra of limits per se is not 'imprecise'. Cathy's words and resignation underline the conflict between school and university mathematical knowledge. As long as the conflict remains unresolved, the novices sometimes withdraw altogether from using their school mathematical knowledge. At the same time this knowledge is taken for granted by tutors or examiners when it is actually inert.

In the same vein the fact that, according to the tutor, the students have not been using inequalities flexibly in their exploration of limits, is an illustration of the inertia to which school mathematical knowledge seems sometimes to have been condemned by the novices. I note here that a great deal of Analysis relies on simple rules such as replacing some quantities with more manageable smaller or greater ones. Some of these inequalities are taught at school. It seems that an imperative need for mathematics teaching at university level is to devise ways with which the conflict outlined above will be more effectively resolved.

A Note on the Tutor's Semantic Interventions in the Students' Writing. The tutor's two interventions in the students' writing ((1), (2), (3), (4) in Extract 7.1) highlight the need to foster in the novices a flair for a creative use of notational as well as lexical language: he replaces the ambiguous $\delta(a, \varepsilon)$ with $0<|x-a|<\delta$ in George's definition of limit and he adds the words 'suppose' and 'then' in Cathy's proof for CD4.1ii. Especially in the latter case his interventions accentuate the implication contained in the definition of limit and can be seen as part of the tutor's encouragement of his students to use language in order to punctuate and clarify mathematical statements.

Conclusion: In the above, the newly introduced formal definition of the concept of *limit* has sparked off a multi-layered discussion which revealed the students' difficulties with the δ–ε mathematical formalism and with assigning meaning to the formal definition; their prejudice for continuous functions and by implication their view of limit as a redundant concept. Moreover their difficulty with finding limits either via the definition or via the algebra of limits was largely attributed to their growing mistrust towards school mathematical knowledge. Using inequalities in order to manipulate quantities, graphing functions, guessing limits and using the algebra of limits are mathematical practices that the students questioned as to their rigour and, hence, as to their acceptability. From a didactical point of view, the need arose to re-establish the legitimacy of school mathematical practices as a way to gain mathematical insight and to introduce formal mathematical language as a way to refine and establish rigorously these insights.

Section (ii) The Novel Notion of Continuity: Proof By Definition Or With the Algebra of Limits. A Battle of Ambivalent Preferences and the Cognitive Effect of a Hidden Agenda

Context:	See Extract 7.2
Structure:	In the following, the tutor leads Kelle through a proof for CD5.1i by, first, graphing intx and, then, by definition. Subsequently, for the rest of the question, she changes her strategy by suggesting using the algebra of limits. Kelle's unease with this change — as well as several difficulties with the novel notion of continuity — then come to the surface.
The Episode:	A Factual Account. See Extract 7.2

An Interpretive Account: The Analysis

Throughout the Episode the tutor seems to recommend quite uniformly a strategy for tackling problems with limits:

> make a drawing
> find out about the limit
> prove the limit formally.

In these tutorials this recommendation is very common. The students usually receive it with hesitation: in a variety of other cases they are asked not to rely on pictures because pictures are inaccurate. Also guessing a limit is a rather vague, non-specific suggestion. As a result this type of recommendation seems to be adding a mystifying veil to the already obscure notion of limit. In this Episode however the use of pictures is not as problematic as the tutor's ambiguous and unjustified alternation between different approaches to the third step of the above procedure: proving the limit. I note that *ambiguous and unjustified* here are not used as criticisms of the tutor's mathematical handling of CD5.1: her alternation between different approaches springs from her long mathematical experience according to which it is sometimes convenient to determine the limit by using the definition, or side limits, or the algebra of limits. But she does not justify this alternation of approaches to Kelle from whose perspective this ambiguity maybe interpreted as the tutor contradicting herself.

Kelle in the beginning of the Episode complies with the recommendation to look at the drawing and notices the discontinuity at the integers. The tutor accepts the observation and transforms it into a specific question:

look for $\lim_{x \to x_0} f(x)$ when x_0 is or is not an integer.

For the latter case, Kelle switches to $\delta - \varepsilon$ mode and tries in vain (K1) to reproduce the formal definition of limit. With the tutor's prompting he realises that he first has to define $f(x)$, then $f(x_0)$ and then look for the limit. Looking at fig.2a makes him lose generality and reply on the basis of looking at the interval (0,1) instead of (n, n+1) twice. K2 then reflects Kelle's preoccupation with whether

Proving that f is discontinuous at the integers

is equivalent to

Proving that f is continuous except at the integers,

which is not elaborated upon by the tutor who proceeds with proving the other case for the limit. Consistent with her declaration of strategy she insists on using the definition (as opposed to Kelle's suggestion for side-limits, probably evoked by fig.2a). Her reason is a preference for avoiding 'quoting big theorems' and again she recommends looking at the drawing.

Kelle complies and contributes to the construction of the drawing for CD5.1ii. Maintaining her control, the tutor outlines the available strategies: formal definition or the algebra of limits. Kelle, influenced by his tutor's expressed preference so far, suggests using the definition but then she announces they will use contradiction and the algebra of limits. Kelle complies again and tries to construct an argument by contradiction but, as indicated earlier by K1, he is in trouble with articulating formal propositions and he fails (he vaguely starts by not negating the proposition). The tutor corrects his faltering logic and states the argument for the contradiction. He adopts the argument and, still influenced by his tutor's expressed preference for the formal definition of limit, suggests its use. The tutor reminds him of her newly expressed preference for the algebra of limits and takes over.

K3-K5 is evidence that Kelle, not confidently and far from independently or efficiently, begins to be able to reproduce the tutor's strategy as exhibited in CD5.1i-iv.

The Essential Role of 'Pictures'. The tutor repeatedly invites Kelle to use 'pictures' and 'intuition' in order to understand the limiting process. Like other tutors she encourages the novices to engage in a 'guessing' intuitive activity before proceeding with formal reasoning. It is very likely that Kelle's inability to start CD5.1 lies in his not using this recommendation which is the only technical (but ambiguous) advice he has received with regard to the exploration of limits. It is remarkable that, once given the 'picture' of int*x*, Kelle immediately observes its points of continuity and discontinuity. I note here that his two mistakes (that $f(x)=0$ and that x lies between 0 and 1) can be attributed to the fact that he was looking at the specific interval (0,1) and not at the general $(n, n+1)$ suggested by the tutor. This underlines I think the novice's internal fixedness to specificity and it addresses the tension between the tutor's pressure for statements of generality and the student's literal interpretation of the information contained in the graph. It also raises the issue of the role that powerful but constraining representations exert on thinking.

The Problematic Handling of Formalism: Definitions and Logic. In this Chapter the students often try to reproduce the formal definitions of limits and continuity they have been given in the lectures. K1 is a very typical example of a novice's attempt to verbalise a formal statement which is disconnected from his concept image and is in striking contrast with Kelle's immediate grasp of the limits involved in CD5.1i while observing the graph of int*x*. So in K1

- he initially omits the universality of the definition of limit $\forall\varepsilon$,
- he confuses the necessity of finding a δ with finding an ε,

but

- he reproduces $f(x)-l<\varepsilon$ recalling then that the inequality must hold $\forall\varepsilon$.

By recalling the $\forall\varepsilon$ at the final stage, he misses the link between δ and ε, namely that introducing δ and ε is a way of quantifying the limiting process: as x approaches x_0, $f(x)$ approaches $f(x_0)$.

Similar difficulty with a formal expression is exhibited later when Kelle tries to formulate the argument for the contradiction in CD5.1ii. However the student demonstrates a remarkable caution for detail when he wonders whether proving that f is discontinuous at the integers is equivalent to proving that f is continuous except at the integers: this is a quite fine linguistic distinction to be made by a novice who has just seemed to be incapable of providing an adequate formal definition of continuity. Assuming that making this linguistic distinction reflects a clarity of his global vision of the proof's structure then this clarity is in contrast with his

struggling and rather failing attempts to use the $\delta-\varepsilon$ definition of limit, and, his inability to negate a formal proposition in order to construct a proof by contradiction.

This contrast, as evidence in this Chapter repeatedly suggests, is characteristic of Kelle's novice behaviour.

First Principles Versus Theorem Quoting: a Battle of Ambivalent Preferences. In the Episode the tutor is seen as altering her preferred strategy for finding the limits in CD5.1. What the tutor may tacitly aspires to convey is the need for flexibility to adapt to a mathematical question's specific needs and to master the use of multiple approaches. What I think Kelle has received though is very ambiguous instructions in which the rules of how to choose an approach (First Principles/Definitions or Theorem Quoting) remain unarticulated. Therefore it is likely that the process of alternating among the various approaches remains mystified for Kelle. There is no actual contradiction in preferring FP to TQ in one case and TQ to FP in another; it is an entitlement and a virtue of the mathematician to master the skill of optimal preference. But it is a didactical action that induces a rather uncreative uncertainty and confusion. Kelle in other words is cognitively a victim of his tutor's hidden agenda.

Changing the strategy to using the algebra of limits and switching to the use of proof by contradiction seems to be a more powerful shift than Kelle can handle and the discussion degenerates into exposition by the tutor for both CD5.1ii and iv. Interaction is re-established for CD5.1iii but despite Kelle's participation in K3-K5 due to the tutor's exercise of firm control and inclination to exposition it is very difficult to evaluate the learning outcome and confirm that Kelle would have been able to carry further his observation that int$\sin x$ is discontinuous at the points where $\sin x = n$ and transform it into a proof.

Conclusion: In the above, the novice's thinking with regard to the newly introduced concept of continuity and the formal definition of limit seemed to be influenced by the following three elements:

- the role of visualisation: the use of pictures in guessing limits; pictures as specific representations potentially opposing generality,
- the role of formal notation and logic as a vehicle for mathematical arguments: the difficulty of evoking the formal definition of limit; the difficulty of constructing a logical statement of an argument for a proof by contradiction;

preoccupation with the use of the term *not-continuous* as equivalent to *discontinuous*.,

- the juxtaposition of proof by theorem quoting and proof by first principles: the effect of a not overtly justified alternation using the definition of limit and using the algebra of limits.

Didactically, the dialectics of very closed questioning and of constant reclaiming of control on the part of the tutor have proved unhelpful in supporting the student's conceptual development with regard to the concept of limit. Also, methodologically, the tutor's style has been interfering with the researcher's attempt to understand the cognitive state of the student's thinking.

Section (iii) LimΣ, Σlim and the Right to Exchange Limits. The Superiority of Proof Via First Principles and the Convention of Foundational Rigour

Context: See Extract 7.3

Structure: In the following, the tutor refutes the students' choice of strategy in CD7.1 on historicoepistemological grounds and addresses the question of principles involved in assuming statements that have not previously been proved in the course. The students have correctly used a theorem which, however, will not be proved until next term. They discuss the limited validity of the theorem and whether they are entitled to use it.

The Episode: A Factual Account. See Extract 7.3

An Interpretive Account: The Analysis

The Student's Foundational Preoccupations and the Tutor's Preference for Proofs by First Principles. It is intriguing that despite the imperfect inductive proofs both students presented in CD7.1, they seem to have been theoretically alerted to the dangers of 'assuming any of the foundations', 'going sort of backwards' (Andrew) and about 'the idea of assuming limits' (Jack is not happy about that). The tutor elaborates upon these preoccupations by demonstrating a proof for CD7.1 which is clearly based on first principles. In fact the proof via the binomial theorem is an almost arithmetical one and it employs mostly school mathematical knowledge. Thus at a metacognitive level the tutor's preference of the proof-by-binomial-theorem over the proof-by-induction signifies a tentative, much needed bridging of the gap between elementary and advanced mathematical thinking. However I note here that, in the tutor's words, his preference for the proof via the binomial theorem is grounded on historicoepistemological, and rather than didactical reasons.

The tutor's preference for proof by first principles in CD7.1 may look as if it is largely a matter of aesthetics. It is the discussion of CD7.2 which makes it more obvious that he is seriously concerned with the implications of the students' practice of assuming unproved knowledge. There is evidence in this Episode that within the course lie conflicts with regard to the knowledge to be assumed and with regard to the varying demands for rigour. The students 'instinctively' use a theorem on the exchange of double limits which holds only under some conditions, without either stating clearly the theorem, or proving it. Formally speaking this approach is unacceptable in a foundational Continuity-Differentiability course in which Analysis is built upon a few, selected first principles. At the same time, in the course on Analytical-Numerical Methods and Differential Equations, the technique of unjustifiedly exchanging limits in order to simplify calculations is very common and one that the students are encouraged to employ. The borderline of the above distinction is for the novices rather blurred and ambiguity often results in treating questions, such as CD7.2, unrigorously. The tutor's statement 'It's the job of this course [Cont-Diff] to question' the conditions of the theorem and his promise that the theorem will be 'reinstated next term with all the full rigour that mathematics is capable of' is a tentative clarification that tries to resolve the ambiguity. The tutor here tries to draw clear boundaries between the different domains of mathematical activity in the course and specify the rules of the game of rigour within each one of these domains. In other words he tries to articulate, more explicitly than most of the other tutors participating in this study, the terms of the didactical contract by stating what is expected of the students in the various domains of their mathematical activity.

The Role of Conventions: the Right to Use the Limited Validity of a Theorem and the Right to Make Unjustified Statements. The students' unconditional exchange of double limits provides evidence for another metamathematical observation with regard to their attitude towards counterexamples: at first both students are convinced that exchanging double limits is 'not necessarily' acceptable. To formalise this conviction the tutor suggests constructing a counterexample. Jack immediately responds with a not very successful but reasonable attempt. Andrew sounds more confused with the act of looking for a counterexample than the example itself. Finally both students appear sceptical when presented with one:

- Andrew, who in previous Episodes has been uncomfortable with the role of a counterexample as disproving totally the truth of a proposition, wonders whether the tutor's argument is 'useful': it ruins their proof; also, in the particular case of CD7.2, the limits can be exchanged.

- Jack's explanation on why he used an unproved theorem is even more unconventional: 'you can not-justify yourself on paper as long as you can justify yourself in a tutorial'. What Jack does not realise is the basic convention of an introductory course that, on the b/b, on paper or in a tutorial, arguments are expected to be self-contained and not rely on tacit explanations. Jack's comment actually reflects the common practice in published mathematical research according to which theorems, that appear in journals for instance, contain statements, the proof of which is merely hinted at, or briefly outlined, by the author. I note however that it is questionable whether Jack is knowledgeable enough to adopt this practice successfully.

In both cases the students doubt the value of the tutor's effort, perhaps because they see it as a bureaucratic hindrance to the unfolding of their argument: whether the theorem has been proved in the course or not, it can actually be proved and that makes their proof for CD7.2 valid. The tutor realises the undermining power of the students' thoughts and presents them with a less questionable proof, disentangled from matters of principle and from conventions.

Conclusion: In the above, evidence was given of the tension between novice and expert practices. In particular:

- the expert appeared capable of distinguishing between different approaches to a proof relevant to the notion of derivative and of expressing a preference for a proof via first principles (via using the binomial theorem), as opposed to a proof by mathematical induction, on historicoepistemological grounds.
- the novices appeared alerted to the dangers of not clarifying the foundations of their proofs but in practice they turned out to be confused and weak in doing so. Especially one student expressed concern with 'assuming limits'.
- the novices generalised the finite case of $\lim(a+b+c)=\lim a+\lim b+\lim c$ into $\lim\Sigma=\Sigma\lim$ which is actually a double limit, thus disregarding unconsciously the conditions for the exchange of limits.
- the novices acted with suspicion and hesitation towards the tutor's request to consider the limited validity of the theorem via a counterexample:
    - one of the students objected to considering the cases where the theorem does not hold since it holds in the particular case of the Episode (CD7.2)
    - the other student negotiated his right to use elliptic arguments (a common professional practice).
- the expert appears capable of distinguishing between expected practices in different domains: for instance, in applied mathematics, retrospective use of unproved results is allowed as opposed to foundational courses where it is not.

The conflict can be seen as an enculturation process during which the novices learn about certain conditions of the didactical and epistemological contract in common mathematical activity: in the interaction it is revealed that they have been unconsciously assuming the validity of exchanging double limits, using thus a theorem which is yet unproved in the course. What the novices seem to be learning here is that they need to exercise control over their mathematical reasoning in order to avoid subconscious and unjustified decisions and that this skill to control is a characteristic of appropriate mathematical behaviour.

Section (iv) Striving for Meaning and Significance: The New Concept of Fourier Series

Context: See Extract 7.4

Structure: In the following, the student almost forces a rather reluctant tutor to improvise on the subject of Fourier Series. As in other sessions with Camille, the student frequently interrupts in order to ask for clarifications. The session stands as a generic example of a novice's struggle to construct meaning and significance for a new concept. A determinant characteristic of this struggle is the intensely emotional responses of the student. In this case the struggle is possibly let down by what seem to be perceived by the student as unsatisfactory responses by the tutor.

The Episode: A Factual Account. See Extract 7.4

An Interpretive Account: The Analysis

Striving for Meaning and Significance. Camille's determination to explore the meaning of Fourier Series of a function is clearly stated in the opening statement of the Episode, C1. Subsequently the student seems to have chosen the strategy of clarifying excruciatingly every relevant term/concept/definition. This results in her frequent interruptions of the tutor who is fairly annoyed and sounds as if she is always on the verge of losing her patience. I note here that this is something not fully conveyed by the dry citation of facts in the transcript. Also I note that the tutor's irritation can be attributed to the fact that Camille has asked her about a topic that she is not entirely and readily familiar with. The reason I begin the analysis of Camille's behaviour with these comments on the tutor's behaviour is that the latter directly influences the former. Moreover Camille's reactions are hard to dissect in terms of the cognitive/affective dichotomy so at least in this Episode there is a combined consideration of both emotional and cognitive issues.

For instance, Camille's three interventions, concerning

the definition of periodic function,
the use of the terms *point-wise continuous* and *continuous* interchangeably and
the notation $f(x+)$,

can be seen at the same time as pedantic reactions AND signs of insecurity. Then comes the discussion of ~ as a sign of ambiguous meaning. Justifiably Camille, as well as other novices in various occasions, confuse the several uses of ~ that vary from course to course. Camille mentions approximation, equivalence/similarity and equality (C2-C4). All are dismissed by the tutor who prefers ~ to mean 'this is the Fourier series of this function' because this allows writing the Fourier Series of a function *f* down without worrying about whether it converges to *f* or not.

Camille again expresses mistrust and turns to Frances — her silent peer who is sitting at the same tutorial with Camille — for support with C5. Frances listlessly admits to no confusion. Camille's perplexity becomes more evident in her subsequent 'but it doesn't really converge' which the tutor does not explore. With her wondering about what is a Fourier Series still pending, Camille turns to the exercise from the lectures (fig.4b). The tutor explains the calculations that have been omitted by the lecturer and Camille's frustration now turns overtly against the lecturer towards whom she is mistrustful. She seems nevertheless to have grasped the notion of possible odd and even extensions of a function. Her next query is again a sign of mistrust and disbelief towards the lecturer: if odd and even extensions are two possibilities why did the lecturer mention that the possibilities are three? A brief examination of the lecture notes leads to the conclusion that the lecturer meant the ordinary Fourier Series as the third possibility. Satisfied for a moment Camille turns to her initial epistemological question on the Fourier Series of a function with C6.

The tutor's subsequent explanations in T5 are characterised by an evasive vagueness culminating in the final statement '...we are sort of skipping things here a bit because you are doing this clearly from the point of view of using it, so you really need to know the minimum to actually do things. Because the actual theory of it gets quite complicated'. The latter illustrates the conflict encountered earlier in the first term of the novices' studies. While in the course on Continuity and Differentiability there is a clear intention to build Analysis carefully and precisely on sound axiomatics, the Analytical and Numerical Methods course is an exercise in the mechanics of differentiation and an application of rules for solving differential equations. The tutor's statement illustrates the double standards that occasionally apply in these introductory courses. In other words, beneath the tutor's words lies an attempt to discredit Camille's inquisitive epistemological approach (highly desired in more rigorous courses like Continuity and Differentiability) as inappropriate for this particular course. This application of double standards becomes problematic since the dichotomy is poorly perceived by the students who justifiably cannot distinguish confidently among the variations of rigour expected from them at the various stages of the course.

In sum, in this Episode a novice's struggle for learning has been seen in the light of both contextual and emotional factors. In particular in this Episode the mathematical context of Analytical and Numerical Methods seems to determine a variety of aspects of Camille's learning, starting from the interpretation of a symbol (~) and up to the degree of detail and rigour required in the definition of new concepts. Camille has not been told about the particular meaning of ~ in the definition of a Fourier Series and she does not realise that all that is expected of her, at this stage, is to learn about the new concept 'from the point of view of using it'. So she engages in an excruciating decomposition of the concept which turns out to be emotionally frustrating since she doesn't share the same concern with her peer in the tutorial or her tutor. Her increasingly persistent inquiry is a result of this frustration and this is why I think that, in this Episode, the cognitive and affective are intertwined and therefore harder to distinguish and dissect.

Conclusion: In the above, the influence of emotional and contextual factors on the learning of the new concept of Fourier Series were explored within the topical area of Analytical and Numerical Methods and with regard to the concept of the Fourier Series of a function. Evidence was given of

- a novice's struggle for attributing meaning and significance to the new concept. The student attempts various interpretations of the symbol ~ in the definition of a Fourier Series (approximation, equivalence, equality) and also probes in detail several related concepts (periodic function, continuity, side limits)
- the various degrees of expected rigour in the different topical areas of the course (applicability of Analytical and Numerical Methods versus formalistic rigour of Continuity and Differentiability). These variations in rigour do not seem to be clearly perceived by the novices.
- the student's frustration with unsatisfactory explanations (from both lecturer and tutor).

In the above, cognitive and affective aspects of learning have been demonstrated as inextricably connected.

From a didactical point of view the contrast between the tutors (see for instance Extracts 7.3 and 7.4) is striking. Whereas the tutor in Extract 7.3 is quite willingly leading the process of enculturating the novices into the conventions of mathematical formalism, the tutor in this Episode responds rather reluctantly.

From the two tutors' different degrees of wilful engagement in the process it is reasonable to deduce that the first considers his interaction with the students at this reflective, meta-topical level a legitimate part of the tutorial. On the contrary, the second tutor seems to view Camille's questions as a deviation. It is a contention of this study that metamathematical reflection is a necessary component of the novices' enculturation — they often ask for it. Deterring them from metamathematical discussions in a tutorial is equivalent to leaving questions unanswered and therefore not catering for their expressed cognitive needs.

Section (v) The Contrast and the Gap Between the Mechanistic and the Conceptual Approach to the Notion of Derivative

Context: See Extract 7.5

Structure: In the following, the tutor and the students discuss the solution to a problem which the students have been working on. The tutor's questions aim mainly at clarifying the justifications for the answers given by the students. I note that despite the numerous opportunities for an exploration of the notion of derivative the tutor seems to be inclined towards a rather restricted presentation of the solution without any substantial deviations. One of the students, Camille, on the other hand, appears more willing to discuss the reasoning and conceptual background of the question but her wish remains largely unfulfilled.

The Episode: A Factual Account. See Extract 7.5

An Interpretive Account: The Analysis

The Mechanistic As Opposed To The Conceptual Approach To The Notion Of Derivative. Camille appears absent-minded in the beginning of the session: for instance she interferes in order to modify Eleanor's suggestion for B6i (from $n \geq 2$ to $n > 2$) and follows her correction up with a proof for $n \geq 1$. Moreover in B6ii, whereas Eleanor has given the answer ($n \geq 2$) before they start the proof, Camille interrupts the tutor in order to ask how they knew 'it was $n \geq 2$ straightaway'. The tutor's wry answer ('because Eleanor knows the answer and I know the answer!') overshadows the possibility that Camille's comment is not simply a sign of absent-mindedness but a genuine expression of discomfort with discovering the answer intuitively and then proving it formally. The tutor, in order to confirm that Camille is following the discussion, asks for the reason they are looking at this limit. Camille's immediate responses indicate that she is following the argument of the discussion.

That Camille has conceptual preoccupations becomes evident in C1. The tutor's response, T1, is rather unbalanced in its emphasis: the tutor does give the mathematically correct answer that the derivative might exist and not necessarily be continuous, but she also unnecessarily emphasises that, in this case, the derivative is continuous. Later she adds the necessary condition 'as long as we make sure that this exponent is at least one' but at that moment her statement sounded absolute. C2 then is a sign of the following misunderstanding:

The Tutor's Interpretation of CI	Camille's Implicit Meaning in CI
Is this particular derivative always continuous at zero?	Is the derivative of a function always continuous at zero?

The tutor's response with regard to the particular function causes Camille to repeat her question explicitly this time. Camille's responses subsequently demonstrate that she knows definitions and calculational methods but it is the tutor who highlights the next logical step in the proof for B6ii of the question: it is now necessary they prove that for $n=1$ the derivative at zero does not exist. This idea appears to be elusive to the students who, having found the right answers in B6ii and B6iii, did not complete the proof formally since they did not cover all the cases. The scene will be repeated a bit later in B6iii.

Another noteworthy occurrence in B6ii is Eleanor's claim that $\cos 1/x^2 \rightarrow \infty$ as $x \rightarrow 0$ (the limit does not exist). The tutor corrects Eleanor, and proceeds without any comment, but it would have been interesting to explore the students' fusing the perceptions of tending-to-infinity and not-tending-to-a-limit. Interestingly this conceptual puzzlement does not seem to affect the student's performance who, with ease and precision, suggests the method for the next part of B6 (calculate the derivative and then find its limit at zero). Both students calculate the derivative impeccably.

The next piece of evidence that, despite evident technical proficiency, their conceptual understanding is not at all at the same level, comes from Camille who calculates the derivative of $x^n \cos 1/x^2$ efficiently. In the end she adds '$\forall n \geq 2$'. Since B6iii follows B6ii, in which the existence of the derivative at zero has been proved $\forall n \geq 2$, Camille extends this claim to the existence of the derivative for any x. This is not the case since the derivative exists $\forall n \in N$ for $x \neq 0$. Camille's reaction is a sign that there is a gap between a conceptual approach to the notion of derivative (determining the value of the derivative point by point by calculating the limit of the ratio) and the mechanics of the algebra of limits/derivatives (here for instance $(fg)' = f'g + fg'$). It is again frustrating that the tutor does not point out the connection between the two approaches and leaves the two concept images of derivative dissociated from each other.

For not very visible reasons, Camille's response to the tutor's next question (under which condition for x, her calculation of the derivative is correct) is $x>0$, whereas the only value of x that must be excluded from the derivative of f_n is zero. However, that Camille has understood the mechanics of the question, is clear from her eloquent 'the question is to find for which n to have the limit of this derivative to be zero' and her answer and explanation to the question ($n \geq 4$). Then again, in a comment that sounds distracting from the impression of clarity and eloquence that Camille gives, she asks what would happen if $x>1$ — ignoring that the limits involved in the process are all taken when $x \to 0$. Here it is possible that Camille's concept image of the domain of f_n being $\Re$-{0} (x can be anything other than zero; therefore it can be >1) overshadows her concept image of the limiting process which requires x to be taken as close to zero as possible. Even though x can be greater than one, the values of f_n for $x>1$ are irrelevant to the question. Another instance of this possible conflict between the point-wise/static and the procedural/ time-dependent perception of the variable in a function that I encountered in these tutorials is the following: when discussing the limit of a sequence s_n ($\lim_{n\to\infty} s_n$), some students ask what the limit is if $n \to 0$. In this case wondering about the values of s_n as $n \to 0$ is not simply irrelevant, as in the case above, but is also evidence of an incomplete image of the domain of the sequence s_n ($n \in N$).

Subsequently, as commented in an earlier paragraph, the tutor reminds to the students that, despite having found that the derivative is continuous at zero for $n \geq 4$, they still have to exclude the cases for $n=2$ and $n=3$.

Then Eleanor claims that the second part of f_2 (namely $(2\sin 1/x^2)/x$) tends to infinity. It is likely that Eleanor has been applying here the argument that was used repeatedly in today's tutorial. That is: Likewise,

$$x^n \cos 1/x^2 \to 0, \text{ because } |\cos 1/x^2|<1 \text{ and } x^n \to 0, \text{ as } x \to 0$$
$$(1/x)\sin 1/x^2 \to \infty, \text{ because } |\sin 1/x^2|<1 \text{ and } x^n \to \infty, \text{ as } x \to 0$$

which is again a mechanistic and inaccurate transfer of argument. Camille appears to have a more concrete image of the situation at this moment. She notes about $(1/x)\sin 1/x^2$: 'But it's not quite vast'. In her unconventional choice of words, I think, lies a conception of $\lim f(x)=\infty$ as f becoming 'vast', taking very large values.

Before closing I would like to highlight, on a socioaffective note, Camille's persistent attention-grabbing interventions. In this last case Eleanor, whose shyness is reflected in her mumbling, has suggested that $(1/x)\sin 1/x^2 \to \infty$. The tutor hasn't heard that and asks for a repetition. Camille, who has heard what Eleanor said, comments upon Eleanor's claim. The tutor, annoyed, bursts with T9. Camille then repeats what Eleanor said (most students when confronted by the tutor in this manner would be quiet and let Eleanor repeat). Camille's persistence has repeatedly caused friction between her and the tutor who seems to prefer more submissive students.

Finally I only briefly point out the tutor's metaphor on well-behaved (continuous) and badly-behaved (discontinuous or tending to infinity) functions. She says about f_2: 'the first part is 'well-behaved and you cannot quite cancel the bad behaviour of' the second part'. And she adds:' If they both behave badly they might somehow cancel out the bad behaviour...but that's not gonna happen here'.

Conclusion: In the above, there are occasions where teaching, despite given the students' expressed preoccupations, leaves the connections between mathematical ideas unexplored. Here the case was the concept of derivative and below I list the occasions when one of the students expressed an epistemological concern which was not always satisfied:

- discomfort with discovering answers to questions (in fact limits) intuitively and then proving them formally,
- query about whether the derivative of a function is 'always' continuous at zero,
- comments on another student's claim that $(1/x)\sin 1/x^2 \to \infty$,
- in general attention-seeking interventions that most of the time are deviations from the solution-executing format of the tutorial.

The undercurrent theme of the session is the contrast between the proficient performance of the students who actually find the right answers and appear as understanding the mechanics of differentiation quite well but also

- do not logically complete their proofs (for instance they find out for which n B6i-iii are true but they do not exclude the possibility that for the rest of the natural numbers they are not true) and
- they do not demonstrate a conceptual understanding of what a derivative is and in particular how the algebra of derivatives is derived from and connects to the definition of derivative as the limit of a ratio.

As a result a didactical need emerged for establishing a link between the meaning and the mechanics of calculus.

Section (vi) The Novices' Difficulty with Grounding Intuitive Arguments on Appropriate Theorems. Decontextualised Knowledge, Regression to Quasi-Formal Familiar Modes of Reasoning and the Examples of the Intermediate Value Theorem and the Inverse Function Theorem

Context: See Extract 7.6

Structure: In the following, the tutor and the students discuss an application of the Intermediate Value Theorem and the Inverse Function Theorem. Both students have drawn the graphs in B7 and from them they have intuitively drawn conclusions. Implicitly they have used theorems which they cannot recall. Moreover, even when they reproduce the statement of a well-known theorem, they cannot link it to the intuition that led to answer the questions in B7.

The Episode: A Factual Account. See Extract 7.6

An Interpretive Account: The Analysis

In parts a and b of question B7, the students have drawn the graphs and given intuitive explanations of their (correct) answers. What they cannot do is support their arguments with formal explanations that go beyond the graphical representations of the functions in question and ground their arguments on theorems that they have been recently taught.

By convention, the questions in a weekly problem sheet usually aim at providing applications of the theorems proved in the lectures. Thus the theorems are seen into a context of applicability and are embedded into the growing domain of the novice's mathematical knowledge. Abidul and Frances do not seem to be in tune with these aims. Both employ A-level techniques (first and second derivative tests) to identify the local max and min of the function and use intuitive notions of bound and limit in order to construct the graphs of x^5-$5x$+10 and $x^{-1}e^x$. So they seem to deviate from the enculturating contextualisation aimed at by the authors of the problem sheet which contains B7.

Characteristically when Abidul, under the pressure of the tutor who insists on Abidul grounding her illustrative, intuitive argument (A1) on a well-known theorem, replies hesitantly 'Rolle's', as if she feels obliged to hide behind some impressive name-dropping. Similarly Frances drops in IVT. Both students seem to drop the names at random, because they have heard them recently and they expect them to appear frequently. F1, A2 and A3 are fragmentary reconstructions of IVT. They are also evidence of how disconnected the students' actual mathematical thinking is from the 'official' mathematical knowledge that is supposed to form and support this mathematical thinking.

The situation repeats itself identically in part b of question B7. Again via the derivative tests, the min of the function has been identified and the graph has been drawn. Frances' argument about the smallness/bigness of $x^{-1}e^x$ with relation to the size of x, is underlain by an intuitive and informal notion of limit and her answers are definitely embedded in the graph of $x^{-1}e^x$. Frances explicitly and exclusively turns to the graph in order to answer the question in the second part of b which requires a construction of a function by reversal.

That the second part of b is the most cognitively difficult task in question B7 is evident in Abidul's giving up and Frances' complete reliance on the graph. Frances carries out the reversal (which in thorough terms is allowed by the Inverse Function Theorem under certain conditions satisfied in question B7) required by the question by simply reflecting the decreasing part of the graph in the y axis and by locating the range of values of x that correspond to $[e, +\infty)$. This is exactly what the IFT allows but, as with the IVT, the students use the theorem unconsciously keeping thus its assumptions implicit. Both IVT and IFT — especially IVT — make quite obvious statements (if f is continuous on $[a,b]$ then it takes all the values between $f(a)$ and $f(b)$; if f is continuously differentiable on an interval then its inverse is defined) that, however, hold under certain conditions. Novices tend to over-generalise the domain of validity of theorems only because the theorems apply to a large number of functions they are familiar with. Almost certainly it did not occur to Abidul and Frances that they needed to address a theorem to ground their answers to B7.

The difference between the tutor's and the student's approaches is of a delicate ontological nature. For the tutor IFT guarantees the existence of f^{-1} in B7 but for Frances (F2) this function already exists: it is there, on the graph, and in order to see it, all she has to do is look at the graph in a slightly new way (x becomes y and so domain becomes range). Frances does not need anything stronger than her own sight to be convinced and she assumes that everyone else shares her sense of conviction.

Conclusion: In the above, the students have been employing quasi-formal techniques (intuitive arguments based on graphical representations of functions) in order to

- find the roots of an equation,
- identify the image of a function, and

- construct the image of the inverse of a function.

Their arguments could have been grounded on the Intermediate Value Theorem and the Inverse Function Theorem but this did not occur to the students who, perpetuating A-level attitudes, used the theorems tacitly. The tutor has played the role of an agent trying to link the students' intuitions with the appropriate theorems and to foster in their minds the necessity of doing so. When they finally realise and put into practice the need to go beyond unconscious and unfounded uses of theorems, it will probably be a sign of adaptation to customary and formally acceptable mathematical practices.

Section (vii) The Gap Between the Novice's Advanced Algorithmic Behaviour and Inadequate Conceptual Understanding. An Example from an Application of the Taylor Series

Context:	See Extract 7.7
Structure:	Under the firm instruction of the tutor, the students execute the calculations involved in the manipulation of the inequalities in B10. At one instance the students suggest a differentiating technique (for finding the maximum of a variable quantity) with which they are familiar from school but the tutor prefers a more basic, algebraic alternative. The proof is completed with very closed questioning.
The Episode:	A Factual Account. See Extract 7.7

An Interpretive Account: The Analysis

Part a of B10 requires a reasonable amount of good algorithmic behaviour: it is about the cautious substitution of x^3 for $f(x)$ and 2 for n. Part b however requires going slightly beyond this mechanism of substitution and engage in the manipulation of the Taylor expression for g, g' and g''. This qualitative differentiation in the question passes rather unnoticed by the tutor whose words might perpetuate the students' impression that simple algorithmic behaviour is still sufficient for part b. The tutor's surprise that the students have not completed a question for which 'there is nothing much you can do' is a trivialisation with which the confused students possibly would not agree. Beth then appears ready to 'translate' the Taylor expansion in part a for function g in part b without being concerned for the interval or the point around which the expansion is taken.

Hence, for Beth, the Taylor series, a tool for approximation, is reduced to a machine of plugging in values, detached from possible uses or meaning. With the tutor's instruction, she manages to define the interval and the point around which the expansion is taken and to calculate the expansion for g. Subsequently Beth's observations are correct insights which however she cannot prove.

Cary's suggestion to use a differentiation technique in order to identify the maximum of $2/h+h/2$ is in classic contrast (in this chapter it has been repeatedly pointed out) with the tutor's preference for an algebraic technique, based on a few arithmetical principles. I note then that her first attempt to find such a technique fails and the tutor returns to Cary's suggestion. Unlike other occasions the tutor seems here to change her preference for a first-principles approach to a theorem-quoting one a bit more easily, probably under the pressing immediacy of Cary's suggestion. In principle, she may prefer a more basic proof but she does not deny the convenience of a practical alternative, even if it is a philosophically less acceptable one.

Finally, Beth again exhibits considerable facility with algebraic manipulations when towards the end of the Episode contributes substantially to the proof that for all x in [0,2] it is possible to choose h with $|h|=1$ so that $x+h$ is in [0,2].

Conclusion: In the above, evidence was given of the contrast between the students' relative facility with algorithmic behaviour (the action of algebraic manipulations) and more procedural and conceptual understanding of the Taylor Series as a tool for approximation. Also part of the significance of the instance lies in the short juxtaposition between the tutor's philosophical preference for proofs based on first principles and one of the students' suggestion for using a more sophisticated (but informally familiar from school) tool — identifying the maximum of a variable quantity via differentiation.

Section (viii) The Contrast Between Novice and Expert Approaches to Mathematical Reasoning. The Example of a Convergent Series

Context: See Extract 7.8

Structure: In the following, the students present, or talk about, their evaluation of four infinite sums. Finitist attitudes come to surface as well as the contrast between the expert tutor's sophisticated and embedded approach to the evaluation and the novices' arithmeticised and decontextualised technique.

The Episode: A Factual Account. See Extract 7.8

An Interpretive Account: The Analysis

The Novices' Finitist Attitudes Towards Infinite Sums. In this Episode, Cliff's and Cathy's attitude towards infinite sums is, in brief, to treat them as finite sums. The students subsequently apply a wide range of arithmetical operations on these finitised infinite sums:

- Cliff 'splits up' $\sum 1/r(r+k)$ as $\sum(1/kr - 1/k(r+k))$.
- Cathy 'breaks' the $(-\infty)-(+\infty)$ sum in two: $(-\infty)-0$ and $0-(+\infty)$. Then she removes || and calculates the two infinite sums.
- Cathy on $\sum r^2/3^r$: $r^2=r^2-1+1=(r-1)(r+1)+1$ and breaks the infinite sum accordingly. Since $\sum 1/3^r=1/2$, she turns to calculating $\sum(r-1)(r+1)/3^r$ which she rewrites as the sum of its term at zero plus the sum from 1 to ∞. Breaking the infinite sum once more leads her to obtaining $1/3\sum r^2/3^r$ on the right hand side of the equation. Finally by calculating $\sum 2r/3^{r+1}$ she reaches the result 3/2.

The students' treatment of the infinite sums, which are limits, as finite quantities is illustrative of the students' attitude towards $\sum$ and the ease with which they use the notion of rearrangement. Didactically, the danger of the overextended use of the 'right to rearrange' can be proved to the novices via exposure to the large number of cases where it does not hold. As seen in cases like continuity and differentiability, the novices' impression that infinite sums can be broken, rearranged etc. reflect their finitist views of infinity. It also reflects culturally and epistemologically embedded conceptions, or primary mathematical intuitions, about certain mathematical properties, such as the differentiability of all continuous functions, that permeate through the history of mathematics. Teaching, which is oriented towards the overcoming of these epistemological obstacles, can influence the novice's mathematical growth away from these conceptions. On the contrary, the novices' constant and biased exposure to sums that can be broken and rearranged, such as the ones in this Episode, is likely to result in the perpetuation of these conceptions.

The Contrast Between the Expert's Embedded and Sophisticated Approach and the Novice's Decontextualised Technique. Cathy's way of evaluating the sum in SS7.1iv is a refreshing, back-to-arithmetical basics approach. It is not terribly elegant (a few of her 'moves' are repetitive and circular such as writing r^2 as r^2-1+1, moving 1/3 inside and outside the Σ several times, etc.) but it is pragmatic and straightforward. It has the feel of handy arithmetic and does show skill and imaginative capacity. I note however that only ostensibly Cathy's solution is basic and arithmetical (the only piece of previous knowledge she explicitly employs is that $\Sigma 1/3^r=1/2$). This is a deceptive appearance since, behind Cathy's rearrangements, lies the theory that makes them possible. What Cathy seems to be doing here is unconsciously reducing infinity to the finite rules of a game she knows well, namely manipulating algebraic quantities.

On the other hand the tutor's approach is a formal and elegant shortcut in resonance with the material the students have been taught at lectures and the techniques they will need. It is, in other words, a contextualised choice of technique which is generalisable to a large number of infinite sums. It has the benefit of hindsight and of globality. It shows an expert handling, an informed awareness of the facilities available to the craftsman ($\Sigma x^r=1/1-x$, letting $f(x)$ be $1/1-x$, calculating f' and f'' and noting that f'' can be written in terms of f and f') as opposed to Cathy's decontextualised, hence slightly primitive approach.

None of the above is meant to diminish Cathy's efficient approach which (the dangers of naive rearrangement of the terms in a series aside) yields the correct answer. It only aims at highlighting the inclination of the novice to resort to familiar (here: handling of algebraic expressions) modes of operating at the expense of adopting new, potentially more contextualised and efficient ones.

Conclusion: In the above, the novices' inclination to treat infinite sums as finite quantities was demonstrated and attributed to deeply embedded epistemological beliefs and to the novices' biased exposure to infinite sums that can be harmlessly evaluated with finite techniques. Moreover two approaches to the evaluation of an infinite sum were juxtaposed:

- the novice's basic and arithmetical finitist one, and
- the expert's contextualised, concise and sophisticated and, possibly generalisable to a number of cases, one.

The novice's attitude was attributed to a habitual regression to familiar modes of thinking (manipulation of algebraic quantities) despite the novel experiences of alternative, newer techniques.

PART III A Synthesis of the Findings in the Area of Calculus. Indications for the Cross-Topical Synthesis in Chapter 10

In this chapter the students' first experiences of fundamental calculus concepts were explored in a series of Episodes from both terms of observation. As noted in the Interlude, the concept that emerged as *paradigmatically problematic*, within the area of Calculus was the notion of limit as encountered in a variety of contexts (limit of a sequence, continuity, convergence).

In particular, in their introduction to the formal definition of limit [1, 2], the students are in difficulty with the $\delta-\varepsilon$ mathematical formalism as well as with assigning meaning to the formal definition. In particular, their difficulty seems to be with comprehending the mechanism of the proposition that is contained in the formal definition and with how this mechanism provides a tool for proving limits. The students also demonstrate [1] a tendency to believe that all functions are continuous and this possibly implies a view of limit as a redundant concept. Moreover visualisation proves [2] to be controversial: the novices debate the use of pictures in guessing limits and see pictures as specific representations and this specificity seems to impede the students' shift towards generality. At the same time they resist formal notation and logic as vehicles for mathematical arguments: so they cannot evoke the formal definition of limit or construct a logical statement of an argument for a proof by contradiction. Their linguistic preoccupations include wondering whether the term *not-continuous* is equivalent to *discontinuous* [2]; or whether the derivative of a function is *always* continuous [5][0].

The students' use of ordinary language has been variably successful in their attempts to describe limiting processes in various contexts: the definition of limit [1] or in specific situations [5]. So even in cases where their use of language successfully conveys their general grasp of an idea, this success is not fully integrated in the process of presenting a consummate formal argument. Another illustration of this is evident in the students' proficient algorithmic behaviour with regard to differentiation whose mechanics they seem to understand but do not actually employ fully [5] in order to complete a formal proof. Also they do not demonstrate a conceptual understanding of what a derivative is and, in particular, how the algebra of derivatives is derived from and connects to the definition of derivative as the limit

[0] where *always* was given different significations by the tutor (local) and the novice (universal).

of a ratio. More evidence of this behaviour was given in the context of Taylor Series [7] and within the new context of convergence of a series[1].

The students seem to have been sensitised to the requirements of university mathematics for rigour: this engenders a hesitation towards school mathematical practices which deters them from referring to and employing previous mathematical knowledge. The novices possibly need to clarify the distinction between rigorous and intuitive arguments, legitimate and illegitimate use of knowledge that is thought of as previously established. On the other hand there are occasions [6] where the novices are still at the stage of unconsciously using knowledge that is taken as so obvious as to be not in need of establishing formally (like the propositions of the *Intermediate Value Theorem* and the *Inverse Function Theorem*). This tacit use of theorems is evidence of a perpetuation of A-level attitudes in considerable distance from the expected mathematical formalism. Similar regression to familiar-from-school modes of action was demonstrated in the novices' inclination [8] to treat infinite sums as finite quantities, possibly influenced by deeply embedded epistemological beliefs about ∞ as well as a biased exposure to infinite sums that can be harmlessly evaluated with finite techniques. Again the novices' finitism was juxtaposed with the expert's contextualised, concise, sophisticated and, possibly, generalisable approach.

The students' most prominent difficulty in the context of limits seems to be finding limits either via the definition or via the algebra of limits [1, 2, 3]. Using inequalities in order to manipulate quantities, graphing functions, guessing limits and using the algebra of limits are mathematical practices that the students question as to their rigour and, hence, as to their acceptability. So due to their growing mistrust towards the practices of school mathematics, they begin to avoid intuitive practices, such as guessing limits. So, for instance, when [2] it is not overtly justified to them why the use of the definition of limit can alternate with the use of the algebra of limits, they seem severely confused with the *tension* between what can be called *Proof-By-First-Principles and Proof-by-Theorem-Quoting*. This tension has been generally observed in these tutorials in various contexts. Similarly the novices are seriously perplexed with [3, 4] distinguishing between the practices that they are supposed to espouse in different mathematical domains: in applied mathematics, retrospective use of unproved results is allowed, as opposed to foundational courses, where it is not. This difficulty echoes the difficulty reported in Chapter 6 to distinguish between what knowledge they are allowed to assume and what has to be established by formal proof.

---

1 Reported in (Nardi 1996) where the student seemed mistrustful towards arguing mathematically in ordinary language and was concerned about the legitimacy of this practice.

The tension between First Principles and Theorem Quoting was also exemplified in the context of juxtaposing novice and expert practices: the expert [3, 7] not only distinguishes between different approaches to a proof but also expresses a preference for a proof (usually via first principles) on historicoepistemological grounds. The students' occasional preference for more sophisticated tools, such as identifying the maximum of a variable quantity via differentiation, not via basic algebraic inequalities [9], is possibly due to their — unrigorous — familiarity with these tools from school.

The novices tend to generalise unrigorously or react to what they perceive as tedious rigour: so they generalised [3] the finite case of $\lim(a+b+c)=\lim a+\lim b+\lim c$ into $\lim\sum=\sum\lim$ which is actually a double limit, thus disregarding unconsciously the conditions for the exchange of limits; they also reacted with suspicion and hesitation towards considering the limited validity of a theorem via a counterexample: characteristically for novices, one of them thought it was not useful to refute a theorem by a counterexample when it holds in the particular case in which he used it; another tried to negotiate his right to use elliptic arguments in writing, as long as he can explain what he did orally [3].

The tensions and the conflicts reported above can be seen as part of an *enculturation process* during which the novices learn about certain conditions of the didactical and epistemological contract in formal mathematical activity. So when, for instance [3], in the interaction it is revealed that they have been unconsciously assuming the validity of exchanging double limits, using thus a theorem which is as yet unproved in the course, the novices seem to be learning that they need to exercise control over their mathematical reasoning in order to avoid subconscious and unjustified decisions and that this skill to control is a characteristic of formal mathematical behaviour.

The students' perplexity with the status of rigour of the various approaches to finding and proving limits [1, 2, 3, 5] — even when they appear alerted to the dangers in not clarifying the foundations of their proofs [3] — makes it necessary to consider it as a task of the first year course

- to establish the legitimacy of school mathematical practices as a way to gain mathematical insight, and
- to introduce formal mathematical language as a way to refine and establish these insights rigorously.

This is in contrast with certain current practices according to which the novices are advised to leave behind their school-mathematical way of thinking and start anew by trying to build mathematics on the solid foundations of mathematical formalism: this is a suggestion by the tutors and lecturers which ought to be understood as a manner of speaking and not as a literal translation into practice. The distinction however is not clearly made by the novices who at times perceive these suggestions as contradictory and are thus led to inefficient practices such as avoiding the algebra of limits or avoiding to intuitively find a limit by looking at a graph.

A substantial part of the novices' action seemed to concentrate on their effort to *assign meaning and significance to the new concepts* [4]. This is usually a process which takes place under the severe influence of both emotional and cognitive/contextual factors [1 and 4 [2]]. More detailed evidence of the meaning bestowing practices in which the novices are constantly engaged follow in Chapters 8 and 9.

[2] where a student attempts various interpretations of the symbol ~ in the definition of a Fourier Series (approximation, equivalence, equality) and also probes in detail several related concepts (periodic function, continuity, side limits).

Chapter 8

The Novices' Encounter With Mathematical Abstraction: Cases from *Linear Algebra*

PART I A Guide to the Paradigmatical Cases (Episodes) Presented in this Chapter

The following table contains contextual information with regard to the 8 Episodes presented in this Chapter.

Episode Number	Time of Incident Term - Week	Participants	Mathematical Content
1	Michaelmas 5	four pairs of students in one college	span of a set
2	Michaelmas 6	four pairs of students in one college	subspaces of a vector space, the Subspace Test
3	Michaelmas 6	Connie	LA5.23&24, spanning sets
4	Michaelmas 7	Jack and Andrew	LA6.26, bases of a vector space
5	Michaelmas 7	Jack and Andrew	LA6.29, bases of a vector space, dimension of a vector space
6	Michaelmas 8	four pairs of students in one college	the matrix of a linear mapping between two vector spaces and an example from $P_3(\Re)$
7	Michaelmas 8	Cathy and George	LA7.35, image, kernel and the *Rank and Nullity Theorem*
8	Hilary 2	two pairs from one college	B3, mappings, matrices and bases

PART II Data and Analysis

In the following I present the factual and interpretive accounts and conclusions for the 8 Episodes of the table in the previous page. In Part III then I synthesise the findings of Part II related to Linear Algebra and briefly discuss the wider cognitive issues that are presented in the overall synthesis of the data analysis in Chapter 10.

Section (i) Constructing the Span of Various Sets as an Example of the Generating Procedure of Spanning and the Debatable Value of the Metaphor of the Plane

Context: See Extract 8.1

Structure: In the following, the unanimous unease of the students with their introduction to the concept of spans and spanning dominates the conversation. The students, with the exception of Camille, once they are given the 'more practical' definition of Span, seem to forget about the abstract definition given in the lectures. The underlying tension then seems to be between
the novelty of the concept of Span and
their familiarity with the plane.
Also even though only peripherally mentioned, the notions of linear independence and basis also appear to cause a general unease.

Note to the Reader: This is a presentation of four tutorials of the same mathematical content, tutor and college. The discussion in the three of the four tutorials, 1, 2 and 4, is presented in parallel (in 3-column tables linked by a common main narrative). Tutorial 3 is presented separately.

The Episode: A Factual Account. See Extract 8.1

An Interpretive Account: The Analysis

As the tutor herself stresses, even though at the moment of her statement she had not realised that the definition of span given in the lectures was the abstract one (see Context), the Linear Algebra course moves from matrix operations to the more general context of vector spaces. As it becomes obvious in this Chapter the transition is not smooth and unproblematic at all. These tutorials are a characteristic example of the immense influence on the students' learning process of the degree of abstraction within the concepts they are introduced to. In the following, I present the evidence for what I think is the underlying tension of these sessions; the tension between the novelty of the concept of Span and the students' familiarity with the geometrical metaphor of the plane.

Patricia's Use of the Term Spanning. As a flavour of the novices' typical treatment of the novelty of spans and spanning sets, I cite Patricia's use of the term *spanning* in P4. Similarly in another part of the tutorial, not included in the Episode, Patricia (asked about the spanning set with regard to an example from the set of solutions in a Differential Equations problem) replies 'ax plus...', that is she starts describing a linear combination in the Span instead of suggesting a Spanning Set. In this Chapter similar cases of confused terminology, evidently reflecting the novices' ambivalence about the meaning of the concepts, will be highlighted frequently. The novices' struggle with new terminology is certainly not surprising but, in the case of spans and spanning sets, this struggle has an additional complication: a set is the span of its spanning set. This syntactical association of the two terms reflects their conceptual dependence: spans are created out of spanning sets. Spanning sets are sets-that-span. So Patricia, in the above mentioned example from Differential Equations, seems to respond to the tutor by pointing out an example of this creation (a linear combination that is an element of the span) instead of the creator (the spanning set). A potential significance in the above distinction is that the novice's difficulty in distinguishing linguistically between the two terms may reflect the more essential lack of understanding the conceptual mechanism that connects the two concepts (spanning set as the cause; span as the effect).

A similar case of the above mentioned ambivalent use of terminology is evident in B1 and B2, which illustrate Beth's fragility of expression in general. Her interpretation of $<v_1>= \varnothing$ when $v_1=0$ may be linguistically interpreted in terms of thinking that

since zero is nothing, it produces nothing (the empty set).

Beth's Inflexibility in Moving from One Dimension to Another. Looking at Beth's development of ideas during the tutorial (from B1 to B4 and then her contribution in finding the span of $\{v_1, v_2, v_3\}$) I note an inflexibility in understanding the construction of spans as a way to describe gradually spaces of a higher dimension. For instance in B3 she seems to understand that $<v_1>$ is the 'line' that carries v_1. When v_2 is added in the picture she also knows that $<v_2>$ is the 'line' that carries v_2 but finds it surprising that $<v_1, v_2>$ is something more-than-its-parts, namely something more than just the two lines together. Possibly Abidul's 'and' in A3 alludes to the same conception. Similarly, when constructing $<v_1, v_2, v_3>$, she fails to see the possibility of $v_3 \notin <v_1, v_2>$, therefore delaying to see the incoming third dimension, something which all the other students in this tutorial noticed immediately (see next paragraph).

I see Beth's cognitive behaviour here as impeding her transition to generality and in connection with the questionable use of the metaphor of the plane that dominates these tutorials, I have some reservations as to whether Beth is constructing the necessary links between the definition of Span given earlier in the tutorial and the example of the line, the plane and the space.

From the Line, to the Plane and the Space: the Convenience of Learning the Recipe Well. As briefly mentioned in the previous paragraph all students when asked to construct $<v_1, v_2, v_3>$ appear ready to make the distinction between

$$v_3 \notin <v_1, v_2> \text{ and}$$
$$v_3 \in <v_1, v_2>.$$

Given that only Camille made this observation in the earlier cases — and she is the only one who mentions a generalisation to k dimensions in the end —, it seems reasonable to assume that the students *learned* to make the distinction from watching the tutor doing so. There is no evidence as to whether the necessity of the distinction has been understood or whether the students would make the same distinction if they were given k instead of 3 vectors to span. In any case the tutor's progressive build-up of the span of one, two and then three vectors seems to have yielded the desired outcome: visualise $<v_1, v_2, v_3>$ as the three-dimensional space. In a sense, this is likely to be learning by mimicking, by acquiring a habit for reasoning in a not completely rationalised way; in other words this is likely to be a case of acculturation to the habit of generalisation in advanced mathematical thinking.

The Questionable Use of the Metaphor of the Line, the Plane and the Space. In the following, I cite evidence of the students' use of language in these tutorials that illustrates the strong geometric approach that dominates their thinking about $<v_1>$, $<v_1, v_2>$ and $<v_1, v_2, v_3>$. I stress that I find rather alarming the possibility that this geometric metaphor is taken as the literal message. Additionally in these tutorials, with the exception of Camille, no reference is made to a transition to the geometrically-not-so-easily-conceivable bigger-than-3 dimension. In the following I have compiled some brief lexical references from the students' expressions with the purpose to convey their strictly geometrical — as opposed to the intended algebraic — frame of mind:

A1 (multitude)
F1 (the distance a times v_1, turning around zero)
E1 (how far to this direction)
C1 (take components)

I note that the students' pointing at, bending towards the drawings of the tutor as well as the use of their hands (for instance P3 and B3) also convey an impression of their close adherence to the metaphor. I also note that the tutor intervenes only in C1, possibly disturbed by the direct reference to school geometry, in order to explicitly distinguish between vectors on the plane and vectors as elements of a vector space. Moreover in P1 Patricia appears to attempt a more abstract description of $<v_1, v_2>$ but the tutor pulls her back to a 'pictorial' description. Patricia probably was in less need of the scaffolding metaphor of the plane and the pulling back can be seen as trying to condition Patricia within the cognitive frame of the tutor's agenda.

A Verbal Proof of $<v_1, v_2>=<v_1>$ *when* $v_2 \in <v_1, v_2>$. A4 is Abidul's idea of why $<v_1, v_2>=<v_1>$ when $v_2 \in <v_1, v_2>$ and it illustrates the commonly encountered phenomenon in these tutorials that the novices have an adequate argument in their minds but have difficulty in expressing this argument in formal mathematical terms. Cathy's explicit doubt whether this way of expressing herself is 'formal enough' (Extract 8.7) is another example of the tension between ordinary and formal expression and its consequences on learning.

The Case of Camille: When Affective Factors Determine Crucial Aspects of the Learner's Cognitive Behaviour. Camille is one of the students that have left the most striking impressions on me with regard to their learning style. In this Extract as well as others (for instance Extract 8.2) Camille appears as the most confident and determined among her peers in this college. She is the most vivid participant and interferes incessantly demanding or giving explanations. She is not afraid of risking interpretations that may prove unacceptable, she tends to think constructively and comparatively and leaps to brave generalisations. She talks explicitly about her difficulties and rarely allows the change of the subject if she does not feel satisfied with her understanding of it. Occasionally her queries are about mathematical terminology in English because she has been educated in French; also her mathematical background is at times different to that of some of the other students. In the following, I present the evidence of Camille's behaviour as outlined in this paragraph.

The Interference of English as a Second Language in Learning. In this discussion Camille's thinking is interrupted four times in order to clarify the use of a number of mathematical terms: *scope* (slip of the tongue for 'span'), *directions* (for 'vectors'), *plan* (with French pronunciation for the 'plane'), *taking projections/ dropping perpendiculars*. These interferences, though rather superficial in their influence on Camille's flow of thinking, illustrate how semantic issues determine understanding. I suggest there is a parallel to these interferences from another language — French in this case — with the interferences on the students' learning from the semantics of mathematics as, among everything else, a symbolic language. The issue reappears in Chapter 9 and is dealt with there in greater detail.

Conceptions of Subspace and Span: Signs of a Constructive Mind. Camille is the only one among the students in these tutorials to look critically at the definitions of Span given in the beginning of the Episode. Her observations about $0 \in S$ and $S \subseteq \langle S \rangle$ as well as her reasoning for them are evidence of a mind in the process of constructing a meaning for $\langle S \rangle$ from trying to understand its constituent parts, its elements and its properties. Uniquely in these tutorials Camille is also addressing the issue of the equivalence of the two definitions of Span and in particular that $\langle S \rangle$ is the smallest subspace of V that contains S.

This comparative and constructive approach is again evident when Camille intervenes in Cleo's question about $\langle S \rangle$ being a subset of V (I note here that several times during this tutorial Cleo appears to be in difficulty to keep up with Camille's thinking: her participation seems to be suppressed by Camille's impulsive reactions. Also Cleo, when personally addressed by the tutor, does not give any evidence that she either follows completely the arguments exchanged in the discussion or is willing to participate extensively): Camille extends the notion of subset of V as a set that, provided that it has the properties of a vector space, becomes a subspace. Apart from the impressive amount of confidence that Camille's reaction exudes — she actually replaces the tutor in this instance by giving explanations to Cleo — it is also a demonstration of her concept image of a subspace as a set endowed with some essential features which she explains in an instrumental and illustrative way.

From a methodological point of view, Camille's confident exposure of her ideas makes her cognitive processes more transparent. This is beneficial to the purposes of the study because this transparency reduces the amount of the interpreter's explanatory intervention; that is, it reduces the degree of necessary analytical processing in the extraction of explanations. Unfortunately Camille is among a minority of students participating in these tutorials who approached tutorials in this way from the beginning. Fortunately, during the two terms of observation, more students developed similar behaviour.

An Idiosyncratic Suggestion for Adding Two Vectors on the Plane. Asked about how can a vector on the plane be expressed in terms of two vectors that span the plane, Camille suggests (fig.1e) taking the projections on the two axes and 'measuring the distance' (C3-C5). Even when given some time to think, and given Cleo's weakness or indifference in changing Camille's mind, she insists. It is likely that Camille's suggestion originates in the common way of identifying the co-ordinates of a point on the plane (or the space) with regard to a Cartesian, orthogonal system. If this is true, then it alludes to the persistence of the geometric metaphors in the students' minds as well as the dubious influence of cross-references to their different aspects (here the Parallelogram Rule as a way of adding two vectors on the plane interferes with the Cartesian orthogonal addition of vectors). C6 is Camille's explanation for C3-C5 and it is reasonable to regard it as a reinforcement of this 'orthogonal' explanation.

I note that the change in Camille's mind and the elicitation of the Parallelogram Rule is achieved by the tutor with three reversals in the flow of her teaching:

- adopting Camille's idea of 'dropping perpendiculars' and
- reversing the question of expressing a vector as a combination of the two vectors that span the plane to a question of how two vectors are added.
- by accepting Camille's 'minus bv_2' and asking her to show the 'plus bv_2' on the drawing.

With the first and the third the tutor helps Camille rid of her misleading ideas and with the second to replace one of them with the desirable one.

Remarkably the tutor exhibits a considerable amount of flexibility in her efforts to modify Camille's ideas by adopting the student's perspective and then questioning it from within. As a result, in the end, the responsibility for Camille's changing mind is successfully transferred to Camille herself. As the evidence suggests, in this and other chapters, most strikingly Chapter 9, this tutor usually employs highly leading techniques of very closed questioning. Therefore the flexibility she exhibits here is rather unconventional. In a sense Camille's unconventional 'orthogonal' conception of vector addition seems to have yielded an equally unconventional (out of character) response by the tutor.

I note that like Patricia in P3, Camille looks at the plane and decides about the sign of the vector: it is minus because it is 'on the left'. Again the relative sense in which representations, such as a drawing, are supposed to be taken is not entirely understood by the novices.

Finally I note that Camille in this and other tutorials is explicitly expressing her fear of anything beyond three dimensions. As with her difficulty to imagine sets of functions 'after 3 it gets confusing'. Paradoxically she is the only one among the students in these tutorials who achieves a reasonable generalisation — notably unprompted by the tutor — in C7. From a cognitive point of view this may not be paradoxical at all: Camille is more capable of achieving generalisation because at first she is the one who is conscious of the difficulties involved in generalising. Once she has become suspicious or aware of the cognitive leap involved in generalising, she can attempt taking the leap. Consciousness in these terms then emerges as an enhancing determinant of cognition.

Conclusion: In the above, evidence was given of the novices' semantic and conceptual difficulties related to the novel notion of span and spanning set and of the influence of geometrical metaphors on advanced mathematical cognition. In particular, the novices seemed to miss the grammatical link between the terms *spanning set* and *span* — which reflects their cause-and-effect conceptual link. This linguistic deficiency possibly mirrors and partly determines the novices' restricted understanding of these new concepts.

In terms of the students' learning process with regard to the necessity to generalise (the strategy of spanning from 1 to 3 dimensions), evidence was given that the students respond to the tutor's stimuli for generalisation with various degrees of readiness. Almost paradoxically, the only student who committed an attempt to generalise with some facility, was the one who complained about the difficulty of generalisation. Consciousness of difficulty was then identified as an enhancement of the possibility to overcome difficulty. The students also appeared to be in difficulty to express formally an intuitively grasped idea regarding spanning sets.

Moreover the novices' literal interpretation of the geometrical metaphor of the plane was illustrated in their use of strictly geometrical language regarding vector addition, their body language and their orientation on the plane. Most vivid of these illustrations was one of the students' use of a Cartesian orthogonal system, in which two different aspects of the geometrical metaphor of the plane interfered with each other and with the novice's understanding of the construction of $<v_1, v_2>$.

From a methodological point of view a contrast was identified — and discussed in terms of its influence on the study — between learners who are willing to reveal and discuss their thought processes and learners who are not as willing.

From a didactical point of view an example was discussed of a successful attempt to modify a student' s perspective from within, that is by adapting their point of view and undermining it with key prompting questions.

Section (ii) The Contrast Between Algorithmic Ability and Conceptual and Contextual Understanding: Applying the Subspace Test and Looking for the Zero Element of $\Re^{\Re}$

Context: See Extract 8.2

Structure: Similarly to Extract 8.1, where the students learned the 'recipe' of finding the span of a set, the students here appear good at organising their algorithmic thinking (see first two applications of the Subspace Test in the Context). Unlike these two applications that involved thinking in terms of $M_n(\Re)$, the application presented in this Extract involves thinking in terms of $\Re^{\Re}$, with which they are less familiar. This lack of familiarity results in the students' unease with handling the question.

Note to the Reader: Same as Section (i).

The Episode: A Factual Account. See Extract 8.2

An Interpretive Account: The Analysis

In the following, I comment upon the students' handling of the zero element of $\Re^{\Re}$ as well as Camille's preoccupation with the contents of $\Re^{\Re}$. I note that the first frictions (though not mentioned in the Context) with regard to the zero element of a vector space appear in the discussion of the first two applications of the Subspace Test: for instance when Abidul uses the term 'nought' for the zero vector of $M_n(\Re)$, the tutor corrects to 'zero' (meaning: the zero matrix). Subsequently Abidul asks whether she can write 'just 0', not the full matrix with zeros everywhere. The tutor says it's fine as long as it is clear what she means.

Another element that seems to be setting up the scene for Extract 8.2 is the fact that the students in the beginning of the tutorials do not know the Subspace Test. However during the tutorial they are given the opportunity to apply it several times and by the time they reach the third application (that is this Extract) they appear to be handling it quite well — for instance, all of the students immediately suggest the evaluation of $af+bg$ when the tutor asks for the second condition of the Subspace Test. Similarly to Extract 8.1, where the students in the course of the tutorial learned the 'recipe' of finding $<v_1, v_2, v_3>$, the students here are quite good at organising their thinking along solidly outlined algorithmic procedures. The problematic element is that, while the previous two applications of the Subspace Test involved thinking in terms of the elements of $M_n(\Re)$, here the vector space in question, $\Re^{\Re}$, is less familiar. In the following, I present the events from the perspective of this lack of familiarity and highlight the outcoming learning difficulties.

Looking for the Zero Element of an Unclearly Defined Vector Space. Some of the students here appeared ready to start applying the Subspace Test on a set, $U \subseteq \Re^{\Re}$, whose contents were not well aware of. As a result of this rather mechanistic choice, they soon impinge upon the identification of the zero element of $\Re^{\Re}$. Even though some of the students can recall the formal definition of the zero element in a vector space ($a+0=0+a=a$, $\forall a \in V$) all of them either remain silent or, after persistent prompting, suggest 'stays the same' or 'x', namely the identity function. Even when Eleanor says 'nought' towards the end of the incident, one is tempted to think that she simply means the real number zero — and not the function $z(x)=0$.

The evidence of the students' difficulty to understand, and furthermore to act upon, $\Re^{\Re}$ is even stronger and more elaborate in the tutorial with Camille. I note that, once the hurdle of identifying the zero element of $\Re^{\Re}$ is overcome, the students are still uncomfortable with the contents of $\Re^{\Re}$ — as their frequent confusion of f with $f(x)$ illustrates. Considering separate values of f as elements of $\Re^{\Re}$ is an indication of their difficulty to view f as an object contained in a set of similar objects, as an element of the set $\Re^{\Re}$. Again the tutorial with Camille provides strong evidence of this difficulty too.

Camille's Preoccupation with the contents of $\Re^{\Re}$. C1-C8 illustrate Camille's struggle to construct a meaningful interpretation of $\Re^{\Re}$; C9-C11 her narrow concept image of a vector and the expansion of this image that the understanding of $\Re^{\Re}$ necessitates.

Given that Camille is a learner who is not afraid of expressing risky ideas (see Extract 8.1), C1 is her first attempt to imbue meaning to $\Re^{\Re}$. C2-C5 are indications that despite the definitions given by the tutor, Camille — probably not having found them helpful — insists on trying to relate $\Re^{\Re}$ with $\Re^2$. Knowing that the set of (x, $f(x)$) is a subset of $\Re^2$, where $f: \Re \rightarrow \Re$, her thinking gets entangled with f as [a set of] ordered pairs and f as an object-element of a set. In C6 she seems to have realised that $\Re^{\Re}$ is a set of functions and she explicitly expresses her difficulty in shifting from a set of elements — for instance matrices — to a set of functions.

C7 then leaves an impression of a second-round effort. Now Camille possibly interprets $\Re^{\Re}$ as a 'power' of $\Re$, thus 'containing' $\Re$ — in the same sense that 2^3 'contains' 2? The chain of tentative interpretations of $\Re^{\Re}$ continues with C8 which seems to be a compromising combination of her entanglement with $\Re^2$ and her newly-acquired idea that $\Re^\Re$ contains functions. In any case C1-C8 is an impressively explicit illustration of a novice's striving for meaningful interpretations of the concepts — and the notation attached — they are being introduced to. Camille's attempts here highlight the failure of the transmissive model as a descriptor of learning. Camille's struggle cannot be described in terms of transmission and reception of knowledge. What the tutor repeats again and again about $\Re^\Re$, is variably perceived by Camille as her unendingly changing interpretations of $\Re^\Re$ indicate. Camille seems to be very much alone in this process even though the steps to the various directions she is taking are obviously fed by the information she receives.

Finally, the narrow image that Camille has of a vector (also evident in C6) is revealed in C9-C11 and, in particular, in her attempt to combine the information that

- the zero vector is a function
- this function has a particular property related to denoting it 'zero vector'

in order to deduce that the zero vector of $\Re^\Re$ is the zero function. Camille still seems to struggle with understanding how the elements of $\Re^\Re$ can be functions. In particular C11 highlights a schism between the notion of vector and the notion of function.

The evidence in this study suggests (for instance Extracts 8.1 and 8.3) that the students have acquired a persistent geometric image of a vector through teaching that focuses on the examples from the line, the plane and the space. Understandably the teachers are repeatedly suggesting to their students to espouse a geometric approach in order to demystify Linear Algebra and especially Vector Spaces which are deemed as too inaccessible and abstract by the learners and to embed it in already familiar contexts. This practice unfortunately proves short-sighted because, within weeks of their instigating a geometric approach, the tutors are forced to generalise their discourse on vector spaces to cases where the geometric metaphor not only does not apply but also creates distracting interferences.

Conclusion: In the above, the students' narrow-minded algorithmic application of the Subspace Test on a subset of $\Re^\Re$ (whose contents they were not aware of) was shown as an action-in-void via their difficulties to identify the zero element of $\Re^\Re$. These difficulties were illustrations of a narrow concept image of vector and of a weakness to perceive function as an object-element of a set. Moreover a student's bold interpretations gave evidence of her uniquely individual struggle for a meaningful construction of $\Re^{\Re}$. As in Section (i), the students' difficulties with an abstract perception of vectors was here partly attributed to teaching that adheres to a large number of examples from the line, the plane and the space. Finally a methodological point similar to Section (i) was also made.

Section (iii) Spanning Sets and the Struggle for a Meaningful Metaphor

Context: See Extract 8.3

Structure: In the following, the thread that connects the fragments of Connie's sporadic querying about various questions in LA5 is a concern for her problematic encounter with the concepts of spanning set and basis. The tutor tries to resolve her queries by adapting an expository approach and by drawing Connie's attention to some formal properties of spanning sets and bases.

The Episode: A Factual Account. See Extract 8.3

An Interpretive Account: The Analysis

In the following, I comment upon Connie's efforts to construct a meaning for spanning sets. The tutor presents definitions and insists on drawing examples from the geometrically familiar case of the plane. He also insists on drawing Connie's attention to some formal properties of spanning sets and bases, such as their non-uniqueness. Dwelling on examples from the plane and insisting on exposition regarding formal properties at times distracts from investigating some of Connie's perceptions as shown below in more detail.

Analytical Versus Geometrical Approach: how resisting alternative approaches impedes understanding. As mentioned in the Context, from the beginning of the session, and in consistence with his recommendations in the morning session, the tutor constantly encourages Connie to think of spanning sets and bases in terms of the two-dimensional Cartesian plane; also to adopt a geometrical perspective. Connie however repeatedly resists (C1, C3). With the tutor's second geometric prompt (T3) she seems to begin seeing the span of (1,0) and (1,1) as the whole plane (C4).

The tutor's subsequent theoretical exposition on the notion of basis does not seem to impress Connie. A few minutes into his exposition Connie interrupts in order to draw his attention back to her specific queries on LA5. Underlying her difficulty with the conditions under which $<<S>>=<S>$ (I note also that what does not appear in the transcript above is Connie's slip-of-the-tongue in which she misinterprets ≤ as 'less than' instead of 'a subspace of') is her discomfort with the notion of spanning. So, despite the fact that it was she who insisted on talking specifically about LA5.24, it is she who interrupts the tutor's exposition on the question in order to discuss the notion of a spanning set (C5). Connie here appears to attempt to construct a parallel between spanning sets and mathematical knowledge she is already familiar with, such as the image of a function ('all about the values that thing can take'). The tutor however does not pursue an inquiry into this construct. On the contrary he prefers picking on Connie's phrasing (THE spanning set) and recommends caution: a spanning set of a set is not unique. His comment, even though mathematically valid, seems to be, didactically, a distraction from illuminating further Connie's attempt to construct a meaning for a spanning set.

However, Connie insists on her image-making (C6). She is struggling for a meaningful metaphor, for an embedding of the concept of spanning set in a familiar and operational context. When the tutor defines <S> as the set of all linear combinations, she wonders 'Why are they useful?' (C7). The tutor then presents her with an explanation according to which spanning sets themselves are not interesting, but bases are. Then again Connie interrupts the tutor's theoretical exposition in order to draw him back to the specificities of her query on how LA5.24 and LA5.25 connect. As said earlier, Connie does not seem to be interested in the tutor's theoretical exposition: instead of another lecture, she seems to need to concentrate on the status of these new concepts. Her not very eloquent expression in C8 hints at her beginning to perceive a basis as a spanning set of a minimum size ('the smallest'). The tutor again tries to shift her attention on non-uniqueness and again she seems to ignore his remark ('THE basis' in C9). Her basic concern is still about the concept of a spanning set and a basis (C10) and the tutor then once again militates for a geometrical approach (plane, space). The inappropriateness of doing so is demonstrated in her extension of

- his parallel of i-j-k vectors from Geometry with a basis to
- her thinking of polar co-ordinates r-θ as a basis.

So Connie here seems to think that every co-ordinate system is a basis. To the latter he reacts alarmed but he does not explain why this parallel does not work (T4), leaving Connie probably in doubt about her idea that every co-ordinate system is a basis, but not clear about why this is not so. Finally C12 and C13 are rather fuzzy verbalisations of her final interpretations of what a basis is. The tutor accepts them — I think they do not contain anything immediately wrong but are indeed very ambiguous in their generality.

I note here that, at one point, the tutor suggests banning the term spanning set altogether and replace the relevant phrasing with 'a subspace is spanned by a set S'. This alludes to a more generally observed tendency of the novices to disregard the notion of a spanning set as redundant. I recall here that, in Greek for instance, there is no word for *spanning set*, only for *span*. Details will follow in the analysis of more material in this chapter.

Conclusion : In the above, evidence was given of a novice's struggle to construct a meaning for the newly introduced concept of spanning set. Within this struggle the novice tried to employ metaphors she is already familiar with, for instance from her study of functions. Some of her concept images also illustrated that she exaggerated the tutor's persistent recommendation to use the metaphor of the plane. She also seemed to construct mutually contradicting images that however coexisted independently. Finally the student persisted on her search for a *raison-d'-être* in the introduction of new concepts.

Didactically, a point similar to Sections (i) and (ii) on the use of the metaphor of the plane was made. Moreover the tutor's tendency to resort to a theoretical exposition on formal definitions was questioned in terms of whether it supports the student's meaningful construction of the new concepts.

Section (iv) Constructing Bases For Various Vector Spaces: the Inadequate Targeting of Essential Intuitive Ideas and Induction Into the Notational Language of Advanced Mathematical Thinking

Context: See Extract 8.4

Structure: In the following, the tutor leads the students to the discovery/construction of bases for sets of matrices. The tutor is flexible to adapt his plan to the students' queries and is focusing mainly on the students' writing and expression. In this sense, this Episode can be seen as an exemplary case of the process of induction into the codes and language of advanced mathematical thinking.

The Episode: A Factual Account. See Extract 8.4

An Interpretive Account: The Analysis

Difficulties with Notation and Inadequate Targeting of Essential Intuitive Ideas. Throughout the session Andrew seems to be in great difficulty with the notation suggested by the tutor. Jack, on the contrary, seems to assimilate the new notation almost instantly. As a result, while Jack is already manufacturing bases by manipulating the matrices E_{pq}, Andrew is still uncertain about the definition of E_{pq}, or having difficulty with why $Tr(E_{ii})=1$.

Jack appears to be more understanding of — and more efficient in resolving — Andrew's difficulty than the tutor: he describes what the elements of E_{pq} are like in ordinary English. Andrew's immediate understanding brings the tutor to the apologetic position of not only rewriting the definition of E_{pq} in a less convoluted but equally rigorous way but also of drawing E_{pq} on the b/b.

So far in the Episode, my impression of Andrew's unease was emphasised by its contrast with Jack's swift performance. Andrew's difficulty can be attributed to the difficulties associated with the notation used in the tutorial. However in his subsequent suggestion 'taking the sum of E_{pq}' Andrew appears to ignore the more fundamental fact that the elements of a basis of a subspace must be elements of the subspace. This is not true of the 'sum of E_{pq}' (whatever that means) or the E_{ii}, none of which are matrices with trace zero.

I think that here Andrew is struggling, not against a complete impasse, but against some very unclear, elusive intuitions. The 'sum of E_{pq}' is perhaps a hint at the commonly used idea that every nxn matrix E_{ij} can be written as

$$\Sigma\Sigma x_{pq}E_{pq}, \text{ where } p,q=1,\dots,n$$

Jack too captures almost successfully (with 'It's one less') the idea that since the n elements of the diagonal do not vary independently — there is *one* condition connecting them — their degree of independence decreases by *one*. Finally Andrew's 'Yes, but I was thinking of E_{ii} from 1 to n-1' seems to be vaguely targeting the idea of E_{ii}-E_{nn} but he doesn't accomplish that fully. The tutor eventually delivers a suggestion for a basis of the set of matrices with trace zero. Andrew's continuing discomfort is illustrated in his request for a detailed presentation of how the set proposed by the tutor spans the set of matrices with trace zero and is linearly independent.

Conclusion: In the above, the novice's encounter with an extremely formalistic notation obstructs the unfolding of his problem-solving. Moreover the novices seem to be targeting, but not fully grasping, intuitive ideas that are essential for the completion of the mathematical task (which finally they do not completely achieve). So, while the expert and the novices agree in the beginning of the tutorial about the method of approaching the problem of finding the dimension of a vector space, they seem to differ in terms of the implementation of the approach: as a result their interaction evolves into an initiation process during which the students, with variable ease, become familiar with a new notational tool. Hence their learning becomes a specific struggle for accommodating this new tool with the vivid intuitive ideas they have about the solution of the problem. An obstacle to this accommodation seems to be the complicated appearance of the new notation.

Section (v) $\dim(X+Y)=\dim X+\dim Y-\dim(X\cap Y)$ and the Varying Persuasion of Mathematical Arguments

Context: See Extract 8.5

Structure: This is a proof and refutation session mostly driven by the tutor. The tutor questions the students' claims twice; one is received with suspicion; the other enthusiastically. Finally tutor and students collectively prove a well-known formula and the students suggest an extension of the formula.

The Episode: A Factual Account. See Extract 8.5

An Interpretive Account: The Analysis

Note: I note here the expression of a general complaint by the students with regard to lecturing. Jack's 'we've never found a basis for anything' (or his stressing that he remembers the theorem because of his reading and not the lectures) is a direct attack against what is perceived by the students as excessive accumulation of theoretical knowledge in lecturing. The tutor's request to return to the specific discussion of the theorem is a common response of tutors in similar situations: the valuable time of tutorials is not to be spent on general criticisms of the course.

Receiving Coldly a Logical but Not Well-Embedded Mathematical Argument. In Part a of Extract 8.5 the tutor elicits from the students the answer to the question ($\dim(X\cap Y)=n-2$) with very closed questioning. However, despite the fact that the answer is uttered by one of them, the argument is coldly received by the students. The reason, that Jack gives, is that the tutor's argument is 'just based on a formula we haven't proved'. In addition to that, I note that underlying the tutor's argument are two tacit theorems:

if $X, Y\leq V$ then $\dim(X+Y)\leq\dim V$, and,
if $U\leq V$ and $\dim U=\dim V$ then $U=V$.

The above theorems constitute necessary knowledge for the construction of the argument and in Part a the students do not give the impression that they possess this knowledge. These epistemological deficiencies coupled with the method of Socratic persuasion by very closed questioning result in the students' scepticism.

What follows in Part b is in striking contrast with Part a: the tutor deconstructs their belief that a basis of a vector space can be reduced to a basis for a subspace by removing some vectors from it via an illustrative counterexample.

The differential efficiency of the two arguments in Part a and Part b can be tentatively explained as follows:

- Firstly, the formula as well as the two tacit theorems in the argument of Part a are not well-embedded in the students' active knowledge. Therefore, while having very limited control of this knowledge, they are led by closed questioning to an answer that, in a sense, they are forced to accept.
- Secondly, the counterexample in Part b is a concrete refutation of their belief. Even though a perception of $\Re^{27}$ may not be graphically attainable, its elements can be conveniently thought of as 27-part strings of real numbers. Hence they are manipulable and the argument about the need for the elements of the basis for a subspace to be elements of the subspace can be more easily conveyed.

In a similar vein, the collective effort to prove the formula used in LA6.29ii turns out to be quite successful even though there is a great deal of leading by the tutor who sets up the scene by refuting the students' belief that a basis for Z can be reduced to a basis for X or Y(part b) and by replacing it with the new technique of extending a basis for a subspace to a basis for the vector space. The students' thinking initially revolves around a direct application of the technique on X until Jack alludes to the possibility of applying the technique on $X \cap Y$. Jack's intuition on $X \cap Y$ is prompted by Andrew who notes that Jack's suggestion to start from a basis for X does not provide good prospects for constructing a basis that contains a basis for Y.

Finally the proof is collectively completed and an imaginative suggestion is made by Jack who wonders whether the formula can be extended to 'more than two subspaces'. The exchange of arguments between Jack and Andrew is an exemplary case of interactive mathematical creativity even though the creations of this interaction in particular are rather tentative. I note that the tutor does not deter the students from experimenting on their idea. On the contrary he suggests the use of a concrete context ($\Re^{27}$) for the experimentation.

Note: there is a remarkable repetition of the misleading phrasing 'THE basis' by the students which is commented upon by the tutor. Similar evidence is available from other tutorials and the interviews.

Conclusion : In the above, evidence was given with regard to the efficiency of various ways of mathematical persuasion. Socratic closed questioning, in which unproved theorems were used tacitly, appeared to be less convincing than Refutation by Counterexample in which the examples vividly and concretely pointed out the argument. Moreover the efficiency of collective proving efforts was exemplified with regard to the strength of the students' conviction as well as their imaginative practising of generalisation.

Section (vi) Looking for the 'Usual' Basis of $P_3(\Re)$: Decontextualised Knowledge and the Ambiguous Nature of 1

Context: See Extract 8.6

Structure: In the following, the strictly tutor-led discussion is interspersed with the students' individual or collective responses that reflect a wide variety of their difficulties with the topic as well as evidence of some novice reasoning attitudes.

The Episode: A Factual Account. See Extract 8.6

An Interpretive Account: The Analysis

In the following I focus on:

- Patricia's and Beth's different responses to the almost identical teaching stimuli regarding the discovery of the basis $\{1, x, x^2, x^3\}$ for $P_3(\Re)$,
- the unanimous interpretation of $T(1)$ as 2 instead of 1, and
- the case of two students' imaginative suggestions for extensions of the knowledge discussed in the tutorial.

The Iteration of an Elicitation Process and its Different Outcomes. Both Patricia and Beth have difficulties with reconstructing the 'usual' basis for $P_3(\Re)$. Responding to the difficulty the tutor prompts them with a recommendation to think again 'what $p(x)$ looks like'. In both tutorials, the request for a general expression of $p(x)$ yields

$$p(x) = ax^3+bx^2+cx+d.$$

Patricia then 'sees the point' and notices (P2) that $p(x)$ is a linear combination of x^3, x^2, x and d ('the constant'). With further prompting she sees d as d-times-one. I note here that the students' difficulty in 'seeing' 1 as x^0 results in the phenomenon elaborated in the next part of this section (on $T(1)$).

Beth however seems to be in more difficulty than Patricia; it takes her longer to find the 'usual' basis. Her case seems to be more complicated anyway: at least Patricia, in her first attempt (P1) to find a basis for $P_3(\Re)$, suggested polynomials. Beth starts from 'matrices' as the tutor calls (1, 0, 0, 0) etc.. It is likely that Beth was misled by the tutor's persistent use of the term 'usual basis' (not specifying for which vector space) and suggested the 'usual basis' for $\Re^4$. However she suggests

$$p(x) = ax^3+bx^2+cx+d.$$

as the general form of a polynomial in $\Re^4$. Subsequently B1 is muttered and it is likely that Beth was trying to utter 'x^s, where s=1, 2, 3'. The tutor rather impatiently interrupts her in order to remind her of what they are looking for: 'simple polynomials that $p(x)$ is made out of'. B2 and B3 illustrate Beth's interpretation of the tutor's 'simplicity': she suggests linear polynomials, or in other words polynomials of degree one. Beth actually refers to the decomposition (factorisation) of a polynomial of the third degree to the product of three linear ones. In that sense B2 and B3 reflect how decontextualised Beth's thinking is and how the tutor's responses to her difficulty have failed to address the essential reasons behind this difficulty: Beth does not seem to have realised so far the rationale behind looking for a basis; or what is a basis and what kind of expression for the elements of a vector space it provides.

The tutor subsequently explains that it is not factorisation they are looking for but the construction of $p(x)$ as a linear combination of four polynomials. The tutor, presumably thinking that Beth may have a problem with the notion of linear combination, asks her what a linear combination is. B4 shows that Beth knows what a linear combination is — actually we already know that she knows from her earlier reply ($p(x) = ax^3+bx^2+cx+d$). So Beth's difficulty in seeing the 'usual basis' behind this linear combination cannot be attributed to the absence of prerequisite knowledge: there is evidence that Beth knows what a linear combination is and she also knows that she needs to break $p(x)$ into 'simple' polynomials. The fact that she suggests factorisation may reflect the possibility that she has not solidly consummated the idea that the operation between the polynomials in $P_3(\Re)$ is addition. So her difficulty to see the basis can be attributed to her lack of the ability to piece together these items of prerequisite knowledge in a meaningful way and in resonance with the needs of the problem she is asked to solve.

Finally Beth does see the correspondence between

$$ad_1+bd_2+cd_3+dd_4$$

and

$$ax^3+bx^2+cx+d$$

and correctly gives the 'usual basis' $\{1, x, x^2, x^2\}$ for $P_3(\Re)$. This however is not necessarily a fact which guarantees that she finally succeeded in co-ordinating the necessary pieces of knowledge. She yielded the answer intended by the tutor but it is questionable (there is no further evidence in Extract 8.6) whether she resolved the mystery of why her previous suggestion (factorisation) was not accepted.

The unanimous interpretation of T(1) as 2. Most of the students in this tutorial when asked to calculate $T(1)$, where $T(p(x))=p(x+1)$, reply '2'. As explained in B6, the students appear to be seeing T as an action according to which 'we are adding one'. Actually in her elliptic expression (B6) Beth captures the essence of her and her peers' difficulty: what is missing from her sentence is the object on which the action of the verb is suggesting: add one to what? Underlying the students' response is seeing 1 as the number on which to apply the action suggested by 'adding one' and not as

$$p(x) = 0x^3+0x^2+0x^1+1x^0 = 1$$

which therefore would help them see $T(1)$ as

$$T(p(x)) = p(x+1) = 0(x+1)^3+0(x+1)^2+0(x+1)^1+1(x+1)^0 = 1$$

Abidul's '$1+x$' is probably closer to the right answer: she seems to have thought:

I have to make x into x+1
I have no x, I have 1
I'll do this once, or leave it the same so x+1

If this is at all close to her thinking, then at least she has noticed the absence of x which the other students have not. In any case the students appear weak in interpreting functional information: they understand the action required by the rule of T but lack the crucial understanding of the objects which this action is to be applied on. The fragility of this understanding becomes evident in cases where the nature of these objects is ambiguous, such as 1 (a number? a polynomial?). In this case, it seems that what seems to be a routine algorithmic procedure (the application of T) is not routine at all: what this raises is the issue that notions like 'routine' and 'simplicity' are far from being unquestionably clear; especially when used in relation to the novices' learning and understanding.

Finally, I briefly note Cleo's and Camille's imaginative extensions of the discussion in this tutorial: Cleo hypothesises about another basis $(1, 1+x, (1+x)^2, (1+x)^3)$ and

Camille reverses the question tackled in this tutorial (find the matrix of a mapping T given a basis) to: given a matrix, a diagonal one, find suitable bases for the vector spaces between which T is defined. I also note that when the tutor and Camille go into the details of how Camille's suggestion can be realised, Cleo, unable to follow, stays behind. This differentiation raises several issues related to teaching; from a learning point of view however what is significant here is that the students reach the end of the tutorial with a capacity for extensions and generalisations.

Conclusion: In the above, the tutor's use of closed questioning results in the elicitation of responses from the students that reflect their difficulties with new algebraic concepts. Also algorithmic procedures, deemed by the tutor to be routine and simple, turn out to be problematic for the novices. While, for instance, looking for the 'usual' basis for $P_3(\mathfrak{R})$ the students either fail to respond to the task or make suggestions that are severely decontextualised and not in resonance with the needs of the task. Moreover the students unanimously appear weak in interpreting functional information (interpretation of $T(1)$ as 2). The fragility of their understanding seems to depend seriously on the ambiguity of some mathematical objects (1, for instance, as a number or as the constant function of value one). In contrast to this difficulty with 'simple' tasks, the novices at times appear capable of remarkable extensions and generalisations. This difficulty with 'simplicity' raises the issue of whether closed questioning allows an exploration (and service) of the students' actual cognitive needs or is simply a rigid implementation of the tutor's predetermined agenda.

Section (vii) Transforming 'beautifully literary' Intuitions Into Mathematical Formalism

Context:	See Extract 8.7
Structure:	This is a proof and refutation session in which the tutor twice pursues the students' suggestion which turn out to be successful. In this way, the tutor's act is to propel the transformation of the students' intuitive ideas into formal mathematical reasoning.
The Episode:	A Factual Account. See Extract 8.7

An Interpretive Account: The Analysis

Transforming 'beautifully literary' Intuitions Into Mathematical Formalism. The interaction of Cathy and the tutor — George's intervention is mathematically crucial but the psychologically rich aspect of the interaction involves mainly Cathy — can be deconstructed as a sequence of actions in which

- the novice has the necessary global grasp and intuitive ideas but is unable to organise them to an effective proof
- the expert organises these ideas in a demystifying process during which the novice learns about LA7.35 as well as about thinking in formalistic terms.

This is the theme of this Section and the analysis here aims at illuminating the details of the interaction from this point of view.

The focus of Extract 8.7 is psychological/affective. Cathy demonstrates a lack of confidence in the rigour of her thinking (not contented with her rigour in the beginning / 'solution then might be wrong' / refuses repeatedly to translate her intuitions to proofs / 'is this formal enough?' she says reluctantly later). These words by Cathy make the Extract 8.7 look like an attempt by the tutor to boost her confidence by flexing her cognitive muscles as well as encouraging her emotionally: Cathy does formalise when encouraged and when clearly introduced to the tools of mathematical formalism. The inseparability of cognition from the affective factors under whose influence it takes place here is evident.

The tutor's intentions seem to be to co-ordinate optimally the students' suggestions as well as act as an expert initiator into the codes of the mathematical culture (in particular as far as the mathematical reasoning used is concerned and as far as the rules of propriety that govern mathematical expression: logic and notation).

Expert Initiation Into the Codes of Mathematical Culture: Mathematical Expression and Notation. A priority in the tutor's agenda seems to be reforming Cathy's written expression of her ideas and definitions. Characteristically Extract 8.7 starts with two such attempts for reform: her writing for ImT and ImT^2.

For the first, Cathy has used x to denote an element of V as well as ImT on the grounds that both vectors 'come from V'. This can be misinterpreted as $x=Tx$ which is not what she aimed to say. Moreover her expression 'for v in V' can be misinterpreted as the universally quantified expression $\forall v \in V$, while she meant $\exists v \in V$. The tutor's critique leads her to more acceptable writing introducing her thus to a notational convention of formal mathematical writing.

Similarly, her writing of ImT^2 seems to be a symbolic representation of her rather illustrative idea that T^2 is 'applying T twice'. With the tutor's suggestion to think of T^2 as a transformation of V, her thinking is clarified and her reconceptualisation of T^2 leads to an acceptable expression. In this case the tutor intervenes in order, not merely to modify a problematic representation of a good intuitive idea, but also to channel Cathy's perception of T^2 to a direction which is more appropriate for the context of LA7.35.

Once the issue of Cathy's problematic symbolic representations of definitions has been settled, the focus of the interaction seems to shift to the transformation of her 'beautifully literary ideas' into mathematically formed argumentation. Cathy's 'ImT^2 has less vectors than ImT' is a rather inaccurate and finitist verbalisation of Im$T^2 \subseteq$ ImT — in fact in the realm of infinite sets cardinality and $\subseteq$ are two issues that cannot be treated adequately with colloquialisms such as 'less than'. For intuitive reasons that possibly relate to the simplicity of the definition of kernel — the elements of a vector space mapped to zero — Cathy finds the proof for ker$T \leq$ kerT^2 'easier to do formally'. As a result, C1 — despite its containing the same 'less than' expression as previously — turns out as a verbalisation of the heart of the argument. The student's question in the end of C1 — 'Is that formal enough?' — illustrates vividly the pressing sense of obligation towards mathematical formalism that the novices begin to feel.

The second time that the interaction between the students and the tutor takes the form of a transformation of intuitive verbalisations into formalism is with C2 and C3. Cathy seems to have grasped the idea that if ker$T$ contains an element other than zero, then the construction of another element is possible which is contained in kerT^2 and not ker$T$, thus contradicting b. It takes, however, the tutor's insight that

Cathy's idea is more than an unjustified guess — and thus it is worth pursuing — and George's contribution to transform Cathy's 'it seems true but...' into a proof for b⇒a. The tutor at first organises C3: let $v \in \mathrm{Im}T \cap \ker T$, $v \neq 0$. Cathy interprets that as $Tv=0$ and then as the correct but rather circular $T^2v=0$. The students' impasse, with regard to how $v \in \mathrm{Im}T$ can be exploited, is resolved by the tutor's reminder that $v \in \mathrm{Im}T$ is an 'existential' statement: $\exists w \in V$ such that $v=Tw$. Having this relational statement available, Cathy moves on swiftly to her next successful verbalisation, C4. The tutor, again in charge of conditioning the novices' expression within the boundaries of mathematical propriety, translates C4. This accelerates their arrival at the conclusion — which also is due to another tutor intervention with the theorem that, if two subspaces intersect trivially, then the dimension of the sum is the sum of the dimensions.

I note that, in the above, the students' weakness in dealing dynamically with the definitions of $\ker T$ and $\mathrm{Im}T$, despite Cathy's eventually successful writing of their definitions in the first part of the tutorial, reflects a possible gap between their relational and their instrumental understanding of the concepts.

Finally, C5 is another verbalisation of a suggestion for a proof (a⇒b). Its relevance to the novices' preferred approaches to mathematical reasoning makes it more appropriate to be mentioned in the following.

Expert Initiation Into the Codes of Mathematical Culture: Mathematical Reasoning and Proof. Unlike other cases, where extremely closed questioning seemed to stifle the novice's possibility of making connections, here the interaction, even though characterised by the tutor's dynamic interventions, is more prone to discovery than elicitation. The fruitful learning outcome of this approach is reflected in the students' success to reproduce the argument used for $\ker T \subseteq \ker T^2$ in their proof for $\mathrm{Im}T^2 \subseteq \mathrm{Im}T$; also in George's seeing the analogue between the proofs for b ⇒ c and c⇒b.

As far as the cyclic process of proving the equivalence of propositions a, b and c (a⇒b⇒c⇒a) — in contrast to evidence from other tutorials where the students found it difficult to come to terms with the logic of this cyclic process — Cathy's (and George's to a lesser extent) difficulties lay elsewhere: in the formalisation of their ideas for instance as explained above. George appeared unsure about where to start in the chain of proofs and how to continue but had a plan even though he did not manage to pursue it. The tutor himself attributes a great deal to an initial intuitive approach: b and c look the same so maybe they are of the same degree of difficulty. He then moves on to organising the information from the *Rank and Nullity Theorem* in a way that appears relevant to the data of the question.

Initiation into acceptable and viable ways of mathematical reasoning culminates in the part of the tutorial relating to the proof of a⇒b. In C5 ('go backwards') and C6 Cathy demonstrates an approach to mathematical arguments widely preferred by novices: contradiction by assuming the negative form of the argument to be proved. The tutor, for the first time in Extract 8.7, intervenes more dramatically with criticising Cathy's suggestion, not as incorrect but as unnecessarily negative and proceeds with his own suggestion. Since the tutor is not explicit about the reasons that make proof by contradiction less acceptable, it is questionable whether Cathy comes to any rationalisation of why her suggestion is not followed.

Finally T1 is strongly graphic evidence that the tutor is determined to intervene dynamically in the students' approach to argumentation: emotionally laden expressions (the impact of the word 'surely' in an appropriately firm tone of voice) do not support or reinforce a mathematical argument which ought to be grounded on proof.

Conclusion: In the above, the novice appeared to have the necessary global and intuitive grasp of a proof but, even though alerted to the requirements of formal expression, she was unable to organise her ideas effectively. The latter was propelled by the tutor in a demystifying process during which the student seemed to learn about the particular proof as well as about thinking in formalistic terms. The interaction between tutor and students was both cognitively and affectively intense. This highlighted the inseparability of cognition from the affective factors. The Episode stands as a metaphor of the novice's initiation into the codes of expression of the mathematical culture by the expert.

Section (viii) Leading Didactical Style as a Potential Propagator of Passive Learning. Resisting the Contingency of Multiple Answers to a Mathematical Question

Context:	See Extract 8.8
Structure:	The group of students featured in this section are particularly non-participant. Also the tutor uses closed questioning extensively. One of the implications is that the students' responses, even when correct, sound unnatural and unconvincing. In the following, the students appear rather inarticulate and the tutor leading.
The Episode:	A Factual Account. See Extract 8.8

An Interpretive Account: The Analysis

Throughout the observation and recording period I found it rather difficult to distinguish between cause and effect in the cases where the tutor used very leading, closed questioning and the students were passive and reluctant participants in these minimal dialectics. My impression however is that generally it is the tutors who set the tone and style of the sessions. It is an exceptional case (some of these cases are reported in this and other chapters) when a student, keen on dynamic dialectics, forces the tutor to adapt.

Here the students are not keen participants. Their responses to the tutor's closed questions sound more like a product of enforcement than conviction. In the following, I present the effects of this approach within the context of question B3.

How Patricia was led to realise the redundancy in her thinking. Patricia's inarticulate answers

- 'you can deduce the one from the other one' and
- 'can we just say that because E is linearly independent...?'

led me to believe that, even though the tutor prompted her to admit the redundancy of proving both conditions, she does not actually know why. In fact the second answer cited above might suggest Patricia believes that linear combinations of linearly independent vectors are linearly independent vectors.

The Success of a More Cognitively-Friendly Approach. In the second part of the question the tutor's strategy slightly changes. She appears as if she is trying to clarify first the students' knowledge of what a matrix of a map is-and-does and then ask them to apply this refreshed knowledge to finding the matrix in question. This proves more successful than her earlier strategy even though the students again rather mechanically transfer the application of the idea from bases E and F to bases E' and F. The contrast between their inarticulate utterances

- 'Te equals f. All the elements in it...'
- 'that's the matrix of the transformation'.
- 'The coefficients of the sums...f_1 plus f_2...'

and the fact that eventually the students dictate the necessary calculations for the evaluation of Te_i is spectacular. The students appear to become more easily better executioners of algorithmic plans than interpreters of the role and significance of the concepts they are introduced to.

Complications Resulting From Multiplicity of Answers. In B3 it is $Te'_3=0$ with respect to the matrix given in the question. So while calculating Te'_1 and Te'_2 has led to finding vectors f'_1 and f'_2 for basis F' of W, Te'_3 has not. Zero cannot be in a set of linearly independent vectors. The two students have tackled this complication in different ways. Beth decided that since $Te'_3=0$ and we cannot have zero, the basis will contain only two elements (B1). Cary on the other hand, convinced that a basis of W must necessarily have three elements, decided to keep zero in the basis (C1).

With the tutor's prompting, Cary notes that, if we include zero in the basis, then our set of vectors ceases to be linearly independent (C2). The tutor wants the students to understand that they have to find f'_3 replacing zero with another vector, not necessarily via the same method that led them to find f'_1 and f'_2. Beth however seems at ease with the idea that since the process determined these two acceptable vectors only, the dimension is 2 (B2). Finally through B3 to B8 she is led to change her mind into the multiplicity of possibilities for f'_3 cannot be zero.

In the above, the students appear to resist the idea of determining f'_3 in a non-unique way. f'_3 can be any element of W as long as it is linearly independent of f'_1 and f'_2. This potential pluralism of answers, while a common mathematical occurrence (examples: there is often more than one solution in a differential equation; there is an infinite number of solutions in an under-defined system of simultaneous equations) seems to be alien to their mathematical experiences and hence they resist it. The expectation of singular answers to mathematical problems, patronisingly designed to present mathematics as a polished, clear-cut deterministic activity, is here revealed to be a detrimental, die-hard habit. I note that the rest of the students in this college, that were confronted with the possibility of various choices for f_3, responded in similar ways.

Conclusion: In the above, a case was explored where the novices seem to resist the potential pluralism of answers to a mathematical problem. Moreover in the process a gap was observed between the students' calculating skills and their ability to reason articulately about their mathematical actions. In terms of teaching style the highly leading and predetermined closed questioning was illustrated as inadequate and unnatural, whereas a more cognitively friendly exploratory style was proved relatively more successful. It is likely that such a teaching approach deepens the above mentioned gap between algorithmic and conceptual understanding. Both passivity and reluctance to participate as well as resistance to the potential of mathematical pluralism reflect a deeply unadventurous and conservative learning style.

PART III A Synthesis of the Findings in the Area of Linear Algebra. Indications for the Cross-Topical Synthesis in Chapter 10

In this chapter the students' first experiences of Linear Algebra concepts were explored in a series of Episodes from the first and the beginning of the second term of observation. As noted in the Interlude, the concept that emerged as *paradigmatically problematic* in the area of Linear Algebra was the concept of *span /spanning set* which was explored from a linguistic and a geometric/visual perspective: the novices [1] often seem to miss the grammatical link between the terms *spanning set* and *span* — which reflects their cause-and-effect conceptual link. This linguistic deficiency mirrors and partly determines the novices' restricted understanding of these new concepts[0]. Specifically, their understanding seems to be also influenced by their responses to the requirements of generalisation. Within the same mathematical task [1][1],

- some students learned to reproduce the strategy, suggested by the tutor, of generating $<v_1, v_2>$ from $<v_1>$, when $v_2 \notin <v_1>$, and $<v_1, v_2, v_3>$ from $<v_1, v_2>$, when $v_3 \notin <v_1, v_2>$. Some did not seem to accommodate the acculturating process of learning and applying the rule as readily as others.
- one of the students seemed to perceive $<v_1, v_2>$ as $<v_1>+<v_2>$. This implies that the student resists the idea of a construction larger than its parts and it also resonates with findings in other contexts, for instance functions, on the persistence of linearity in the novices' thought processes.

So the students appear as if they respond to the tutor's stimuli for generalisation with various degrees of readiness. On what seemed initially to be a paradox, among the observed students, the only one who attempted to generalise with some facility, was the one who complained about the difficulty of generalisation. Consciousness of difficulty was then conjectured as enhancing the possibility of overcoming difficulty.

Tutors and students also seem to perceive routine-ness and simplicity of task differently: for instance [6], when looking for the 'usual' basis for $P_3(\Re)$, the students either failed to respond or made severely decontextualised suggestions (for instance

---

[0] Similar semantic and linguistic interpretations were attached to the novices' confusion of $\varnothing$ and $\{0\}$.

[1] the discussion of $<v_1>$, $<v_1, v_2>$ and $<v_1, v_2, v_3>$

factorisation). Moreover the students unanimously appeared weak in interpreting functional information (e.g. in the interpretation of $T(1)$ as 2). The fragility of their understanding seems to depend seriously on the ambiguity of some mathematical objects (1, for instance, as a number or as the constant function of value one). These differences in perception make it necessary to reconsider what is traditionally thought of as a simple task. Especially when at the same time that the students appear incapable to perform a 'simple' task, they are capable of remarkable extensions and generalisations.

In their attempts to understand the vectorial context, the novices seemed to adhere strongly to the metaphor of the plane which they frequently interpreted rather literally: they use strictly geometrical language regarding, for instance, vector addition. This attitude was vividly illustrated in a students' suggestion [1] to add vectors by 'dropping perpendiculars', that is by using a Cartesian orthogonal system, instead of using the Parallelogram Rule: as a result, two different aspects of the geometrical metaphor of the plane (Cartesian orthogonal system, Parallelogram Rule) interfered with each other and with the novice's understanding of the construction of $<v_1, v_2>$. Other illustrations of this attitude include the students' body language as well as their orientation on the plane [1] (minus for left, plus for right).

Novices tend to adapt metaphors in an excessively literal manner mainly because the suggested metaphor — the plane in this case [1, 2, 3] — has been associated with convenient and familiar algorithms. The likelihood of this explanation is reinforced by the students' tendency to apply their algorithmic competence even within contexts they do not have a good grasp of: e.g.[2] applying the Subspace Test on a subset of $\mathfrak{R}^{\mathfrak{R}}$ whose contents they are not aware of. Soon this action-in-void brings their conceptual difficulties to the surface: difficulties with the zero element of $\mathfrak{R}^{\mathfrak{R}}$ and confusion of f with f(x) [2].

Some students however try to avoid this meaningless action, when trapped in a context they are not aware of, and attempt to accommodate the new concept in what they already know [2]: for instance, while struggling with seeing the zero function as the zero element of $\mathfrak{R}^{\mathfrak{R}}$, a student attempted a meaningful construction of $\mathfrak{R}^{\mathfrak{R}}$. In the process it is revealed how her concept image of a vector on the plane (a directed line segment) interfered with the notion of vector as an element of a vector space; also her weakness in perceiving function as an object-element of a set.

Similar struggles with constructing a meaning for the newly introduced concept of spanning set (the difficulty with the generating aspects of the spanning process seems to characterise these struggles) mostly consisted of efforts to employ familiar metaphors, for instance from the study of functions. In one instance [3] — characteristic of the tendencies among the novices — the student exaggerated the tutor's persistent recommendation to use the metaphor of the plane and appeared to believe that every co-ordinate system is a spanning set, including polar co-ordinates. In the process she also appeared to see basis as the smallest spanning set and, as her consistent use of the article 'the' for spanning sets indicates, she seemed to believe that a spanning set and a basis are unique. This contradicts her image of basis as the smallest spanning set which implies the existence of more than one spanning sets. The two images seemed to co-exist independently. Finally the student wondered about the utility of spanning sets. This reflects an inclination to search for a purpose in the introduction of new concepts[2] whose usefulness is not readily visible to the novices.

Apart from wondering about the utility of the new concepts the students also seem to acquire a variety of explicit or implicit images related to the concepts that co-exist, or exist despite of, the concept definition. Examples:

- the elements of a basis of a subspace do not necessarily belong to the subspace [4],
- a basis of a vector space can be reduced to a basis of a subspace [5].

Along with these sometimes inadequate or contradictory concept images, the students' inadequate knowledge [7] of necessary definitions (ker$T$, ImT) and the linguistic and graphical inconsistencies associated with these concepts also becomes an obstacle to building up meaning for the new concepts as well as proving arguments. Trying to work within $\Re^{\Re}$, while thinking of the concept of function disconnected from relevant notions such as domain and range [2], is an example of the role played by these obstacles.

In the above, the novices appear to respond in a variety of ways to the introduction of new concepts: by regressing to the adoption of familiar metaphors (substituting thus the power of the abstraction in the new concepts with the convenience of a familiar context[3]); by concentrating on a competent if narrow minded execution of algorithmic tasks; by engaging in a struggle for a meaningful construction of the new concept. For instance[4] while discussing the new concept of linear mapping

[2] reported extensively also in Chapters 7 (limit)and Chapter 9 (cosets)

[3] I note that in other contexts, such as limits (Chapter 7), the novices were less prone to resort to previously established techniques.

[4] Condensed evidence from a number of episodes not reported in the Chapter.

some students adhered to an application of the definition, even when they could not give an example of the concept; others tried to critically embed the new concept in their previous knowledge (for instance, by asking whether particular examples of mappings they know about fit in the definition of linearity). In understanding the notion of linear mapping the novices' familiarity with the metaphor of the two-dimensional plane seemed to play a positive role. This specificity of context however seemed to have a more controversial effect in cases like finding the zero element of a vector space, where the novices associate the zero vector with the number zero. So, in this case, specificity of context impedes the novice's understanding of the zero vector as, for instance, a function (also in [3]) or a matrix.

The students also appear to be in difficulty to express formally an intuitively grasped idea regarding spanning sets [1, 4, 7, 8], even in the cases where they appear alerted to the necessity of a formal argument [7]. One of the students' linguistic unease [1] — because English is her second language — was illustrated as a metaphor for the difficulties engendered by mathematics as a symbolic and formal language too. The extreme formalistic nature of some of the new notation, for instance the unnecessarily convoluted use of *Kronecker's Delta* in [4] also seems to obstruct the development of a student's problem-solving thinking. In this case, the difference between the expert's facility with formalism and the novice's unease is so obvious that while the expert and the novices may agree [4] about the method of approaching a problem, they differ in terms of the implementation of the approach: as a result, their interaction evolves into an initiation process during which the students, with variable ease, become familiar with the new notational tools of mathematical formalism. Hence their learning becomes a specific struggle for accommodating to this new tool — whose appearance maybe intimidating — the vivid intuitive ideas they have about the solution of the problem.

In the above, the students' learning is described in terms of their interaction with an expert. In this interaction, the efficiency of the different ways of mathematical persuasion varies [5]: Socratic closed questioning, in which unproved theorems were used tacitly, appeared to be less convincing for the novices than Refutation by Counterexample, in which the examples vividly and concretely pointed out the argument. In the latter case the strength of the students' conviction as well as their imaginative practising of generalisation were exemplified [5]. At the same time the logic of implication and of proof by contradiction are reasoning techniques that the novices appear to struggle to master [7].

The students' interaction with their tutor was in most cases described as an enculturation process in which the tutor organises the students' intuitive ideas in a demystifying manner and in which the novice seems to learn about particular proofs as well as about thinking in formalistic terms. In some cases [7] the interaction between tutor and students is both cognitively and affectively quite intense and in a manner which highlights the inseparability of cognition from the affective factors under whose influence it takes place.

Chapter 9

The Novices' Encounter With Mathematical Abstraction: Cases from *Group Theory*

PART I A Guide to the Paradigmatical Cases (Episodes) Presented in this Chapter

The following table contains contextual information with regard to the 8 Episodes presented in this Chapter.

Episode Number	Time of Incident Term - Week	Participants	Mathematical Content
1	Hilary 5	Connie	GRF5.1, order of a group and order of an element, cyclic groups, *Lagrange's Theorem*
2	Hilary 6	Camille	equivalence classes, cosets
3	Hilary 6	Jack and Andrew	GRF5.6, order of a group and order of an element, cyclic groups, *Lagrange's Theorem*
4	Hilary 6	Connie	centralisers, conjugates of an element in a group
5	Hilary 6	Jack and Andrew	GRF5.8d, order of a group and order of an element, subgroups, *Lagrange's Theorem*
6	Hilary 7	four pairs of students in one college	*First Isomorphism Theorem for Groups*
7	Hilary 7	Connie	GRF7.3, normal subgroups, cosets, *Lagrange's Theorem*
8	Hilary 8	Cathy and Cliff	GRF8.5, order of an element in a group

PART II Data and Analysis

In the following I present the factual and interpretive accounts and conclusions for the 8 Episodes of the table in the previous page. In Part III then I synthesise the findings of Part II related to Analysis and briefly discuss the wider cognitive issues that are presented in the overall synthesis of the data analysis in Chapter 10.

Section (i) A Gradually Revealing Example of the Linguistic and Conceptual Dimensions of Difficulty With *Order of an Element*, Generating $\langle g\rangle$ and the Group Operation

Context:	See Extract 9.1
Structure:	At first the discussion — see Context — is triggered by Connie's problematic understanding of GRF5.1. However in the Extract it gradually becomes evident that Connie is in confusion about the notion of order (of group G or element g) and of the group operation.
The Episode:	A Factual Account. See Extract 9.1

An Interpretive Account: The Analysis

This episode stands as a metaphor for a cognitive journey back to the roots of the student's confusion with the new group theoretical concepts she has been introduced to: ostensibly about GRF5.1, this is mostly about $\langle g\rangle$, $|g|$ and the group operation.

A Problematic Use of the Term Order of an Element. As the tutor notes (his observation is based on Connie's reaction to Alan's solution in the group tutorial that morning) Connie lacks a clear understanding of the notion of order of an element. He then chooses to re-introduce her to the notion, via the definition given in the lectures, and explain that $|g|=|\langle g\rangle|$. In C1 and C2 Connie seems to be slightly surprised with the connection between $|g|$ and $\langle g\rangle$.In fact her confusion and surprise can be attributed to the linguistic use of the term *order of an element g* which actually is an abbreviation of the more accurate *order of the group generated by an element g*. C4 illustrates how Connie is in trouble with understanding how an element can have an order — which so far has been identified as a property of groups: groups have *order* and the order of a group is a number equal to the number of its elements. Connie's confusion can be accounted for as an effect of her not realising the tacit abbreviating of the term. C3 is evidence of her seeing *order* as the number of elements in a group and actually as a finite number: she uses the word 'count' in order to talk about the act of finding the order of a group. I note that her finitist conception of *order* is perfectly justifiable in the context of finite groups she has been recently working in.

Finding however the order of a cyclic group by counting the number of its elements obscures the fact that if $\langle g\rangle$ is of order p that means that it has p elements *because* the powers of g start 'repeating themselves' after p. In this sense *order of an element* is a concept that contains both a static characteristic of $\langle g\rangle$ (the number of its elements) AND information about the process of obtaining these elements (how many times it is necessary to take the powers of g in order to cover <g>). In Connie's words, most strikingly C4, this duality is missing.

Generating <g> and the Notion of Group Operation. Subsequently in C5 to C9, it turns out that behind Connie's unease with the concept of *order* lies the even more fundamental unease with the notion of generating a cyclic group from the powers of an element $g \in G$. Even further her troubled notion of generating can be largely attributed to her muddled perception of the operation in a group.

In C5-C8 'times' and 'to the power of' are used unclearly interchangeably; it is also not clear at all in each one of them whether Connie refers to the 'multiples' or 'powers' of g or g^2. This can be due to the tutor's effort to convey the idea that, in a subgroup of order p, where p is a prime number, every element, other than e, is of order p too. Therefore, whether we consider the powers of g or g^2, eventually the generated set will be H. In this sense the tutor and Connie are far from understanding each other and communicating fruitfully: Connie is still struggling with her exploration of the notion of generating a group from an element; she is still vague about how this process takes place (operational stage). The tutor on the other hand assumes the clarity of the process of generating a group from g and a group from g^2 and attempts a demonstration of how these two groups coincide.

It seems fair to say that the discourse between the two interlocutors takes place past each other. Perhaps most striking is the exchange of words in C5 to C8 with regard to 'times' and 'to the [power of]': in C5 Connie explicitly uses 'times' and in T2 the tutor responds with taking powers. In C6 and C7 she insists on 'times' and the tutor shifts from 'multiplying' elements that are powers of g to manipulating the powers to which these elements are taken. While doing so he seems to assume the clarity of these operations. C9 is evidence of how Connie, far into the discussion, is still struggling with clarifying the objects on which the operations are applied. Since the conversation is completed with C10 — it seems that the tutor has been expecting a verbal signal of understanding from Connie so that he can move on to other topics — there is no evidence of whether Connie's perception of the group operation, the generating process and ultimately cyclic groups and their order has been clarified and enriched.

Conclusion: In the above, a novice's problematic perception of $|g|$, <g> and the group operation were gradually revealed. Linguistic (abbreviated use of the term) and conceptual (static and operational duality) interpretations of the student's difficulty with the notion of order of an element have been given. Further, the operation of generating <*g*> seemed to be problematically perceived by the novice who uses metaphorical expressions like 'times' and 'powers of' very unclearly and at times interchangeably. Didactically the dialectics between tutor and student illustrate a communicational gap which leaves the question, whether the student's perceptions have been enriched, pending.

Section (ii) A Novice's Inquiry on the Concept of Equivalence Class and of Coset: Bestowing Meaning Through Ambivalent Uses of Geometrical Metaphors

Context:	See Extract 9.2
Structure:	In the following, Camille enquires the tutor about the ideas she has found problematic in the lectures: the proof of one of the isomorphism theorems and then the definitions of centraliser and conjugate. Underlying her questions seems to be her difficulty with the new notions of equivalence class and, most severely, with the notion of coset.
The Episode:	A Factual Account. See Extract 9.2

An Interpretive Account: The Analysis

I note that the boxed parts of the discussion are the three highlights of the analysis.

Literal Interpretation of a Drawing I: Equivalence Classes as Straight Lines. Fig.2a is the tutor's visual representation of f, a mapping of G on itself. In order to avoid a diagram in which G and Imf would be separated (a misleading idea since Im$f \subseteq G$), the tutor prefers to represent an element a of the domain as a dot and its equivalence class (generally defined as the set of elements in the domain that are mapped on the same value as a) as a line segment. This metaphorical representation however seems to escape Camille who interprets fig.2a literally and wonders (C1) why equivalence classes are straight lines. T1 and fig.2b are attempts to set the record straight and emphasise the metaphorical nature of the representation. Soon, however, Camille seems to repeat analogous interpretations with regard to the notion of coset (C10 onwards).

C1 and C2 reveal Camille's preoccupation with the notion of an equivalence class which extends later, more intensely, to the relevant notion of a coset. C3-C5 however reflect her puzzlement with the main aspect of the isomorphism theorem (there is an isomorphism between the elements of a group and their equivalence classes). I note that, in this Chapter, the students are repeatedly unsettled by the idea of

> defining a mapping between two groups (or on a group itself),
> then defining a new relation between sets of elements of these groups and
> then defining a type of morphism
> between these sets or

between elements of the group and sets of elements

(for instance Extract 9.6). Here C3 is a sign of this confusion. Camille quotes the lecturer's definition of g ($g(e_a)=f(a)$) and then claims 'we don't know what g is'. It seems that commonly used phrasing such as 'define a correspondence between the elements of a group and their equivalence classes' is not perceived by the novice as a clear establishment of a function; or in C3 Camille does not see the = sign as a sign of definition but as a sign of equality. Her confusion then is the outcome of knowing what lies on the right hand of the equality and not knowing what lies on its left.

Moreover Camille's confusion can be justified on another basis: the tutor's expression 'a correspondence between the elements of a group and their equivalence classes' is equivalent but not identical to the lecturer's $g(e_a)=f(a)$. What the lecturer seems to have said is the following:

> *I define e_a as the set of elements x of the group for which $f(x)=f(a)$ where f is a homomorphism from G to G. I then define g as the correspondence that assigns e_a to this common value $f(a)$.*

This is more specific than the tutor's expression and there is not much to guarantee that a novice necessarily ought to see that the two expressions in essence coincide. The tutor's subsequent explanations clarify the definition of g, as well as its properties, but it is noticeable that it is due to Camille's persistence that these clarifications are finally being uttered.

Literal Interpretation of a Drawing II: Cosets as Squares. A Multi-Faceted Tentative Construction of a Meaningful Image of the Concept of Coset. The discussion of the correspondence between the elements of a group and their equivalence classes evokes in Camille a query on another correspondence: 'the 1-1 correspondence between the conjugates of x and x''. Remarkably Camille demonstrates precise knowledge of the relevant definitions (centraliser, conjugate) as well as a relation between the two concepts. I note that, unlike Camille, most of the students in these tutorials at this stage were incapable of reproducing definitions of even simpler group-theoretical constructs mentioned in the lectures.

However Camille in her demonstration of knowledge has not used the term *coset* at all. The term occurs for the first time in the tutor's words and captures Camille's attention. Subsequently and in the rest of the Episode it seems that the notion of *coset* constitutes a large part of her preoccupation : C6-C12 seem to be persistent, multiple attempts to imbue it with some meaning. C6 comes through as a surprisingly philosophical and abstract question which raises a very fundamental existential issue with regard to the notion of coset: what is surprising about C6 is that it comes in the middle of the tutor's describing a quite sophisticated construction (establishing a correspondence between the cosets of the centraliser and the conjugates of an element x in a group) and shifts the conversation from the strictly and specifically mathematical (represented by the tutor) to the metamathematical. Camille has been attentively listening to the tutor's demonstration of the construction and has given the very strong impression that, throughout, she has been processing the dense information provided by the tutor. C6 however illustrates that this processing must have been motivated mostly by the desire to construct a representation of coset — visual, 'material' — than consume the tutor's argument. From then on, as said earlier, C6-C12 is a series of successive attempts at interpreting the concept of coset.

C6 is a nearly platonistic enquiry on the nature of cosets as objects, as entities. Camille's entities in C6 do not necessarily act or interact. In C7 the questioning of the nature of these objects takes the form of an exploration of their *raison-d'-être* (very similar to Connie's enquiry on *conjugates* in Extract 9.4). C8 is a dissection of a coset which equates a coset with how it comes into existence. I note that so far T4-T6 do not seem to have a direct impact on the genesis of Camille's ideas of what a *coset* is. C9 is a geometrical interpretation of C8 derived from the notion (and notation for) transformations, and in particular translations. The tutor carefully tunes in (T7) but Camille accelerates her tentative condensation of her conception of coset in a geometrical image in a questionable way: C10 (in parallel with C1) illustrate how the line between a metaphorical and a literal interpretation of a picture is thin and severely disguised under the heavy weight impact of visual imagery. The tutor is surprised and alarmed (T8) by Camille's intention to 'apply [this idea] on squares'. C11 is evidence that Camille is too preoccupied with her image construction to be influenced by T8 and she furthers the interpretation of her fig.2c in a less controversial but highly ambivalent way. T9 is one more effort on the tutor's side to tune in and transform the student's images from within. Surprisingly then Camille turns in a shift to a more abstract property of cosets in which however the geometrical jargon ('size' in C12) is maintained. The tutor (T10) has completely adopted Camille's metaphor and contributes another observation on cosets.

Finally Camille ceases the effort to interpret further the notion of *coset* once she acquires an image of *cosets* that is satisfying and clear to her. That Camille is content with what she has acquired can be assumed on the basis of the evidence, given during observation, that this student does not bring a conversation to an end until she acquires a satisfactory (to her) understanding. The issue that C6-C12 raise is whether the quality of the acquired perception of a coset — via a multiplicity of metaphors and visual representations — justifies Camille's eventual sense of content. Given that the tutor cautiously surrenders in adopting Camille's metaphor but does not cross-check whether the intended (by the tutor) and the acquired (by Camille) image of a *coset* coincide, the questions raised by this issue ought to remain open.

Conclusion: In the above, a student, who exhibits a remarkable knowledge of the definitions of the concepts involved in the discussion, is engaging in a meaning bestowing process with regard to the notions of equivalence class and of coset. The student asks the tutor about the *raison-d'-être* of the concepts and her efforts are characterised by a tendency to use metaphors of some regular geometrical shapes in order to construct a mental representation of the concepts (equivalence classes as *straight lines*, cosets as *squares*). Evidence was given that these geometrical representations are interpreted literally by the student. This raises the issue of a potential cognitive danger built in the use of geometrical metaphors.

Moreover the novice's difficulties were identified with regard to conceptualising a mapping between elements of a group and sets of elements of the group. This was seen in two examples of such mappings (involving the concepts of conjugate, centraliser and equivalence class). The tutor has demonstrated a considerable flexibility in thinking in the terms of the student's metaphors (actually it is the tutor who sparks off the use of geometrical representation in this tutorial) but in the end there doesn't seem to exist any guarantee that the didactical use of metaphorical discourse has resulted in the tutor's intended concept image of the notion of coset.

Section (iii) A Contrast Between Expert and Novice Approaches to Proof: The Fine Details of a Lateralisation of Cases in GRF5.6

Context: See Extract 9.3

Structure: In the following, the student seems to have acquired an adequate grasp of the mathematical argument in GRF5.6 but lags behind in terms of an appropriately formal and complete presentation. It is debatable whether in the interaction with the tutor, the finesse of the argument escapes Jack, or it is simply a matter of emphasis and clarity in the student's words.

The Episode: A Factual Account. See Extract 9.3

An Interpretive Account: The Analysis

The Details of the Argument for GRF5.6 that Sparked the Controversy. As illustrated in fig.3, the argument for GRF5.6 (Prove that a group of order 35 contains elements of order 5 and of order 7) presented in the tutorial is as follows:

> G is either cyclic or not cyclic.
> If G is cyclic then $\exists x \in G$ such that $G=\langle x \rangle$ and $|x^5|=7$ and $|x^7|=5$.
> If G is not cyclic then there does not exist $x \in G$ such that $G=\langle x \rangle$, therefore there does not exist $x \in G$ such that $|x|=35$. Hence $\forall g \in G$, $|g|=1$ or 5 or 7. Let *gŒG. Then* $|g|=1$ or 5 or 7. If $|g|=5$ then, from GRF5.1, the number of elements in G of order 5 is a multiple of 5-1=4. But, since $|e|=1$ and 4 does not divide 35-1=34, G contains a number of elements of order 5, a number which is a multiple of 4, and the rest, since there is no element of order 35 and the only element of order 1 is e, must be of order 7. The same conclusion would have been reached if g was assumed in the beginning to be of order 7.

Jack's argument (J1) grasps the essential part of proving that, in all cases, G contains elements of order 5 and 7, but is not clear in his presentation of the cases he takes. However, as the tutor notes (T1 and T2), selecting an element of order 5 and proving that the rest of the elements in G are of order 5 or 7, by not excluding the possibility that among them there maybe an element of order 35, is not legitimate. In Jack's mind this exclusion seems to have taken place in the beginning, when he let $x \in G$ and lateralised the cases for $|x|$. But this is different from what the tutor says: choosing an element in terms of its order says nothing about the order of the rest of the elements in G. Jack proves that, in case $|x|=35$, then G is cyclic; therefore it contains elements of order 5 (x^7) and of order 7 (x^5). When he chooses $|x|=5$, he does not exclude the possibility of another $y \in G$ whose order is neither 5 nor 7 and $g \neq e$.

In a sense, the presentation of the argument as in fig.3 or above is less controversial and clearer than Jack's. It seems that the exclusion process described above has taken place in Jack's mind (J2 can be taken as a declaration of his taking mutually exclusive cases and that once a case has been looked at, for instance $|x|=35$, it is not necessary that it reappears in the proof) but is not emphasised enough in his presentation.

Lateralising cases in terms of G being or not being cyclic, is, as the tutor's argument shows in T1 and T2, a more unambiguous approach. Jack's however lateralisation seems more natural: the question asks for a proof that a group of order 35 contains elements of order 5 and elements of order 7. So Jack sets out to point these elements out. This forces him to take cases. In the course of his checking out all the possibilities, if he were more precise, he would have to distinguish between G being or not a cyclic group. So the distinction, if we follow Jack's train of thought, is born out of the argument in his proof as a necessity; it is deeply incidental.

In contrast, the tutor's argument, as succinctly put in T2, suggests that the proof should start with making the distinction between the case where G is cyclic and G is not cyclic. This fine, logical suggestion however is the benefit of hindsight: knowing that subsequently the proof requires a lateralisation of cases, the tutor suggests that this can be neatly done in the beginning. There doesn't seem to be any psychological reason why someone would start this proof from distinguishing between cyclic and non-cyclic groups.

The above juxtaposition illustrates the contrast between expert and novice approaches: the former logically economical, succinct and benefiting from hindsight; the latter naturalistic and exploratory. Jack in the end doesn't appear as if he is clear about the difference between his approach and the tutor's. In fact the tutor, who calls this difference a 'minor point', is not emphasising the logical underpinnings of his and of Jack's approach: he seems to be more preoccupied with the full coverage of the cases in a formal way (he insists that Jack is clear that, when he deduces the existence of an element of order 7 from the existence of an element of order 5, he has excluded the possibility of the existence of an element of order 35) and less with the naturality in the genesis of the proving argument.

Conclusion: In the above, an argument whose finesse seems to be not fully grasped or not clearly emphasised by the student, sparks off an elaborate conversation of the details of lateralisation of cases in GRF5.6. In the process a contrast becomes visible between the genesis of arguments by an expert (logical, succinct and benefiting from hindsight) and novice (naturalistically born out of the proof).

Section (iv) A Novice's Struggle for a Meaningful Interpretation of the Definitions of Centraliser and Conjugacy Class: Request For Examples and For a Teleological Rationale Behind the Definitions

Context: See Extract 9.4

Structure: In the following, evidence is given of Connie's effort to imbue some meaning and purpose to the new concepts of centraliser of a group and conjugacy class of an element.

The Episode: A Factual Account. See Extract 9.4

An Interpretive Account: The Analysis

The Episode is triggered by C1, Connie's frustrated declaration of non-understanding. I note the emphasis she has put on 'really' as well as her phrasing: 'what they are': this is a purely existential statement which deviates considerably from the standard approach, of the lectures for instance, which is to familiarise with a new concept through its definition. The tutor responds with the definition of $C(x)$ and C2 is a comment on the definition. I note that in C2 as well as in C7 Connie's use of the words 'swap back' and 'sends it back' possibly reflect an action-dominated concept image of the group operation. In $xy=yx$, 'swapping back' is the result of the commutativity of the particular x and y. And in C7 the inverse x^{-1} is 'sending x back to itself'. This latter verbalisation can perhaps be associated with the action perceived aspect of the inverse function (if f sends x to $f(x)$, then f^{-1} sends $f(x)$ back to x, so f^{-1} is a way of coming 'back' to x).

With regard to the definition of centraliser the symmetry of the expression $xy=yx$ seems to engender the false impression that $C(x)$ and $C(y)$ are the same. I think Connie here is confused with the quantifiers behind x and y: in the definition of $C(x)$, x is fixed and y runs through G. Some of the y, the ones that commute with x, are elements on $C(x)$. The tutor (T1) illustrates that $C(y)$ is the set of all the elements that commute with y, whereas $C(x)$ is the set of all the elements of the group that commute with x. In a sense the verbal presentation of the definition of $C(x)$ seems to be less prone to the engendering of the false impression because it puts the emphasis on the fixed nature of x and the variable nature of y.

Giving the definition of a centraliser however does not satisfy Connie who in C3 emphatically requests some examples 'so that I can understand'. It is a rare occasion in these tutorials that the novices turn up with such eloquent and clear cut declarations of what helps them understand. In this case, Connie makes a statement about the value of examples in her constructing a meaning for a new concept. The tutor's example of an Abelian group is slightly disappointing because it does not enrich the concept image of a centraliser as a new entity; it possibly trivialises it. With the second example however Connie, who also corrects the tutor's claim that $C((12))$ is $\{e\}$ and suggests $(12) \in C((12))$, participates in the construction of $C((12))$ and requests no further explanations on the concept of centraliser. It may be reasonable to assume that at least temporarily she is content with her newly acquired image of the concept.

C4 and C5 reflect Connie's preoccupation with the role of g in

$$C_x = \{ gxg^{-1}, g \in G \}$$

I note that in both C4 and C5 it is not clear whether Connie is talking about C_g or C_x. The tutor seems to shift from one to the other adapting each time to Connie's words but never pointing at his shift explicitly. His clearest statement is the process-oriented T3 in which he defines C_x by suggesting a way to construct it: run through all $g \in G$ and construct gxg^{-1}. The set of gxg^{-1} is C_x.

C6 is evidence of the impact that the quasi-algorithmic T3 had: Connie realises that the construction of C_x is an x-centred action and also that running through $g \in G$ may generate more than one conjugate for x. I note that Connie's confusion with the concept of conjugate seems to be underlain by the same confusion with regard to the notion of fixed and variable element as in the definition of centraliser. 'how do you choose g'', 'is it for every g' in C4 and C5 as well as C6 possibly reflect an interplay in her mind between the fixed and the variable, between the single- and the multi-valued. In C6 she seems to have accepted the idea of the multi-valuedness of the conjugates of x: 'x has several conjugates'.

Symmetrically to the definition of centraliser, the tutor proceeds with examples (one trivial, one not) — this time without Connie asking for them. Unfortunately using examples here is not as successful as with the notion of centraliser, because Connie finds the calculations for $C(12)$ too complicated. Still concentrating on trying to understand what a conjugate is, she abandons the example and returns to a more theoretical exploration of the concept (C7 and C8). C7 reveals that underlying her questioning about conjugates is her confusing them with inverses. To her the *raison-d'-être* of an inverse is that it 'sends x back to itself' (see comments above). However despite the presence of an inverse in the definition of conjugate she does not see what a conjugate actually does, what it is for and in what sense it differs from an inverse. In C8 she explicitly demands a justification for the introduction of conjugates. C8 encapsulates, like C3, another learning problem that the novices' seem to be preoccupied with in various mathematical contexts: the absence of a teleological rationale in the introduction of most new concepts.

The tutor's reaction (he notes that she is 'leaping ahead') to her request for justifying the purpose of introducing conjugates is typical of the dominant approach in university mathematics according to which concepts are introduced arbitrarily and gradually begin to make sense as organic parts of the mathematical discourse. Here, for instance, conjugates will begin to make sense when normal subgroups are introduced. For the tutor, Connie is 'leaping ahead' because defining a concept logically precedes justifying it. Psychologically however things are probably different: a large part of Connie's preoccupation with the new concept seems to be constructing a meaning of conjugate that encompasses a reason for its existence and not necessarily a fully-blown expression of the definition. In this sense, Connie does not 'leap ahead' at all: she is simply struggling to imbue the new concept with a meaningful, action-oriented interpretation and, for her, this meaning-imbuing process is an essential part of her understanding of the concept.

Conclusion: In the above, a novice has been explicitly requesting clarifications on the newly introduced concepts of centraliser and conjugate of an element in a group. In particular, she has been asking for examples and for the raison-*d'-être* of the concepts. Her concept images of centraliser and conjugate seem to be dominated by a confused perception of the fixed and variable elements in the definitions as well as with a difficulty to accept the multi-valuedness of the defined elements (that there maybe more than one element commuting with an element x and that x might have more than one conjugates). There is some evidence here that this persistence of single-valuedness may have been reinforced by the student's association of commutativity and inversion with colloquial expressions such as 'back to itself' and 'swapping back'.

Section (v) A Controversial Step Into Mathematical Maturity. A Novice Realises the Pitfalls of Pretentious Formalism Through a Conflict Between Ordinary and Formal Language

Context: See Extract 9.5

Structure: In the following, Andrew presents his formal proof which is refuted in parts by Jack and the tutor. Finally he abandons the effort to formalise and explains his argument in ordinary language.

The Episode: A Factual Account. See Extract 9.5

An Interpretive Account: The Analysis

Note on the delicate mathematical argument in GRF5.8d. The converse of *Lagrange's Theorem* is not true. It is not true that for every factor of $n=|G|$ there is a subgroup of that order. What Andrew has been trying to convey is the idea that if n is not a prime, then, for at least one of its divisors, there is an element of G of order equal to this divisor.

The problem with Andrew's presentation is that he is trying to convey this idea in what he thinks is a properly formal way. So he assumes $|G|$ is not prime and he hopes to reach a contradiction to the initial hypothesis that G has exactly one proper subgroup. In the following table I juxtapose Andrew's written and verbal expressions:

Andrew's Verbal Expression	Andrew's Written Expression
A1	$\|G\|=pq$ $p,q\in \mathbb{N}\setminus\{0,1\}$ $x\in G\setminus\{e\}$ $\|x\|=1$, pq, proper factor of pq
A2	if $\|x\|=pq$, then $\|x\|=pq$ then $gr<x>$, $gr<x^{i}>$
A3	if $\|x\|=n$, n/pq $\{e\}$ $gr<x^{n}>$

Andrew's effort to reproduce a formalistic presentation is evident in the right-hand cells of the above table. He is introducing names for each one the numbers or elements in A1-A3 and tries to define them with precision (example: his definition of p, q and x in the first cell). He is also interrupting the flow of his thinking in order to discuss notational issues or terminology (proper factor and notation for order of an element in A1). He is trying to invest his approach with respectability by stressing that the idea used in this proof was also used in GRF5.1 (he refers to the lateralisation of cases in Extract 9.3).

Jack and the tutor seem to be gradually frustrated by Andrew's chaotic monologue and his equally unclear effort to formalise on the b/b. As the tutor said in the beginning, Andrew seemed to have a generally correct intuitive grasp of the argument in his writing but his presentation was confused. In A1 another element comes through: Andrew indeed is trying to reach contradiction by pointing at the existence of more than one proper subgroup but he seems to ignore that there is already one proper subgroup of G, $\{e\}$, and that he only needs to construct one more. In A1 Andrew sounds as if he is anxiously trying to point at two proper subgroups and because he has assumed n, the order of G, to be the product of two numbers p and q, he is struggling to mould these two proper subgroups so that they are of order p and q. In a sense he is the victim of his own notation because p and q are the notation he used to express the hypothesis that n is not a prime: if n had been assumed to be 36, for instance, then p and q could have been 6 and 6, or 3 and 12, 4 and 9, 2 and 18. Any of these combinations give the product of 36, but isn't it arbitrary to choose one of them and then try and find elements in the group of that order?

So Andrew has introduced quite heavy-handed notation without necessarily having control over it. Coupled with his muddled monologue in A1, he causes confusion to Jack whose two interventions ('can p equal q' and 'not happy with the logic behind that') perhaps do not address the main drawback of Andrew's presentation but reflect Jack's "gut reaction" of discomfort with it.

Things worsen in A2 where n acquires a second role, as a divisor of pq which is now, in Andrew's words the order of G. Andrew insists on trying to construct a proper subgroup of order p (again arbitrarily he chooses to pursue an element of order p and not of q). Again he tries to support his approach by comparing it to the already accepted approach used in GRF5.6. Jack's and the tutor's patience seems to be almost exhausted. I also note that in the tutor's response there is evidence that the tutor still thinks that n in Andrew's words is the order of G; Andrew has not made it explicit — it is very likely he hasn't even noticed- that he changed the role of n.

A3 is evidence of how Andrew's disillusionment begins to collapse. His intuitive belief is fortunately still strong: 'I must be able to choose something!' he exclaims.

With 'what am I doing?' he seems to realise that his formalising effort has failed and has also damaged the clarity of his thinking. By turning his back to the b/b, he disregards what he did so far and with 'so, if n is the order of G and n is not a prime then for a factor of G there must be a proper subgroup of G of order equal to that factor' regresses to ordinary language but in an illuminating way. Even though he does not truly explain in depth where the contradiction lies, the strategy of the proof is there, transparent and consummate.

Andrew's abandoning the effort to express formally can be seen as a regression to more "primitive" forms of expression: it shows that Andrew as a novice has not been enculturated into the conventions of mathematical formalism. He seems to pursue an enculturation, possibly more keenly than other students in these tutorials, but he hasn't quite mastered it yet. His writing on the b/b is a clumsy imitation of textbook writing and this clumsiness almost puts Andrew's idea into the risk of being perceived as using the converse of *Lagrange's Theorem*. Also his monologues are at times incoherent and possibly reflect the lack of clarity in his mind about the formalistic approach he wishes to espouse.

I finally note that Andrew's shifts of approach, whether seen as a regaining of confidence in ordinary logic and language, or seen as a regression to a familiar but not quite formally acceptable way of mathematical expression, only take place because of the tolerant learning environment created by the tutor. Whether Andrew took a step towards mathematical maturity in Extract 9.5, or he conservatively abandoned his effort to formalise and regressed to the convenience and familiarity of ordinary language, the incident has been illuminating (possibly for him too) as to the tensions that tantalise the novices' decision making with regard to their choice of expression.

Conclusion: In a debate-friendly learning environment, a student attempts a formalistic presentation of his argument for GRF5.8d and fails (possibly mostly because of his tendency to imitate textbook writing indecipherably and of his inconsistencies in the introduction of new notation). Unable to continue, he switches to ordinary language which allows the argument to be expressed in more clarity and precision.

Section (vi) The First Isomorphism Theorem as a Container of Compressed Conceptual Group-Theoretical Obstacles

Context: See Extract 9.6

Structure: The tutor leads the students through the proof of the First Isomorphism Theorem via very closed questioning. The overall impression of the sessions is that even when the students respond to the questions correctly they do not seem to have a global understanding of what is happening and how these questions relate to the proof.

Note to the reader: since the following is a fragmentary presentation from the four tutorials, the proof of the theorem is not fully presented in any of them. Thus, in order to follow the Factual Account, I recommend reading the proof in fig.6 first.

The Episode: A Factual Account. See Extract 9.6

An Interpretive Account: The Analysis

In the following, I comment on the sessions individually as well as examine them comparatively.

Memory Retrieval as a Necessary but not Sufficient Condition for Meaningful Understanding. Successful retrieval of the theorem varies in the fours sessions. Except Beth, the other students either remain silent or contribute very weak associations (Eleanor with normal subgroup, for instance). Patricia's association with 'division' (possibly invoked by the use of '/' in the statement of the theorem) is accompanied by her query on the meaning of ~. I note that similar evidence was given by these students with regard to *Lagrange's Theorem.*

Patricia's interpretation (~ means =) does not make sense because G/K contains sets of sets and Imϕ contains elements of the group. It nevertheless reflects the novices' muddled perception of ~ stemming from the diverse meanings that this symbol has in various mathematical contexts. Student Camille in Episode 7.4 is querying the tutor about the meaning and various uses of the ~ symbol. As elaborated upon below, Patricia's non-sensical interpretation of ~ as = conveys how problematic the perception of the newly introduced notion of isomorphism is.

In sum, retrieval of the theorem is generally problematic. The tutor, most notably in Tutorial 4, sounds alarmed and firmly reminds the students (similarly to the tutorials on *Lagrange's Theorem*) of the importance of 'a theorem with a name attached to it'. The tutor's firm and strict recommendation is a pragmatic and not an epistemological argument. In fact the novices have given contradictory evidence as to the power of persuasion of the different arguments: so — whilst, for instance, Connie in Extract 9.4 does not sound satisfied at all with pragmatic explanations of the type 'you need to learn this because it will be of use later' and asks for the *raison-d'-être* of most concepts she is introduced to — in most other occasions students seem to feel motivated to explore and understand some new concepts simply on the basis of its highly probable appearance on an exam paper. I do not see these two motivational forces as mutually exclusive but it seems to be didactically more appropriate to employ motivation of an epistemological rather than a pragmatic nature since the former is more likely to induce a more critically reflective approaching of the concepts than the former.

Incidentally I note how the students reproduce with relative ease information that has been given to them in catchy phrases:

$Kg_1=Kg_2$ iff $g_1g_2^{-1}\in K$
the 'obvious map' from G/K to G is to map Kg to g

Unfortunately immediate retrieval does not precede immediate, deep or in fact *any* understanding of what the phrases mean. It remains at best an automatic and effective act, at worst an act void of meaning that fosters a false impression of achievement.

Similarly the students apply the rule of coset product, for instance, and calculate correctly but cannot make any decision about how to use it in order to support the completion of the proof. They comfortably provide answers to questions that constitute progressive steps of the proof as long as these steps have been predesigned by the tutor. Only once the tutor attempts a more global comment regarding the internal structure and resonance of the *First Isomorphism Theorem* (Tutorial 3) when she says that the 'point' of the proof lies in the strength of the homomorphic mechanism.

Another feeble but noteworthy sign of instrumental understanding of the homomorphic property is given by Abidul when she helps Frances (who is 'stuck' with $\phi(g^{-1})$) by pointing out that 'it [the $^{-1}$] doesn't matter', therefore $\phi(g^{-1})$ can be written as $\phi(g)^{-1}$. Abidul's phrasing ('doesn't matter') reflects a procedural, cause -and-effect view of ϕ.

Finally, the tutor's comment on the mechanism of the proof as 'very standard' of proofs involving quotient groups is an attempt to generalise the approach used in the proof and hint at its potential as a methodological tool. In fact the 'standardisation' comes from the fact that the mapping $Kg \rightarrow g$, which she has been calling the 'obvious map', is used quite often in group-theoretical proofs, when cosets need to be mapped on the elements of the group. She rarely makes similar comments and even in this case she does not elaborate further.

The Problematic $\Leftarrow$ *Direction of (*) and the Properties of an Isomorphism.* Except Frances and Beth the rest of the students have problems in interpreting the $\Leftarrow$ direction of (*) as the definition of 1-1 correspondence. Eleanor confuses it with onto but soon changes her mind and in Tutorial 3 a discussion is triggered that reveals a more general confusion.

First the students are interpreting (*) as providing information about homomorphism ϕ, when in fact (*) provides information about the well-definedness and the 1-1 property of ψ. The students are deceived possibly because their interpretation is an interpretation by appearances. Moreover through P2-P6 Patricia and Cleo clarify, via the tutor's very leading questions (the multiple choice question they are given in T9 verges on the grotesque), their definition of an isomorphism. I note that, in this case of very directed questioning, the students' difficulties with the properties of an isomorphism (onto and 1-1 for instance) are not really explored because the teaching is solely oriented towards the elicitation of the answers that will further the progression of the proof.

Similarly in Tutorial 2 Beth appears to be severely concerned and confused as to the information contained in (*) as well as the homomorphic property of ψ. Like Patricia and Cleo, she does not carry out the switch from ψ to ϕ flexibly and cannot understand how manipulating ϕ can lead to an understanding of the properties of ψ.

Finally in tutorial 4, Abidul's A1 is another piece of evidence of the mechanical, despatched from conceptual understanding, conceptualisation of isomorphism ψ. Significantly in A1 the student starts dictating the necessary calculations for proving that ψ is a homomorphism: instead of

$$\psi(Kg_1Kg_2)=\psi(Kg_1)\,\psi(Kg_2)$$

she starts dictating

$$\psi(Kg_1g_2)=\psi(Kg_1)\ldots$$

as if the homomorphic property has to be proved for g_1 and g_2 and not for Kg_1 and Kg_2; in other words as if ψ is defined on G, not G/K.

Actually in the co-ordination and understanding of the link between ψ and ϕ, as well as the clarification about the definition of ψ lies largely the students' difficulty with the *First Isomorphism Theorem* and by implication with a large part of the newly introduced Group Theory. The degree of complexity in a problem which requires a well-co-ordinated manipulation of mappings between different sets is extremely high. ϕ is defined between the elements of a group (or two groups). ψ is defined between the cosets of the kernel of ϕ and the image of ϕ. This link between ψ and ϕ and the implications and importance of shifting back and forth from ψ to ϕ need to be explicitly made to the learner. The shift from one level of abstraction to another is not self-evident. In the absence of a didactically illuminating decomposition of the theorem to its constituent elements, it is not surprising that the students are not capable of making these shifts to more abstract levels.

An Incident with kerϕ Reveals Problematic Conceptions of Mapping. In Tutorial 3, given the students' difficulty with the meaning of $g_1 g_2^{-1} \in K$, the tutor initiates a discussion of the definition of $K = \ker\phi$. The tutor initially objects to Cleo's use of the term 'zero' for the identity element of G. Cleo quickly corrects (C2) — this is a common terminological mistake that the students habitually make and equally habitually correct when prompted by the tutor. Then in the explication of the definition (T3) it becomes evident that C1 was phrased in a rather problematic way not only because of the student's using the term 'zero': in Cleo's sentences the subject of the verb 'maps' in C1 and C3 is not clear at all. T4 is an interpretation that equates

Cleo's 'maps the elements to the identity'

with

'something is equal to identity'.

C4 possibly means 'element g goes to the identity element' and C5 that 'it goes via ϕ'. The tutor does not explore any further Cleo's grammatical state of mind with regard to the definition of kerϕ and completes the writing on her own initiative. Therefore it remains an open question whether Cleo's perception of ϕ as a mapping has been at all illuminated by the exchange of short verses in C1-C5. As in Extract 9.2, the student's words reflect an undecided perception of mapping as a machine. In this perception however it is not clear what is mapped where. Cleo's antonyms (it) as the subject of the verb 'maps' (C1 and C3) as well as her substituting the verbs (C4 and C5) with 'does' are lexical substitutions that possibly reflect and determine the ambiguity of her thinking.

Conclusion: In the above, evidence was given of a number of difficulties in the conceptualisation of properties associated to the notion of mapping (homomorphic property, onto, 1-1, well-definedness of a mapping); also of the varying degrees of abstraction involved in the definition of a mapping between elements of a group or the cosets of a subgroup and the elements of the group. The high degree of abstraction and the conceptual difficulties have then been linked to the students' cognitive enpuzzlement in Group Theory which culminates at the introduction and proof of the First Isomorphism Theorem for Groups. A didactical decomposition of the constituent elements of the theorem was then pointed out as a potentially helpful tool for the understanding of its content and proof.

Section (vii) A Frustrating Vicious Circle of a Novice's Struggle to Construct a Meaning of a Coset

Context:	See Extract 9.7
Structure:	In the following the tutor and the student engage in repetitive and circular dialectics during which the student tries to construct a meaning for the new concept of coset.
The Episode:	A Factual Account. See Extract 9.7

An Interpretive Account: The Analysis

First I note that the structure of this session is circular and repetitive. I also note that the tutor is patient and compassionate of Connie's repeated folding back to the same questions. In fact he could probably be deemed partly responsible for the repetitive circularity of the session: when there is evidence that Connie does not satisfactorily understand his argument for GRF7.3ii, he merely repeats the argument — almost invariably. It takes two of Connie's requests for clarity (C1 and her disappointed sighing shortly after that) to realise that he has to 'go into it in detail'; namely to re-address his arguments by referring back to basic definitions and concepts such as the coset. Throughout the session it is not clear that he realises the degree to which Connie's difficulty with the argument in GRF7.3ii can be attributed to her unrequited struggle for the construction of a meaning for coset. My main aim here is to give an account of this struggle.

Connie's Struggle for a Meaningful Conceptualisation of Cosets. This is a session on Connie's very problematic confrontation of the new concept of cosets. The evidence of her difficulty is ample but in a way it is also not very revealing: hints at potential sources of Connie's difficulty are not really suggested in Extract 9.7 but I think, maybe a bit paradoxically, this situation-of-non-disclosure renders this piece of data quite powerful. In its elusiveness this is a piece of data which reflects a very common cognitive situation: the cognising subject feeling unease and requesting enlightenment from the tutor. Unable to articulate the problematic aspects of her cognition of a new concept, the student's request comes through as vague and imprecise. In turn the tutor, being himself unaware of the sources of the problem, keeps providing feedback that seems to frustrate the student's attempt to make sense of the new concept. The outcome is a cognitive situation trapped in a vicious circle of mutual misunderstanding. The following presentation illustrates this vicious circle.

Before proceeding with the account of Connie's attempts at making sense of cosets, I briefly mention three incidents that are, peripherally, part of the picture of her difficulty:

- her confusion (C2) of the notation for the quotient group G/K with the notation for the complement of a set, G-K. So due to the iconic similarity of the two symbols Connie confuses a coset (G-K) with the set of cosets (G/K). This actually hints at how perplexed Connie is in the beginning of this session,
- her misuse of the term 'index' to mean the order of a group (C15),
- her uncertainty whether gK is called a 'left' or a 'right' coset (C24)

In the following, I present Connie's meaning-making attempts.

C1 is an imprecise interpretation of the requested proposition in GRF7.3ii. Despite the signs of vagueness and confusion in Connie's words the tutor presents his argument for the proof of GRF7.3ii (from now on referred to as The Argument). Connie is then evidently confused with G/K and G-K and the tutor repeats The Argument. Only when Connie sighs with disappointment he suggests 'getting into more detail' (a sign that the tutor is perhaps more sensitive to affective as opposed to cognitive signs of perplexity).

His back-to-basics trip is short though (it includes a vivid metaphor of cosets as parcels produced from multiplying the elements of a subgroup with elements of the group) and, when he returns to The Argument, Connie (C4 and C5) folds back to redefining the basic concepts involved in The Argument. I note here that, in these two utterances, Connie appears concerned about the idea of K being a subgroup AND a coset. She listens to the tutor's explanation that K is both a subgroup and a coset, but generally cosets are not subgroups because they do not contain the identity element of the group. She then returns to The Argument and disappointingly observes that 'the coset of K in G is K' which is a sign that the tutor's explanations have not been entirely received and also that Connie thinks in terms of one coset only (maybe similarly to her thinking about conjugates in Extract 9.4).

The tutor returns to the presentation of The Argument and talks briefly about a slightly more general case than $[G:K]=2$, $[G:K]=3$. The latter, the case for $[G:K]=3$, seems to attract Connie's attention and a bit later she appears repeatedly asking about this case.

I note that the tutor's drawing (see fig.7) seems to enhance Connie's concept image of cosets (C7) even though her use of the term 'subgroups' instead of 'cosets' may not be an entirely coincidental slip of the tongue. The tutor corrects 'subgroups' to 'cosets' and also notes — with regard to 'half' — that her phrasing applies to 'finite situations'.

He then responds to Connie's queries about the case for $[G:K]=3$ with an explanation of how the cosets would look like in that case. He also juxtaposes the cases $[G:K]=3$ and $[G:K]=2$ by returning to The Argument. Similarly to her question about K and G-K, Connie asks whether the cosets g_1K and g_2K in the $[G:K]=3$ case are also subgroups (C11: I note her use of the vague word 'things'. She still hasn't sorted out what they are. No progress seems to have been made.)

The tutor's subsequent explanations include a second metaphor for cosets (the first one was cosets-as-parcels): 'translates of subgroup by a group element which is not the identity element'. Given the tutor's picture that lies on the paper in front of her and his latest metaphor, Connie continues her inquiry (C12). Connie is striving for a rediscovery of the concept, in her own terms, in the order of her own thinking. In this quest some elements of ambition can be traced: she has been playing around with numbers 2 and 3 and then she is asking about the case where the 'index' (meaning the order' of the group) is a prime (C14 and C15). So she seems to attempt a generalisation.

C18 is disappointing and ambiguous: she hasn't realised that for every subgroup there are plenty of cosets. So to her, one subgroup means one coset. It is true that if G is the subgroup then there is only one coset and it is G itself but Connie sounds very muddled and it is not very likely that she can distinguish this case. The rather frustrating circularity in her mind becomes evident when she asks again about the case $[G:K]=3$ (C20). C22 may be evidence that her impasse lies in her weakness to understand why left cosets and right cosets aren't <u>always</u> the same. Maybe she is interpreting sameness as uniting-to-give-the-same-whole, namely G (C23).

C24 and C25 illustrate that Connie is still struggling with understanding the operation according to which a coset gK is constructed (multiply an element $g \in G$ with all the elements in a subgroup K of G; the products of these multiplications are the elements of gK). C26 and C27 then is a return to The Argument. C28 is then a brief, but possibly not very reliable, reassurance that Connie is satisfied with her understanding of the Argument.

Conclusion: In the above, a novice's struggle for the construction of a meaning for the new concept of coset has been accounted for as a vicious circle of mutual misunderstanding with her tutor. The tutor has employed a variety of devices in order to convey the meaning of a coset (cosets as parcels; cosets as translates). The sources of the novice's difficulties are not largely disclosed but the circular dialectics seem also to spiral down to an exploration of gradually more basic knowledge relating to the problem sheet question: from The Argument, the discussion, towards the end, is about the construction of cosets. This spiral journey as well as the persistence of ineffective ideas in the student's mind graphically illustrate the abysmal complexity of the novice's cognition and also the didactical need for an emphasis on constructive learning processes (that is processes that cautiously build on solid previous knowledge or allow revisiting and reconstructing previous knowledge with facility).

Section (viii) An Example of the Tension Between Novice and Expert Approaches to Mathematical Reasoning: The Need to Learn How to Achieve Mathematical Resonance by Creatively Co-ordinating and Manipulating Relevant Knowledge

Context:	See Extract 9.8
Structure:	This Extract follows Extract 7.8. Cathy presents her proof and the tutor suggests an alternative. The two proofs reflect some of the differences between a novice and an expert approach to mathematical reasoning.
The Episode:	A Factual Account. See Extract 9.8

An Interpretive Account: The Analysis

I think that it bears some significance that Cathy is a bit reluctant to present her solution in the beginning of Extract 9.8; given that Extract 9.8 follows the Extract 7.8, it is possible that she begins to suspect that, though correct, her approach is not exactly up to the standards of elegance and resonance with the material she has been taught recently. This was the crucial point made in Episode 7.8. Possibly under the influence of the discussion there Cathy makes the rather aesthetic comment ' you may not like it' which can be seen as a sign of a developing taste for a certain mode of reasoning.

Similarly to Episode 7.8 Cathy resorts to a solid arithmetical handling (to prove that integers a and b are equal, it is sufficient to prove that $a \leq b$ and $b \leq a$). The tutor on the other hand employs a theorem and invests the arithmetical relationship given in the question ($hcf(o(x),o(y))=1$) with its group-theoretical meaning ($\langle x \rangle \cap \langle y \rangle = \{e\}$). Again, the juxtaposition highlights the differences between an expert and a novice approach. There are some redundancies in Cathy's way as well as some unclarified points — not necessarily visible in the abbreviated version in Extract 9.8. Most important though maybe her starting to conceptualise the need, beyond the merely aesthetic, for an embedded, contextualised mode of reasoning and for an organically connected argumentation, in a way which will turn the coherence and connectedness of mathematical theories to her benefit. In other words via this juxtaposition of approaches, she might begin to learn about the benefits of mathematical expertise.

Conclusion: In the above, a novice manages with minimum principles (mostly arithmetical), albeit not in an impeccable formal way, to reach the conclusion and complete a proof. This approach is juxtaposed with the tutor's proof which is well embedded in the mathematical context of the course (lecture content, problem sheet material, textbook approaches) and thus shorter and more relevant. Therefore as an example of the differences between novice and expert approaches to mathematical reasoning, this Episode is evidence to a novice characteristic: that, even when undoubtedly imaginative and skilful, novices have not acquired yet the skill of association and co-ordination of relevant knowledge which equips mathematical reasoning with the power of resonance.

PART III A Synthesis of the Findings in the Area of Group Theory. Indications for the Cross-Topical Synthesis in Chapter 10

In this chapter the students' first experiences of fundamental group-theoretical concepts were explored in a series of Episodes from the second term of observation. The novices here appeared as especially lacking in basic knowledge which resulted in their lack of understanding of an argument for a proof or the proposition of a theorem [7]. As *paradigmatically problematic* in this sense emerged the concept of *coset*. So, for example within the context of *Lagrange's Theorem*[0] the students' perceptions of certain set-theoretical properties (distinct and disjoint sets, $\cap$ and $\cup$) seemed to interfere with their understanding of cosets.

While constructing cosets the students also appeared to be in difficulty with the abstract nature of the operation between elements of a group [1]: interferences from the properties of numerical operations were observed. In fear of reinforcing these interferences, the tutors discourage the students from using metaphorical expressions such as *divided by*. Somehow paradoxically, however, they encourage them to say *multiply with the inverse* which does not rule out, I think, the possibility of numerical interferences. Similarly problematic turned out to be the linguistic metaphors such as *times* and *powers of* — also used sometimes unclearly interchangeably by the students — with regard to cyclic groups [1].

Linguistic condensation of meaning also causes difficulties, for instance [1] in the context of the concept of *order of an element* of a group, which is an abbreviation for *the order of the group generated by an element*. Notationally the abbreviation is similar: $|g|$ is the commonly used notation instead of $|\langle g\rangle|$. Moreover another aspect of the very problematic encounter with the notion of $|g|$ seems to be the static and operational duality of the concept: $|g|$ is the number of elements in $\langle g\rangle$ and, at the same time, the number of times the power of g has to be taken in order to cover all the elements of $\langle g\rangle$. After $|g|$, the powers of g in $\langle g\rangle$ start repeating themselves. So, in a sense, *order of an element* is a notion containing both information about a static characteristic of $\langle g\rangle$ (its cardinality) and information about a way to construct $\langle g\rangle$ (take the power of g, $|g|$ times). This type of duality is commonly seen as a source of cognitive strain for novices (functions, limits) and it is likely that *order of an element* is not an exception.

[0] Condensed evidence from a number of episodes not reported in the Chapter.

As in Chapters 6-8, the novices are primarily engaged in a meaning bestowing process [2, 4] with regard to the newly introduced concepts: they inquire about the *raison-d'-être* of the concepts. Examples:

- Equivalence class and coset [2]: the efforts were characterised by a use of geometrical metaphors aiming at the construction of a mental representation of the new concepts (equivalence classes as *straight lines*, cosets as *squares*). As in Chapter 8, these geometrical representations were interpreted literally and again the use of geometrical metaphors emerged as a dubious means of image construction for some particularly abstract concepts.

- Centraliser and conjugate of an element in a group [4]. In a characteristic instance, a student's concept images seemed to be dominated by a confused perception of the fixed and variable elements in the definitions as well as with a difficulty to accept the multi-valuedness of the defined elements (that there maybe more than one element commuting with an element x and that x might have more than one conjugate). Her persistence of single-valuedness, explored in other contexts in Chapter 8, may have been reinforced by the student's association of commutativity and inversion with colloquial expressions such as 'back to itself' and 'swapping back'.

The novices also appeared to be in difficulty in conceptualising a mapping between elements of a group and sets of elements of the group [2, 6] (two characteristic examples of such mappings involving the concepts of conjugate, centraliser and equivalence class and in the context of the First Isomorphism Theorem for Groups were presented). As in previous occasions in these tutorials, where the notion of sets of sets and of mappings between sets whose elements are gradually departing further from the simple arithmetical correspondences the students are familiar with from school mathematics, the increasing degree of abstraction as well as the students' problematic perceptions of the notions of domain and range of these mappings rendered their understanding an extremely complex process.

Within the abstract context of Group Theory a contrast between the novice and the expert approaches to abstract mathematical reasoning was observed: so the expert's logical, succinct and benefiting from hindsight argumentation was contrasted with the novice's more-naturalistically-born-out-of-the-proof argumentation [3]. In one instance [3], a novice managed with minimum principles (mostly arithmetical), albeit not in an impeccable formal way, to reach the conclusion and complete a proof. Her approach was juxtaposed with the tutor's proof which was well embedded in the mathematical context of the course (lecture content, problem sheet material, textbook approaches) and thus was shorter and more relevant. Typically most novices have not acquired yet the skill of association and co-ordination of relevant knowledge which equips mathematical reasoning with the power of resonance. Also the contrast between expert and novice approaches has an enculturating function: in the process the novices are clarifying the modes of formal expression at which they are expected to aim.

The students' participation in this enculturating effort is not successful when it is based on a superficial imitation of formalising practices. The instance [5] of a novice's attempt at a formalistic presentation of his argument — which fails possibly because of his tendency to imitate textbook writing indecipherably and of his inconsistencies in the introduction of new notation — is characteristic: unable to continue, he switched to ordinary language, regressing thus to a form of expression with which he felt his argument would be articulated in more clarity and precision.

Learning emerged here as inseparable from the social forces under the influence of which it takes place — here tutoring: the instance [7] of a novice's viciously circular struggle for the construction of a meaning for the new concept of coset was presented as a characteristic example. This instance was described as a succession of mutual misunderstandings between the student and the tutor who repeated identically the argument for the proof of the question on which the discussion was based. The sources of the novice's difficulties were thus not largely disclosed but the circular dialectics finally — and after the student's persistence — seemed to spiral down to an exploration of gradually more basic knowledge relating to the question and in particular to the construction of cosets (for which the tutor employed a variety of devices: cosets as parcels; cosets as translates). This spiral journey coupled with the persistence of ineffective ideas in the student's mind suggest the complexity of the didactical interaction and also, possibly, the need for an emphasis on constructive learning processes (that is processes that cautiously build on solid previous knowledge or allow revisiting and reconstructing previous knowledge with facility). A section of Chapter 10 — the cross-topical synthesis of the observations on the novice mathematician's cognition presented in Chapters 6-9 — is allocated to the brief elaboration on this and other didactical observations.

Chapter 10

Patterns in the Novice Mathematician's Cognition in the Transition to Advanced Mathematical Thinking: *A Cross-Topical Synthesis*

Introduction

In Chapters 6-9, the themes, outlined in the Interlude and related to the novice mathematician's cognition, were elaborated within the four mathematical areas of Introductory Analysis, Calculus, Linear Algebra and Group Theory. In Part III of each one of these chapters the *paradigmatically problematic* concepts within each area (*supremum and infimum*, *limit*, *spanning set* and *coset*) provided the contextual basis for a *micro-discourse* on the novice mathematician's cognition. In this final chapter the focus is on a *macro-discourse*, a cross-topical synthesis of the observations made in Chapters 6-9. In Part I the novice mathematician is described as an individual learner struggling to come to terms with the intellectual demands of mathematical formalism, whereas in Part II the novice's encounter with mathematical abstraction is described as an enculturation process which takes place through the mediation of an expert, the tutor. So in a sense the perspectives in Parts I and II reflect the distinction of perspectives, made in Chapter 1, between learning as a cognitive process and learning as a sociocultural process. By implication — if Parts I and II form a coherent discourse on the novice mathematician's cognition — the approach taken in this study gives evidence that a description of learning cannot take place in the absence of either of the two perspectives and that, in fact, the dissection of the two perspectives is dubious and unnatural.

In Parts I and II the novices' cognition is explored from a learning point of view. A brief and general reference to issues from a teaching point of view is the focus of Part III. Part IV contains a few methodological observations, made in the light of the experience from this study. The chapter closes with a short statement (Part V) about how the synthesis of the findings in this Chapter is embedded in the research reviewed in Chapter 1 and about possible extensions of the study in the future.

Part I The Novice Mathematician's Encounter With Mathematical Abstraction As the Individual Learner's Sense-Making of a New Way of Thinking

The novice mathematician's encounter with mathematical abstraction is presented here in a network of themes — briefly outlined in the Interlude — that emerged from the data analysis exemplified in Chapters 6-9. The complexity of the observed phenomena does not allow to present these themes in a linear manner because of the multiple links and interpretational overlappings between them. Below I give a brief guide through the presentation. The text in italic in this guide corresponds to the titles of the themes as presented in more detail in the next pages.

Learning in this study is seen as taking place in the encounter of the novice mathematician with a new way of thinking mathematically, mathematical abstraction. The novice's entrance into this new world occurs as an encounter with mathematical formalism in terms of reasoning and in terms of the definition of new concepts. A learning process consists of the novice's attempts to make sense of this new world; or in other words to construct meaningful images. This process is characterised by a range of tensions between the old and the new world and there is evidence in this study of a number of these tensions. Concept-image construction takes place in part due to the exposure to the concept definition and it assumes a certain amount of prerequisite knowledge. *In the absence or weakness of this prerequisite knowledge concept-image construction is obstructed and this results in the problematic interaction of the learner with the concept definition.*

Concept-image construction is described here as an attempt to construct meaningful metaphors and to explore the *raison-d'-être* of the new concepts and the new reasoning. This *image construction through acquisition of — visual or other — metaphors and through existential meaning bestowing processes* occurs as a battle between the old and the new ways of the two worlds. It is thus characterised by *the tension between the informal-intuitive-and-verbal and formal-abstract-and-symbolic modes of thinking.* This tension has been explored here as a *tension between verbal and formal/symbolic language* and as a *tension between informal and formal modes of reasoning.* The novices' difficulties in formalising have been identified to be *difficulties with the mechanics of formal mathematical reasoning* as well as *difficulties of applying the mechanics of formal mathematical reasoning in a well-integrated and contextualised manner.* Specifically in this study the embeddedness *of the novice's reasoning* has been *linked with their ambiguous perception of what knowledge can be assumed (the tension between Proof by First Principles and Proof* *by Theorem Quoting* is seen as a grand example of this ambiguity), *and with the fragility of their knowledge as a result of this ambiguity.*

In the above description my perspective on the novice mathematician's cognition is dominated by a polarisation between the old (school mathematics: informal, concrete and intuitive) and the new (university mathematics: formal, abstract and deductive), the uninitiated and the initiated, the novice and the expert. In Part II this *juxtaposition of the differences between novice and expert approaches* is seen *within the context of describing the encounter with mathematical abstraction as an enculturation process.*

In the following, the themes presented above in italics are elaborated on the grounds of the evidence given in Chapters 6-9.

Ia. Concept-Image Construction

The novices' concept-image construction is discussed here as an interaction with the concept definition, as a metaphor acquisition and as a meaning bestowing process.

Ia.i Concept-Image Construction and the Interference of Not Solidly Established Previous Knowledge: the Problematic Interaction With the Concept Definition

Within the discourse on concept image / concept definition, as reviewed in Chapter 1, the novices' acquisition of new concepts takes place as a process of *concept-image construction*. As elaborated in this Part, the evidence in this study resonates with this description. In their familiarisation with new concepts the students seem to construct a variety of explicit or implicit images related to, co-existing, or existing despite the concept definition. The evidence ranges over the wide spectrum of new concepts in the several mathematical contexts observed in this study and especially the *paradigmatically problematic concepts*. Along with these sometimes inadequate or contradictory concept images, the students' inadequate knowledge of necessary and previously established definitions interferes with concept-image construction (see for details on the various concept images of the *paradigmatically problematic concepts* and of the interferences of inadequate previous knowledge in the relevant chapters).

As an example here I mention the notion of function: in Chapter 8, the students' difficulty in reproducing definitions of linearity, onto and 1-1 function, ker$T$, ImT and T^{-1} and the absence of basic knowledge about characteristics of mappings (domain, image, examples on $\Re^{\Re}$) functioned as obstacles in their efforts to handle questions related to the new notion of a linear mapping. The importance of some mathematical notions as the bare essentials of mathematical thinking was mostly emphasised in the cases where the students' thinking of the concept of function, disconnected from relevant notions such as domain and range, led to *cul-de-sac* cognitive situations; especially in the contexts of Linear Algebra (mappings) and Group Theory (isomorphisms). Similar evidence was given in the collected material on homeomorphisms – not reported here – in the context of Topology. The notion of *function* reappears in the following as a cross-topical *paradigmatically problematic concept.*

Ia.ii Concept-Image Construction Through Acquisition of Visual or Other Metaphors and Through Existential Meaning Bestowing Processes

Apart from, or in parallel with, interacting with the concept definition and building on previously established knowledge, the novices construct images of the new concepts

- by adopting metaphors that they seem to be familiar with and
- by exploring the justification for the coming-into-existence of the new concepts.

Within the latter, the novices engage in concept-image construction by searching for the *raison-d'-être* of the new concepts. The topical discourse on the *paradigmatically problematic* concepts in Chapters 6-9 gives ample evidence of these inquiries: in the novice's – similarly to the mathematician's as an *homo historicus* — economy of mind a new concept justifies its coming into existence if it is of some utility. So, for instance the students' prejudice for continuous functions may ensue their view of limit as a redundant concept – what is the point of learning about $lim_{x \to a}f(x)$, when simply calculating f(a) gives the same result? The absence of a utilitarian justification, and the evidence that the novices' meaning and purpose bestowing actions usually take place under the severe and combined influence of both emotional and cognitive factors, raises serious didactical questions which are dealt with briefly in Part III.

The former (adoption of familiar metaphors) was extensively exemplified in Chapter 8, where the students reacted to the novelty of the abstractions in Vector Analysis by embedding their action almost exclusively in the geometrical context of the line, the plane and the three-dimensional space. Their literal interpretation of the geometrical metaphor of the plane was illustrated in their use of strictly geometrical language (most graphic example: the merge of the student's images of the Cartesian orthogonal system and the Parallelogram Rule in the construction of <v_1, v_2>); also in their body language and their orientation on the plane. This excessively literal interpretation of the geometric metaphor — which, I think, is intended by the tutors more to be an illustration than a generic example — may be due to the fact that the students have associated it with convenient and familiar algorithms. The evidence in this study reinforces this explanation: the novices tend to regress to familiar modes of expression and reasoning when confronted with new situations – for instance, by trying to critically embed the new concept in their previous knowledge (examples: when introduced to the notion of linear mappings, by asking whether particular examples of mappings they know about fit in the definition of linearity; when introduced to the notion of a Fourier Series by attempting interpretations of ~ as an approximation, an equivalence or an equality). On the other hand others try to suppress their unfamiliarity with the novel and abstract context: so some apply uncritically a new definition, even when they cannot give an example of the new concept.

Algorithmic competence is one of these familiar modes of behaviour and the students tend to apply it even within contexts of which they do not have a good grasp: most striking in Chapters 6-9 was the example of the students attempts to use the Subspace Test on a subset of $\Re^{\Re}$ whose contents they were not aware of: they confused f with $f(x)$ and their action-in-void soon led them to a *cul-de-sac* within the context of what was perceived by the tutor as a routine part of the Test (identifying the zero element of $\Re^{\Re}$). Again the cognitive embarrassment of a novel and abstract context resulted in an attempt to accommodate the new concept in what they already know: so, for instance, their difficulty with seeing the zero function as the zero element of $\Re^{\Re}$ reflected how a concept image of a vector on the plane (a directed line segment) interfered with the notion of vector as an element of a vector space. In the above the notion of function (this time in its process/object duality) emerged as problematic. Image construction of a number of other concepts (spanning set and basis, linear mappings and cosets are other striking examples) seemed to follow this pattern of adoption/regression to familiar metaphors.

Concept-image construction seems also to be influenced not only by the tendency for a regression to familiar contexts but also by a tendency for a perpetuation of familiar behaviours, mostly school-mathematical stereotypes such as the difficulty to accept the multiplicity built into a definition (there is more than one basis/spanning set for a set; there may be more than one element commuting with an element x in a group; also x might have more than one conjugate). This persistence of single-valuedness – since it has been interpreted as an implication of the tension between the student's school-preformal and newly-acquired-formal mathematical behaviours — reappears later in this Part and it can be embedded in a wider discourse on the novice's unease with the multiple nature of some mathematical concepts: see for example in Chapter 9, the discussion of the no(ta)tion for the *order of an element* in a group.

Finally I note that substituting the power of the abstraction in the new concepts with the convenience of a familiar context and concentrating on a competent if narrow-minded execution of algorithmic tasks within this context, is more controversial than it actually sounds at first: indeed, specificity of context is a powerful initiator into new concepts; but maybe a too powerful one. In this study, for instance, the extensive adoption of the geometric metaphor in the context of Vector Analysis (spanning) and in the context of Group Theory (equivalence classes and cosets) seemed to impede the cognitive leap into the Abstract-Algebraic context. Also within the Calculus context visualisation proved to be controversial: the novices debate the use of pictures in guessing limits and see pictures as specific representations. This specificity seems to impede the students' shift towards generality.

Ib. The Encounter With Mathematical Formalism

The novices' encounter with mathematical formalism is described here in terms of language and reasoning.

Ib.i The Encounter With Mathematical Formalism: Advanced Mathematical Semantics and the Tension Between the Informal-Intuitive-and-Verbal and Formal-Abstract-and-Symbolic Language

The novices do not seem to adapt unproblematically to the novel requirements of formal mathematical expression. Their problematic reactions, which I explain subsequently, can be classified as

- avoidance of formalisation and
- uncritical and precipitate adoption of formalisation.

I note that these reactions do not consistently characterise the novices, that is, it is likely that the same student, at one instance, may give evidence of avoidance and, in another, of adoption. Moreover these reactions are similar to their reactions to the novel requirements of formal mathematical reasoning, reported in the following section. This similarity can be seen as evidence of the inseparability of reasoning (thought) and expression (language), an idea which resonates with the description of mathematical cognition given in Chapter 1.

The novices' difficulties with adapting to the requirements of formal mathematical expression were most graphically illustrated in their response to their introduction to the formal definition of limit (see Chapter 7 for their difficulties with the δ–ε formalism as well as with assigning meaning to the definition; especially with comprehending the mechanism of the proposition that is contained in the formal definition and how this mechanism provides a tool for proving limits). The novices find it hard to see formal notation and logic as a vehicle for mathematical arguments: this results in their reluctance to become familiar with and employ formal definitions or attempt to construct logical arguments for a proof. Consequently they avoid formalisation and adhere to familiar modes of expression, such as ordinary language.

It is then significant — because it generates cognitive conflict — that their attempts to use ordinary language are not invariably successful (again Chapter 7 provides ample evidence to this). Even in cases where their use of language successfully conveys their general grasp of an idea, this success is not fully integrated in the process of presenting a consummate formal argument. This usually implies that their argument remains formally unsatisfactory: sometimes they realise this is so and sometimes they do not. When they do, they usually appear alerted to the necessity of a formal argument and at this moment it is likely that they also realise the power of formalism (in these instances the role of the tutor/enculturator — see Part II — is important).

A number of novices who seemed to be sensitised to the requirements of rigour, but also seemed to find it difficult to express formally with efficiency, demonstrated another tendency — examples were given in the context of convergence of series, of spanning sets and of Group Theory: in their writing or on the b/b they resort to brief and not explicit enough explanations. Asked for elaboration they argue orally in ordinary language and, even though the explanations are articulate, the students are hesitant and mistrustful about the legitimacy of this practice. In writing they prefer to leave their proofs void of explanation if this explanation is in ordinary language; orally they use it only after persistent prompting. Moreover they tried to negotiate the right to use an elliptic writing style — that is to use incomplete arguments in writing —, as long as they can explain what they did orally.

In other cases the novices who seem to be sensitised to the requirements of rigour resort to a superficial, indecipherable and inconsistent imitation of textbook or lecture formalism. In most of the cases looked at in Chapters 6-9, the novices then abandoned their proofs once they realised the impasse. In some other cases (for instance Chapter 9) the novice, unable to continue, regressed to ordinary language — which he had been trying to avoid — in order to complete an explanation of his proof.

Finally I mention here briefly (refer to the relevant chapters for more detailed analysis of each one of them) a number of characteristic cases where various aspects of ordinary language (terminology and grammar) merged with their mathematical counterparts; the merge — in most cases unconscious — generated confusion:

- the equivalence of the terms *not-continuous* and *discontinuous*, the locality or universality of the term *always* continuous,
- the grammatical link between the terms *spanning set* and *span* — which reflects their cause-and-effect conceptual link,
- the use of the term (by the novices) *empty set* to denote $\varnothing$ and $\{0\}$
- the numerical interferences with the group operation, e.g. commutativity, potentially perpetuated in metaphorical expressions such as *divided by* (used by the novices) and *multiply with the inverse* (suggested by the tutors); *times* and *powers of* within the context of cyclic groups.
- the linguistic and notational condensation of meaning in terms such as order *of an element* of a group, which is an abbreviation for *the order of the group generated by an element.*
- the persistence of single-valuedness reinforced by the association of commutativity and inversion with colloquial expressions such as 'back to itself' and 'swapping back' in the context of groups.

In the above, the linguistic difficulties of the induction into the culture of advanced mathematical expression were looked at in juxtaposition to — and as a tension between — the persistent and rich forms of ordinary language in which the novices have learned to express. As a way of thinking, advanced mathematics is deeply determined by its expression through its complex symbolic language. As a result, one should add to the difficulties resulting from this tension, the difficulties resulting purely from the extreme formalistic nature of some of the new notation the novices are expected to adopt — and possibly from the way they are introduced to this notation.

Ib.ii The Encounter With Mathematical Formalism: Advanced Mathematical Reasoning and the Tension Between Its Informal-Intuitive-and-Verbal and Formal-Abstract-and-Symbolic Modes

As said in the previous section, the novices' reactions to the necessity to reason formally are similar to their reactions to the necessity to express formally. Here I recapitulate briefly the ideas from the previous section and then embed them in the particular context of the mechanics of formal mathematical reasoning.

Difficulty in formalising leads to denial of formalisation and regression to more concrete and familiar modes of reasoning. Sometimes this denial is unconscious — the novices plainly and uncritically extend their school mathematical practices to university mathematics — or conscious — the novices reject formal reasoning as redundant once they are personally convinced. So, for example, they make tacit use of theorems which they believe are obviously true. This is possibly a perpetuation of A-level attitudes and a regression to familiar from school modes of action.

Even when they have conceptualised the necessity to formalise, they still struggle with the materialisation of this conceptualisation: so, for instance, they assume in their proofs what is to them intuitively obvious or what they are actually being asked to prove. Sometimes, at least in the beginning, their rather over-zealous allegiance to rigour — as they perceive it — yields hesitation towards their school mathematical practices which in turn deters them from, for instance, using some basic arithmetical facts. So they seem to need to clarify the distinction between rigorous and intuitive arguments, legitimate and illegitimate use of knowledge that is thought of as previously established. In other words, they need an explicit articulation of the new didactical contract of advanced mathematics.

In sum, even sensitisation to the need for rigour does not always imply that the novices are willing to attempt formalisation — these tendencies were most graphically illustrated in Chapter 6 where the novices were observed in their first encounters with the requirements for rigour. In the following, I elaborate on the novices' *difficulties with the mechanics of formal mathematical reasoning* and their *difficulty to embed the newly acquired mechanics in their mathematical action*. Finally this weakness to embed successfully is *attributed* to the above mentioned *fragility of their knowledge with regard to their ambivalence about the terms of the advanced mathematical didactical contract.*

Ib.ii.1. Difficulties With the Mechanics of Formal Mathematical Reasoning

In this section I refer briefly to characteristic difficulties of the novices with the mechanics of advanced mathematical reasoning. Details can be found in the analysis of Chapters 6-9. This reference does not intend to be exclusive – even in terms of the collected data — but, as it is based on the paradigmatical cases exemplified in Chapters 6-9, it reflects some consistently repeated features of the novices' behaviour.

In previous sections of this part, I have referred to the general *tendency for single-valuedness of the novice mind.* This tendency seems to extend to the reasoning domain too: so the novices, mostly in Chapter 6, which marks the beginning of their attempts to formalise, were found in difficulty when confronted with the multiplicity of options in the course of a proof and with the need to co-ordinate a variety of information in order to pick an effective option. Even the rather exceptional case of the novice, who commented on $\leq$ as a sign that is used in the Triangle Inequality not to denote inequality but to denote a variety of options for equality, can be implicitly interpreted as an expression of mild nuisance with this multiplicity. In general however the novices seem to be persistently resisting the potential pluralism of answers to a mathematical problem – for instance, the multiple possibilities for the third element of a basis in a problem with under-defined conditions for the basis.

In the same vein, they seem to be generally not at ease with choosing arbitrary numbers, establishing thus the universality of their proof, and maintaining this arbitrariness through the proof with consistency. The cognitive leap to generalisation, impeded, as said earlier, sometimes by the novices' adherence to specific examples (e.g. the geometric metaphor of the plane), was shown problematic in a variety of mathematical contexts. Significantly when a novice becomes conscious of the intricacies of generalisation, she seems to have better potential to disentangle these intricacies.

As mostly exemplified in Chapter 7, in the context of the various uses of the δ–ε definitions, the mechanics of implication (*proposition a implies proposition b, modus ponens*), that is the manipulation of the logic and the quantifiers contained in categorical propositions were shown to be deeply problematic. Negating and inverting these propositions is a nearly impossible task for most novices. Their difficulties with specific tools of formal mathematical proof, such as Mathematical Induction, Proof By Contradiction, and Proof by Counterexample were attributed to a lack of awareness of basic Logic Rules. So, for instance, some novices wondered whether the Base Clause of Mathematical Induction should be stated for n=1 or n=2;

or perceived the Base Clause and the $n=k$ step as perfunctory (only the k+1 step of the proof was deemed important).

Similarly, different — but mathematically equivalent — counterexamples seemed to carry different degrees of conviction. The novices were also reluctant to accept the limited validity of a theorem via a counterexample because they did not think it was useful to refute a theorem by a counterexample when it holds in a particular case of a problem. Moreover, in the same way that a counterexample disproves a proposition, some novices seemed to think that several examples prove it, thus not seeing the need for a deductive proof.

Ib.ii.2. Deficient Embeddedness of the Novice's Reasoning Linked With Their Ambiguous Perception of What Knowledge Can Be Assumed: the Tension Between Proof by First Principles and Proof by Theorem Quoting

In the previous section a brief description was given of the novices' difficulties with the mechanics of formal mathematical reasoning. Emphasis was placed on the purely logical aspect of these difficulties, so, in a sense, mathematical reasoning was viewed in a rather decontextualised manner, that is detached from the specificity of the mathematical topics in which it was applied. Here the emphasis is on the novices' observed deficiency to embed the newly-acquired tools of formal mathematical reasoning. This deficiency was partly attributed in Chapters 6-9, and mostly Chapter 6, to the students' ambiguity with regard to what knowledge they are allowed to assume.

In a number of occasions – mostly in the context of proving statements within Foundational Analysis — the students' proofs turned out inefficient not only because of their limited proving skills (as elaborated in the previous section) but because the students were not at ease with the assumptions they were allowed to make. Part of the problem seemed to be caused by the absence of clarifications about, for instance, what statements regarding the real numbers can be assumed. The students seem to be extremely vulnerable to such an absence and as a result over- and under- react to the requirements for rigour. In both cases, they seem to have been sensitised to the increased requirements of rigour in the new course but then abandoned to clarify these requirements on their own. Their vulnerability results in inefficiency or avoidance of rigour.

A basic aspect of their perplexity is to what degree assuming knowledge is compatible with the requirements for axiomatic rigour made by the lecturers and tutors in the beginning of the course. In this sense the tension between Proof-By-

First-Principles and Proof-By-Theorem-Quoting is a graphic illustration of their ambiguity. This tension was mostly observed in the context of limits, in Chapter 7 where the students were undecided as to whether they should find limits via the formal definition (First Principles) or via the algebra of limits (Theorem Quoting): using inequalities in order to manipulate quantities, graphing functions, guessing limits and using the algebra of limits are mathematical practices questioned by the novices as to their rigour and, hence, as to their acceptability. Due to their growing mistrust towards the practices of school mathematics (see earlier section of this Part), they avoid intuitive practices such as guessing limits and then proving them. Moreover they are not at ease with the alternation of practices (the use of the definition of limit alternating with the use of the algebra of limits). They also seem weak in distinguishing between the practices that they are supposed to espouse in different mathematical domains: in applied mathematics, retrospective use of unproved results is allowed, as opposed to foundational courses, where it is not.

A result of the students' ambiguity about the legitimacy of certain practices results in the fragility of their knowledge and, hence, in the inefficiency of their action. A characteristic of mathematical expertise is the ability to switch representations or modes of thinking with the aim to suit the context in question (expert and novice practices are juxtaposed from this perspective in Part II). The students' fragile knowledge about the nature of the objects they have to deal with and about the rules of the formal mathematical game, does not allow them this flexibility and so their action often turns out severely decontextualised (especially in Chapters 8 and 9) and not in resonance with the recent lecture or textbook material.

Mathematical objects, such as 0 and 1, misleadingly perceived as mythic carriers of simplicity, reappear, disguised in their plain semiotic appearance, in a variety of mathematical contexts and demand of the novice an acute perception of their multiple entity. But the novice, oblivious and not warned about the possibility of multiplicity, remains prisoner of his single-minded innocence. Zero can be a number, a function, a vector, a matrix, a polynomial... the list is long and the items not necessarily distinct from each other. But this multiplicity, while taken for granted, remains part of the hidden agenda of an utterly unnecessary mathematical mysticism. Similarly, the rules of the game of mathematical formalism, while taken for granted, remain tacit as if the essence of the game is to discover its rules and not to play it. In Part II, the novices' induction to the game of mathematical formalism is described as an enculturation process the responsibility for which is shared between the novice and the expert.

Part II The Novice Mathematician's Encounter With Mathematical Abstraction as an Enculturation Process

In Part I, the novice mathematicians' encounter with mathematical abstraction was described in terms of the tensions and the difficulties of their induction to the advanced mathematical culture; so the focus was on the novice mathematician as an individual learner-in-action within the context of tutorials. A complementary, and not necessarily distinct, way of describing the novices' encounter with mathematical abstraction is to juxtapose their practices with the practices of mathematical expertise – as represented in the tutorials by the tutors. This juxtaposition can be useful because it is possible that we learn something about the novice's cognition by contrasting it (contrast as accentuation) to what the novice is expected to aim at (mathematical expertise); also because through this contrast it may be possible to explore types of interaction between the novice and the expert that are congenial to the novice's cognitive needs.

As an example I refer here to the instance from Chapter 9 where a novice's viciously circular struggle for the construction of a meaning for cosets was described as a succession of mutual misunderstandings between the student and the tutor who insisted on repeating identically the argument for the proof of the question on which the discussion was based. Eventually the circular dialectics – largely due to the student's insistence — spiralled down to an exploration of gradually more basic knowledge relating to the question and in particular to *cosets*. This spiral journey illustrated graphically the need for an emphasis on constructive learning processes — that is processes that cautiously build on solid previous knowledge or allow revisiting and reconstructing previous knowledge with facility. So, in a sense, by describing this interaction in terms of what eventually became an enculturation process, some access was gained to the optimisation of this process.

In terms of the contrast between expert and novice approaches, the brief description of their differences given in Chapter 1, Part III was confirmed: so, in Chapter 6, in dealing with the basic inequalities of Foundational Analysis, the novices, unlike their tutors, seemed to be lacking in the kind of mathematical experience that empowers hindsight, reinforces a more fruitful use of intuition and secures the embeddedness of mathematical knowledge; in Chapter 7 their finitism when dealing with sequences and series was juxtaposed with the tutors' contextualised, concise, sophisticated and, possibly, generalisable approaches to testing convergence and divergence. In the context of Continuity and Differentiability this contrast was evident in the cases where the tutors justified their preferences for Proofs Based on First Principles on historical and epistemological grounds. Actually the novices' approach could not be described as a preference because no selecting seemed to be involved in their question-solving. So in this sense the richness of their repertory is what distinguished experts from novices. Demonstrating then the potential of a rich repertory became a necessary aim of the enculturation process.

Through the tensions and conflicts of this enculturation process the novices seem to learn about certain conditions of the didactical and epistemological contract of formal mathematical activity. So, for example in Chapter 7, when it is revealed that the students have been unconsciously assuming the validity of as yet unproved theorems about limits, the novices seem to be learning about an essential aspect of formal mathematical behaviour: that they need to exercise control over their mathematical reasoning in order to avoid subconscious and unjustified decisions.

Tension in the interaction between novice and expert is generated when the difference in their facility to formalise is exposed: while the expert and the novices may agree about the method of approaching a problem, they seem to differ in terms of the implementation of the approach. As a result their interaction evolves into an initiation process during which the students, with variable ease, become familiar with the new notational tools of mathematical formalism– examples of this were given in Chapter 8 in the context of Linear Algebra. Hence their learning becomes a specific struggle for accommodating into this new tool — whose appearance maybe intimidating — the vivid intuitive ideas they have about the solution of the problem.

Another necessary aspect of the enculturation process seems to be reconciliation between intuitive mathematical practices as a way to gain mathematical insight and formal mathematical language as a way to refine and establish these insights rigorously. The novices seem to be deeply perplexed — see Part I — about the status of rigour the various approaches carry: this was extensively exemplified in Chapter 7, in the context of finding and proving limits. The current state of affairs seems to be one of a misunderstanding: the novices are advised to leave behind their school-mathematical way of thinking and start anew by trying to build mathematics on the solid foundations of mathematical formalism. The novices interpret this suggestion in an exaggerated literal manner and turn suspicious about intuitive mathematical practices. As a result they are cognitively torn between what they instinctively know as a powerful way into mathematical insight (intuition) and their desire to be accepted in the culture of mathematical formalism. So, for example, within this schizoid discourse, they perceive the Algebra of Limits as not-formal-enough-hence-avoidable or they refrain from guessing a limit by looking at a graph. The expert's enculturating role then is elevated from the strictly mathematically-topical to the meta-topical, to demystifying not only particular proofs and solutions but also the rules of the game. I stress that crossing through the mathematical contexts explored in this study is the strongly emotional/affective dimension of this demystification and its hard distinction from the purely cognitive one.

The experts then seem to carry the responsibility for convincing the novices about the necessity and the efficiency of various ways of mathematical persuasion. As explained in Part III, Socratic closed questioning, in which unproved theorems were used tacitly, seems to be less convincing than other more openly interactive approaches. Significantly Refutation by Counterexample seemed to bear strong potential of persuasion. However its persuasiveness seems to vary, for both tutors and students, depending on rather personal, and not epistemological, features of the counterexamples used by the expert. In general the expert's invitation into the new and abstract forms of advanced mathematical thinking seemed to be received by the novices with various degrees of readiness. The examples presented in Chapters 6-9, and especially within the more abstract contexts of Linear Algebra and Group Theory, accentuated the very subjective character of the enculturation process.

The perspective in this Part has been to look at cognition as an enculturation process and also to consider potential optimal features of this process. In a sense, this Part can be seen as bridging the perspectives between Part I (cognition from a learning point of view) and Part III (cognition from a teaching point of view). Below I recapitulate briefly some didactical observations made in this study with regard to the teaching of advanced mathematics in the context of tutorials.

Part III Didactical Implications: Observations Related to the Teaching of Advanced Mathematics Derived From the Study of the Novice Mathematician's Cognition

In the previous section, the interaction between the students and the tutor was examined as an enculturation process of the novice into the expert's culture. Here I present briefly some didactical observations linked to the teaching of advanced mathematics. I note that these observations were originally made in the Analysis of the Episodes but, for the sake of conciseness, were not mentioned in Parts III of Chapters 6-9. So this section is a brief cross-topical synthesis of the didactical observations made in the study. In the brackets I refer to the Episodes where these observations originate. The aim here is to juxtapose Interactive Dialectics — the materialisation of the novice's induction to the culture of advanced mathematics as an enculturation process — with Exposition — the materialisation of the novice's induction to the culture of advanced mathematics as an acculturation process (as defined in the Interlude)— and suggest that, on the basis of the evidence in this study, the former has generally more didactical potential than the latter.

As elaborated in Parts I and II, a major difficulty of the novice's enculturation into advanced mathematics is the lack of clarity with regard to the increased requirements of rigour in the new course that the novices have to confront — for instance [6.2] with regard to the knowledge that they are allowed to assume. This raises the didactical question as to how, in this state of uncertainty about the rules, the novices are expected to play successfully the game of advanced mathematical formalism. In some cases [6.4] the tutors prefer to demonstrate these requirements of rigour by presenting exemplars of flawless and rigorous answers — even in cases [6.4] where the student seems to prefer a more conversational style of first refuting his proof and then presenting an acceptable one. In other cases [6.5], dynamic interaction between the tutor and the students proves a fruitful way of refuting the students' flawed approaches but still exposition seems rather inevitable for the presentation of *correct* proofs. In general this type of dialectics, exemplified in cases [6.2, 6.8] where the tutor and a student discretely scaffolded another student's overcoming of an obstacle in a proof, seem to be a successful means to the enculturating end. In the above, a diversity of tutoring approaches is featured. The diversity of the students' observed cognitive needs implies a need for flexible tutoring approaches that adapt to these needs [9.2].

The students were also found to be in difficulty with establishing a creative co-ordination of intuitive and formal practices: very closed questioning and constant reclaiming of control on the part of the tutor seemed to be not very productive as opposed to providing sufficiently detailed explanations that aimed at sustaining the specific students' meaning making process [7.2, 7.4].

The evidence from these tutorials suggests that there are tutors who view interaction with the students at a reflective, meta-topical level as a legitimate part of the tutorial; others view this interaction as a deviation from a carefully predetermined tutoring plan. It is possible that these views entail that the former [7.3, 7.6] engage in exploring and supporting the students' meaning making processes with more zest than the latter [7.4, 7.5]. I note here that there have been cases[0] in which the tutor engages so substantially in the interactive process that, partly as a result of his own uncertainty — about a particular proof of a statement, not the truth of the statement — he yields almost equal didactical control to the students, who then emerge as dynamic perpetrators of these newly balanced dialectics.

[0] Due to limitations of space, it was not possible to present the very lengthy evidence of these cases; these shifts of certainty-hence-control were observed during hourly sessions.

Yielding control to the learner seems to be of considerable didactical potential: for instance, when the tutor [8.1] manages to modify a student's perspective from within, that is by adapting their point of view and challenging it with key prompting questions; or by creating a more debate-friendly learning environment [9.3, 9.5]. Closed questioning on the other hand and highly directive instruction [8.6, 8.10, 9.1] seems to be less efficient and a perpetrator of decontextualised algorithmic behaviour. Directive instruction is mostly based on the expert's prophecies, about, for instance, the simplicity or difficulty of a task [8.6], and these prophecies can prove misleading.

A major didactical point, regarding the potential cognitive danger built in the use of geometrical metaphors as a visual aid for the introduction of new and abstract concepts, was made mostly in the context of Linear Algebra (Vector Analysis): some of the students' difficulties with an abstract perception of vectors (beyond the geometric approach) was [8.2] partly attributed to teaching that focuses on a large number of examples from the line, the plane and the space. However, flexibility on the tutor's part in thinking in the terms of the student's personal metaphors [9.2] seemed to boost a more confident demonstration of thinking on the part of the student. So there seems to be value in the didactical use of metaphorical discourse but only when it does not impede the construction of the intended-by-the-tutor abstract concept image.

In the same vein a less biased use of examples — for instance in the context of convergence and divergence of series where the students' overexposure to examples of convergence seemed to encourage finitist attitudes towards infinite sums — was suggested as a way to curtail some of the novices' prejudices, for instance, with regard to the number of convergent series, or continuous functions, or cyclic groups.

The students were found to be largely engaged in an exploration of the *raisons-d'-être* of newly introduced concepts. To support this exploration, the tutors often decompose the various problematic concepts or theorems (for example: coset or the First Isomorphism Theorem for Groups [9.6]) into their basic elements. This, even though sometimes based on the tutor's preconceptions of what constitutes the problematic elements of the concept or the theorem, seems to be quite efficient. The students seem to appreciate a new concept when it is launched as a useful apparatus, not as an ideal that exists only because of its definition. One of the reasons that spanning sets emerged as paradigmatically problematic in Chapter 8 was the students' tendency to disregard them and favour instead the concept of basis which is of more obvious utility: a suggestion emerged then to introduce the notion of basis before the notion of a spanning set or even to ban the use of the term spanning set altogether from introductory courses. Research focused on the didactics of particular mathematical areas can substantiate and enrich this type of suggestion.

Similar existential inquiries were carried out by the students with regard to some newly introduced theorems: in justifying the importance of these theorems (Lagrange and First Isomorphism Theorem for Groups [9.6]) the tutors often use pragmatic — as opposed to epistemological — arguments. A pragmatic argument was described as an attempt of the tutor to convince the students of the significance of a theorem by repeating that it definitely appears on exam papers because, for instance, it has a 'name attached to it'; an epistemological argument was described as giving an existential rationale to some newly introduced concepts such as justifying the introduction of cosets as a substantial element of studying normality in Group Theory. The latter were suggested as cognitively more powerful, whereas the former were acknowledged as strong motivators.

Part IV Methodological Implications: A Comment on Observation As A Means to Gain Access to Advanced Mathematical Cognition

Minimally participant observation, as described in Chapters 2 and 4 and used in this study, allows natural access to expressions of cognition made by learners in action. How deep or illuminating this access was, relied to a great extent on the participants: the tutors and the students. Below I have collected a few observations with regard to the participants' influence on the study:

- Highly leading discursive practices, such as closed questioning, seemed to impede my attempt to observe the students' thinking because, by channelling the students' responses, the expressions of their thinking come under the clinical control of the tutor and thus lose a great amount of their authenticity. So the sense of the students' difficulties I was making seemed to be derivative of the sense being made by the tutor: by that I mean that the tutor chose a particular direction of closed questioning on the basis of her interpretation regarding the student's difficulty. Some of the closed-questioning material however seems to provide insights on the students' cognition and parts of it have been presented in Chapters 6-9.

- The influence of the students on the quality of the observation seemed to be dependent on the degree of openness with which they participated in

> the tutorials. It turned out that the higher the degree of openness and participation of the student in a piece of data, the higher the probability that this piece of data would be filtered through the final stages of analysis: a selection of these filtered pieces of data, the Episodes, was presented in Chapters 6-9 and were deemed as data in which the students' discourse is rich and at moments transparent as to their difficulties. Certainly, and given the considerations of this study with regard to the accessibility of cognition explained in Chapter 2, this transparency is relative and bounded by the students own explanatory power.

There was a plethora of moments during observation that further elaboration on the students' expressions would have enhanced my potential for subsequent interpretation. Given the non-participant intentions of the study this intervention was not possible — it was more possible in the interviews. In Part V I explain briefly how the naturalistic observations made in this study can be extended to more focused research around the themes elaborated in the analysis.

Part V Embedding the Cross-Topical Synthesis in the Literature Reviewed in Chapter 1. Possible Extensions of the Study.

Parts I-IV are a synthesis of the findings as these were presented elaborately in the analysis in Chapters 6-9. The analysis as well as the topical and cross-topical syntheses were strongly data-driven processes, simultaneously informed by the theoretical framework of the study as outlined in Chapter 1.

As emphasised in the Introduction, the structure of this Chapter reflects the juxtaposition of perspectives on learning presented in Chapter 1. The psychology of the individual learner is clearly influenced by the Piagetian ideas on how the transition to abstract forms of thinking takes place (mostly the notion of Reflective Abstraction as specifically transformed for the needs of Advanced Mathematics by the PME-AMT theories outlined in Chapter 1, concept-image and metaphor construction being the dominant ones). The Vygotskian impact can be seen in the recognition of the strong interdependence between formal reasoning/thought and formal language in the presentation of the findings. The Lacanian psychoanalytical approach has influenced the part of the analytical discourse in this study with relation to the role of the unconscious in learning and its control by the learner as a means for conceptualising and overcoming Epistemological Obstacles. The novice's induction to Mathematical Abstraction has been described as a process of conceptualisation and confrontation of Epistemological Obstacles as well as an enculturation process. The perspective on the novice's learning as an enculturation process was generally drawn from Hall's (transition from informal to technical level) and Foucault's (from the rules of sense to the rules of rationality) archaeologies of knowledge. The perspective on this enculturation as a learning process built upon a didactical contract — the clarification of the conditions of this contract are described here as the ultimate task of the tutor-enculturator — is drawn from the eponymous theory of Guy Brousseau.

The broadness of the theoretical influences of the study — as recapitulated briefly in the above and reviewed in Chapter 1 — as well as its wide scope — as evident in the presentation of the findings, the analysis and the syntheses — allow a range of possible extensions:

- a refinement of findings within one of the mathematical areas presented in Chapters 6-9,
- a refined pursuit of one of the cross-topical themes,
- an embedding of the findings of this study into the methodological discourse within PME-AMT on the search for a unified framework,
- a transformation of the findings of this study in a language understood by mathematics teachers at university as a stepping stone for a piece of research, analogous in scale and content to this study but focusing on the interaction between the learner and the teacher rather than strictly on learning.

Epilogue

Ludwig Wittgenstein wrote about the fundamental mathematical act of *inferring** :

> *When we ask what inferring consists in, we hear it said e.g.: 'If I have recognised the truth of the propositions..., then I am justified to further write down...' — In what sense justified? Had I no right to write that down before? —' Those propositions convince me of the truth of this proposition'. But of course that is not what is in question either. —' The mind carries out the special activity of logical inference according to these laws.' That is certainly interesting and important; but then, it true? Does the mind always infer according to these laws? And what does the special activity of inferring consist in? —* ***This is why it is necessary to look and see how we carry out inferences in the practice of language; what kind of procedure in the language-game inferring is*** *[my emphasis]*
>
> (Wittgenstein 1978 p.43)

This work is an exploration into these 'kinds of procedure' with regard to the novice mathematician. It started as generally as indicated by this quotation and in the process it was refined into the spectrum of themes synthesised in Chapter 10. Further refinement and substantiation of these themes is necessary.

I take *inferring* in the above to mean more than simply *deduce via the rules of logic*. I see it as both reasoning formally and meaning-making. In this study these two acts were seen as the components of starting to act mathematically at an abstract level and they were shown problematic.

Mathematical abstraction 'supplies us with a *new picture*, a *new form of expression*' and 'there is nothing so absurd as to try and describe this new schema, this new kind of scaffolding, by means of the old expression' (p.138). Not because 'the finite cannot grasp the infinite' (p.263) — what is the point of attempting to understand then? — but because a new language is introduced in order to express new meanings. In advanced mathematics a large number of these new meanings are about the multiplicity of representations and the effectiveness of flexibly alternating between these representations. The novice does not always seem to attribute this type of significance to the new language and tries to acquire the new meanings by trying to describe it in terms of the old pictures and the old words. All the new

* Wittgenstein, with his particular interest in mathematics and its psychology, has been a constant source of inspiration in this study. All the references in the Epilogue are his words from (Wittgenstein, 1978)

language is about is translating 'vague ordinary prose' into clear — hence usable — expressions. The novice is, to start with, an applied mathematician: nothing makes sense unless it has a purpose. Concept formation then is the search for 'the limit of the empirical' (p.237) and, this, didactically, possibly implies the necessity to treat learning as an empirical extension.

The above are, simultaneously, findings of this study and challenging questions. I note that if 'we ask these questions at all, this points to the fact that the answers are not ready to hand' (p.133) and because 'philosophical dissatisfaction disappears by our seeing *more*' (p.118) I can only wish for *more* as the way ahead.

Bibliography

Abbott-Chapman, J. 1993. 'Is the debate on quantitative versus qualitative research really necessary?', *Australian Educational Researcher*, 20(1), 49-63

Abkemeier, M. K. & Bell, F. H. 1976. 'Relationships between variables in learning and modes of presenting mathematical concepts.', *International Journal of Mathematics Education in Science and Technology*, 7(3), 257-270

Adelman, C. 1981. *Uttering, Muttering, Using and Reporting Talk for Social and Educational Research*, London: Grant McIntyre

Anderson, G. 1995. *Fundamentals of Educational Research*, London: The Falmer Press

Anderson, J. A. 1994. 'The Answer is not the Solution: Inequalities and Proof in Undergraduate Mathematics', *International Journal of Mathematics Education in Science and Technology*, 25(5), 655-663

Artigue, M. & Viennot, L. 1987. *Some aspects of students' conceptions and difficulties about differentials*. The Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Cornell, Vol.3, 1-7

Ascher, M. 1991. *Ethnomathematics: A Multicultural View of Mathematical Ideas*, Pacific Grove, California: Brooks & Cole Publishing Company

Bachelard, G. 1983/1938. *La Formation De l' Esprit Scientifique*, Paris: Presses Universitaires de France

Balacheff, N. 1986. 'Cognitive versus situational analysis of problem solving behaviours.', *For the Learning of Mathematics*, 6(3), 10-12

Balacheff, N. 1988. 'Aspects of proof in pupils' practice of school mathematics', in D. Pimm (ed.) *Mathematics, Teachers and Children*, London: Hodder and Stoughton, 216-230

Balacheff, N. 1990a. 'Beyond a Psychological Approach: the Psychology of Mathematics Education', *For the Learning of Mathematics*, 10(3), 2-8

Balacheff, N. 1990b. 'Future perspectives for research in the psychology of mathematics education.', in J. Kilpatrick and P. Nesher (ed.) *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education*, Cambridge: Cambridge University Press, 135-148

Barbin, E. 1989. 'I mathimatiki apodeixi: epistemologikes simasies kai didaktika zitimata.', *Omilos Gia tin Istoria ton Mathimatikon*(18, March 1989)

Barnard, T. 1995. *The Impact of 'Meaning' on Students' Ability to Negate Statements*. PME19, Recife, Vol.2, 3-10

Barnes, M. 1988. 'Understanding the Function Concept: Some Results of Interviews With Secondary and Tertiary Students', *Research in Mathematics Education in Australia*(June), 24-33

Batanero, M. C., Godino, J. D., Steiner, H. G. & Wenzelburger, E. 1994. 'The Training of Researchers in Mathematics Education - Results From an Educational Survey', *Educational Studies in Mathematics*, 26, 95-102

Becker, J. R. & Pence, B. J. 1994. 'The Teaching and Learning of College Mathematics: Current Status and Future Directions', in J. J. Kaput and E. Dubinsky (ed.) *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results*, Washington D.C.: MAA Notes No 33: The Mathematical Association of America, 5-13

Bell, A. W. 1976. 'A study of pupils' proof explanations in mathematical situations.', *Educational Studies in Mathematics*, 7, 23-40

Bell, A. W. 1979. 'The learning of process aspects of mathematics.', *Educational Studies in Mathematics10*, 361-3385

Bishop, A. J. 1989. 'Review of visualization in mathematics education.', *Focus on Learning Problems in Mathematics*, 11(1), 7-16

Bishop, A. J. 1991. *Mathematical Enculturation*, Dordrecht / Boston / London: Kluwer Academic Publishers

Bjorkqvist, O. 1993. *Interpretations of double modalities*. PME19, Tsukuba, Vol.1, 107-114

Bliss, J., Monk, M. & Ogborn, J. 1983. *Qualitative Data Analysis for Educational Research*, London: Croom Helm

Boero, P. 1983. 'Storia Della Matematica, Didattica e Formazione dei Concetti Matematici', *Ed. Mat. Suppl. IV-1*, 3-18

Borg, W. R. 1963. *Educational Research: an Introduction*, New York: David McKay Company, Inc

Brousseau, G. 1986. 'Fondements et Méthodes de la Didactique des Mathématiques', *Recherches en Didactique des Mathématiques*, 7(2), 33-115

Brousseau, G. 1989. 'Le Contrat Didactique: Le Milieu', *Recherches en Didactique des Mathématiques*, 9(3), 309-336

Brousseau, G. 1991. 'Foundations and Methods of the Didactics of Mathematics (Themelia kai Methodoi tis Didaktikis ton Mathimatikon)', in A. Gagatsis (ed.) *Themata tis Didaktikis ton Mathimatikon*, Thessaloniki, Greece: Adelfoi Kyriakidi, 61-134

Brown, T., Hardy, T. & Wilson, D. 1993. 'Mathematics on Lacan's Couch', *For the Learning of Mathematics*, 13(1), 11-14

Bruner, J. 1960a. *The Process of Education*, Cambridge, MA: Harvard University Press

Bruner, J. 1960b. *Studies in Cognitive Growth*, New York: John Wiley and Sons

Bullock, J. O. 1994. 'Literacy in the Language of Mathematics', *American Mathematical Monthly*, 101(8), 735-743

Burrell, G. & Morgan, G. 1979. *Sociological Paradigms and Organizational Analysis*, London: Heinemann Educational Books

Burton, M. B. 1988. 'A Linguistic Basis for Student Difficulties with Algebra', *For the Learning of Mathematics*, 8(1), 2-7

Byers, V. & Erlwanger, S. 1984. 'Content anf form in mathematics.', *Educational Studies in Mathematics15*, 259-275

Cajori, F. 1993. *A History of Mathematical Notations*, New York: Dover Publications

Carducci, O. M. 1993. 'Four Elementary Linear Algebra Projects', *PRIMUS*, 3(4), 337-344

Carnap, R. 1983. 'The Logicist Foundations of Mathematics', in P. Benacerraf and H. Putnam (ed.) *Philosophy of Mathematics: Selsected Readings*, Cambridge: Cambridge University Press, 41-52

Chazan, D. 1993. 'High School Geometry Students' Justification for Their Views of Empirical Evidence and Mathematical Proof', *Educational Studies in Mathematics*, 24(4), 359-387

Chevallard, Y. 1985. *La Transposition Didactique*, Grenoble: La Pensée Sauvage Éditions

Clement, J., Lochhead, J. & Monk, G. S. 1981. 'Translation difficulties in learning mathematics.', *American Mathematical Monthly* (April), 286-290

Cobb, P., Yackel, E. & Wood, T. 1992. 'A constructivist alternative to the representational view of mind in mathematics education.', *Journal of Research in Mathematics Education*, 23(1), 2-33

Coe, R. 1992. *Students' understanding and use of mathematical proof.*, Unpublished MPhil Thesis, Cambridge, UK

Coe, R. & Ruthven, K. 1994. 'Proof practices and constructs of advanced mathematics students.', *British Educational Research Journal*, 20(1), 41-53

Cohen, L. & Manion L. 1989. *Research Methods in Education*, London: Routledge

Cole, M. & Bruner, J. S. 1971. 'Cultural Differences and Inferences About Psychological Processes', *American Psychologist*, 26, 867-876

Collis, K. F. 1974. 'The development of a preference for logical consistency in school mathematics.', *Child Development*, 45, 978-983

Coulthard, M. 1985. *An Introduction to Discourse Analysis*, London: Longman

d'Ambrosio, U. 1985. *Socio-cultural Bases for Mathematics Education*, Campinas, Brazil: Unicamp

Dalen, D. B. v. 1966. *Understanding Educational Research*, New York: McGraw-Hill Book Company

Davis, R. B. 1989. *Learning Mathematics: The Cognitive Science Approach to Mathematics Education*, London and New York: Routledge

Davis, R. B. & McKnight, C. 1984. 'The influence of semantic content on algorithmic behaviour.', *The Journal of Mathematical Behaviour*, 3(1), 39-87

Davis, R. B. & Vinner, S. 1986. 'The notion of limit: some seemingly unavoidable misconception stages.', *The Journal of Mathematical Behaviour*, 5, 281-303

Davydov, V. V. & Radzikhovskii, L. A. 1985. 'Vygotsky's Theory and the Activity-Oriented Approach in Psychology', in J. V. Wertsch (ed.) *Culture, Communication and Cognition. Vygotskian Perspectives*, London: Cambridge University Press.

Dee-Lucas, D. & Larkin, J. H. 1991. 'Equations in Scientific Proofs: Effects on Comprehension', *American Educational Research Journal*, 28(3), 661-682

Dijk, T. A. v. 1985. *Handbook of Discourse Analysis Vols 1-4*, London: Academic Press

Dorier, J. L., Robert, A. & Rogalski, M. 1994. *The Teaching of Linear Algebra in First Year of French Science University: Epistemological Difficulties, th euse of the 'meta-level', long term organization*. PME18, Lisbon - Portugal, Vol.4, 137-144

Dormolen, J. V. 1977. 'Learning to Understand What Giving a Proof Really Means', *Educational Studies in Mathematics*, 8(1), 27-34

Dreyfus, T. 1990. 'Advanced Mathematical Thinking', in J. Kilpatrick and P. Nesher (ed.) *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education*, Cambridge: Cambridge University Press, 113-134

Dreyfus, T. 1991. 'Advanced Mathematical Thinking Processes', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 25-41

Dreyfus, T. & Eisenberg, T. 1983. 'The function concept in college students: linearity, smoothness and periodicity.', *Focus on Learning Problems in Mathematics*, 5(3&4), 119-132

Dubinsky, E. 1991. 'Reflective Abstraction in Advanced Mathematical Thinking', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 95-123

Dubinsky, E., Dautermann, J., Leron, U. & Zazkis, R. 1994. 'On Learning Fundamental Concepts of Group Theory', *Educational Studies in Mathematics*, 27, 267-305

Dubinsky, E. & Harel, G. (ed.) 1992. *The Concept of Function: Aspects of Epistemology and Pedagogy*, Washington D.C.: Mathematical Association of America: MAA Notes No 25

Dubinsky, E. & Lewin, P. 1986. 'Reflective Abstraction and Mathematics Education: The genetic decomposition of induction and compactness.', *The Journal of Mathematical Behaviour.*, 5, 55-92

Duval, R. 1991. 'Structure du raisonnement deductif et apprentissage de la demonstration.', *Educational Studies in Mathematics*, 22, 233-261

Ellis, W. D. 1938. *A Sourcebook of Gestalt Psychology*, New York: Harcourt, Brace

Ericsson, K. A. & Simon, H. A. 1993. *Protocol Analysis: Verbal Reports as Data*, Cambridge, MA: MIT Press

Ernest, P. 1991. *The Philosophy of Mathematics Education*, London / New York / Philadelphia: The Falmer Press

Ervynck, G. 1991. 'Mathematical Creativity', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 42-53

Even, R. 1993. 'Subject-matter knowledge and pedagogical content knowledge: prospective secondary teachers and the function concept.', *Journal of Research in Mathematics Education*, 24(2), 94-116

Ferrini-Mundy, J. & Graham, K. G. 1991. 'An overview of the calculus reform effort: issues for learning, teaching, and curriculum development.', *American Mathematical Monthly*, 98(7), 627-635

Fischbein, E. 1982. 'Intuition and Proof', *For the Learning of Mathematics*, 3(2), 9-18

Fischbein, E., Tirosh, D. & Hess, P. 1979. 'The Intuition of Infinity', *Educational Studies in Mathematics*, 10, 3-40

Foucault, M. 1973. *The Order of Things: An Archaeology of human Sciences*, New York: Vintage Books

Fraenkel, J. R. & Wallen, N. E. 1990. *How to Design and Evaluate Research in Education*, New York: McGraw-Hill Publishing Company

Freudenthal, H. 1983. *Didactical Phenomenology of Mathematical Structures*, Dordrecht: D. Reidel Publishing Company

Furinghetti, F. & Paola, D. 1991. *On some obstacles in understanding mathematical texts*. PME15, Assisi, Vol.2, 56-63

Galbraith, P. L. 1981. 'Aspects of proving: a clinical investigation of process.', *Educational Studies in Mathematics*, 12, 1-28

Garfinkel, H. 1968. *Studies in Ethnomethodology*, Englewood Cliffs, NJ: Prentice-Hall

Gay, J. & Cole, M. 1967. *The New Mathematics in an Old Culture*, New York: Holt, Rinehart and Winston

Gerdes, P. 1986. 'How to Recognise Hidden Geometrical Thinking: A Contribution to the Development of Anthropological Mathematics', *For the Learning of Mathematics*, 6(2), 10-17

Gerdes, P. 1988. 'On Culture, Geometrical Thinking and Mathematics Education', *Educational Studies in Mathematics*, 19(3), 123-135

Ghosh, S. & Giri, S. 1987. 'Understanding secondary mathematics: analysis of linguistic difficulties vis-a-vis errors.', *International Journal of Mathematics Education in Science and Technology*, 18(4), 573-579

Ginsburg, H. 1981. 'The Clinical Interview in Psychological Research on Mathematical Thinking: Aims, Rationales, Techniques', *For the Learning of Mathematics*, 1(3), 4-11

Glaser, B. G. & Strauss, A. L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*, New York: Aldine de Gruyter

Glasersfeld, E. Von 1983. 'Learning as a Constructive Activity', in B. a. Herscovics (ed.) *Proceedings of the 5th PME-NA Volume 1*, Montreal: PME-NA, 41-69

Glasersfeld, E. Von 1987. 'Learning as a Constructive Activity', in C. Janvier (ed.) *Problems of Representation in the Teaching And Learning of Mathematics*, Hillsdale, NJ: Erlbaum, 3-17

Glasersfeld, E. Von 1991. *Radical Constructivism in Mathematics Education*, Dordrecht / Boston / London: Kluwer Academic Publishers

Glasersfeld, E. Von 1995. *Radical Constructivism: a Way of Knowing and Learning* London, Washington D.C.: The Falmer Press

Gödel, K. 1947. 'What is Cantor's Continuum Problem?', *American Mathematical Monthly*, 54, 515-525

Gray, E. M. & Tall, D. 1994. 'Duality, ambiguity, and flexibility: a 'proceptual' view of simple arithmetic.', *Journal of Research in Mathematics Education*, 25(2), 116-140

Greeno, J. G. 1989. 'A perspective on thinking', *American Psychologist*, 44(2), 134-141

Greer, B. 1981. 'Cognitive Psychology and Mathematical Thinking', *For the Learning of Mathematics*, 1(3), 19-26

Hadamard, J. 1954. *An Essay on the Psychology of Invention in the Mathematical Field*, New York: Dover Publications, Inc.

Hall, E. T. 1981/1959. *The Silent Language*, New York: Anchor Press, Doubleday

Hammersley, M. 1990. *Classroom Ethnography: Empirical and Methodological Essays*, Milton Keynes - Philadelphia: Open University Press

Hammersley, M. 1993. *Educational Research: Current Issues*, Paul Chapman Publishing Ltd

Hanna, G. 1989a. 'More than formal proof', *For the Learning of Mathematics*, 9(1), 20-23

Hanna, G. 1989b. *Proofs that prove and proofs that explain*. PME13, Paris, Vol.2, 45-51

Hanna, G. 1991. 'Mathematical Proof', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 54-61

Harel, G. 1989. 'Learning and teaching Linear Algebra: difficulties and an alternative approach to visualizing concepts and processes.', *Focus on Learning Problems in Mathematics*, 11(1-2), 139-148

Harel, G. & Tall, D. 1991. 'The general, the abstract, and the generic in advanced mathematics.', *For the Learning of Mathematics*, 11(1), 38-42

Harel, G. & Kaput, J. 1991. 'The Role of Conceptual Entities and their Symbols in Building Advanced Mathematical Concepts', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 82-94

Hazzan, O. 1994. *A Student's Belief About the Solutions of the Equation* $x=x^{-1}$ *in a group*. PME18, Lisbon - Portugal, Vol.3, 49-56

Hilbert, D. 1918. 'Axiomatisches Denken', *Mathematische Annalen*, 78

Hillel, J. & Sierpinska, A. 1994. *On one Persistent Mistake in Linear Algebra*. PME18, Lisbon, Portugal, Vol.3, 65-72

Hitchcock, G. & Hughes, D. 1991. *Research and the Teacher: A Qualitative Introduction to School-Based Research*, London, New York: Routledge

Jacob, E. 1987. 'Qualitative Research Traditions: a Review', *Review of Educational Research*, 57(1), 1-50

Janvier, C. 1987. *Problems of Representation in the Teaching And Learning of Mathematics*, Hillsdale, NJ: Erlbaum

Joseph, G. G. 1992. *The Crest of the Peacock: Non-European Roots of Mathematics*, London: Penguin Books

Kaput, J. J. 1987. 'Representation Systems and Mathematics', in C. Janvier (ed.) *Problems of Representation in the Teaching And Learning of Mathematics*, Hillsdale, NJ: Erlbaum, 19-26

Kaput, J. J. & Dubinsky, E. (ed.) 1994. *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results*, Washington D.C.: MAA Notes No 33: The Mathematical Association of America

Kline, M. 1962. *Mathematics: A Cultural Approach*, New York: Addison Wesley

Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times*, New York: Oxford University Press

Kline, M. 1982. *Mathematics: The Loss of Certainty*, Oxford / New York: Oxford University Press

Kuhn, T. S. 1962. *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press

Laborde, C. 1990. 'Language and Mathematics', in J. Kilpatrick and P. Nesher (ed.) *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education*, Cambridge: Cambridge University Press, 53-69

Lacan, J. 1977. *Écrits: A Selection*, London: Routledge

Lakatos, I. 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge: Cambridge University Press

Lancy, D. F. 1983. *Cross-cultural Studies in Cognition and Mathematics*, New York: Academic Press

Langford, P. 1987. 'Mathematics (Limits, infinity, calculus, concepts of proof).', in (ed.) *Concept development in the secondary school.*, Croom Holm, Beckenham, Chapter 6: 104-167. Advanced Mathematics: 114-116 &132-133.

Lerman, S. 1996. 'Intersubjectivity in Mathematics Learning: A Challenge to the Radical Constructivist Paradigm', *Journal of Research in Mathematics Education*, 27(2), 133-150

Leron, U., Hazzan, O.& Zazkis, R. 1994. *Students' Constructions of Group Isomorphism*. PME18, Lisbon, Portugal, Vol3, 152-159

Leron, U. & Dubinsky, E. 1995. 'An Abstract Algebra Story', *American Mathematical Monthly*, 102(3), 227-242

Lofland, J. & Lofland, L. H. 1995. *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*, California and London: Wadsworth Belmont

Macgregor, M. 1990. 'Writing in Natural Language Helps Students Construct Algebraic Equations', *Mathematics Education Research Journal*, 2(2), 1-11

Malik, M. A. 1980. 'Historical and pedagogical aspects of the definition of function.', *International Journal of Mathematics Education in Science and Technology*, 11(4), 489-492

Markovits, Z., Eylon, B.-S. & Bruckheimer, M. 1986. 'Functions today and yesterday.', *For the Learning of Mathematics*, 6(2), 18-24

Martin, W. G. & Harel, G. 1989. *The role of the figure in students' concepts of geometric proof.* PME13, Paris, Vol.2, 266-273

Martin, W. G. & Wheeler, M. M. 1987. *Infinity concepts among preservice elementary school teachers.* PME11, Montreal, Vol.3, 362-368

Marton, F. 1988. *Phenomenography and the 'Art of Teaching All Things to All Men'*, Gothenburg University: Department of Education and Educational Research

Mason, J. & Pimm, D. 1984. 'Generic examples: seeing the general in the particular.', *Educational Studies in Mathematics*, 15, 277-289

McLeod, D. B. 1987. *Beliefs, attitudes and emotions: affective factors in mathematics learning.* PME11, Vol.1, 170-180

Mead, G. H. 1934. *Mind, Self and Society*, Chicago: University of Chicago Press

Medley, D. M. & Mitzel, H. E. 1963. 'Measuring Classroom Behaviour by Systematic Observation', in N. L. Gage (ed.) *Handbook of Research on Teaching*, Chicago: Rand McNally and Company, 247-328, ch.6

Mercer, N. 1991. 'Researching Common Knowledge: Studying the Content and Context of Educational Discourse', in G. Walford (ed.) *Doing Educational Research*, London, New York: Routledge, 41-58, ch.3

Merriam, S. B. 1988. *Case Study Research in Education: a Qualitative Approach*, San-Francisco, London: Josey-Bass Publishers

Miles, M. B. & Huberman, A. M. 1984. *Qualitative Data Analysis: A Sourcebook of New Methods*, Beverly Hills London New Delhi: SAGE Publications

Miyazaki, M. 1991. *The explanation by 'example' -for establishing the generality of conjectures.* PME15, Assisi, Vol.3, 9-16

Monk, G. S. 1990. *Students' Understanding of a Function Given By a Physical Model.* Personal communication of a draft for an unpublished paper. Used in (Nardi 1992) — see below.

Moore, R. C. 1994. 'Making the Transition to Formal Proof', *Educational Studies in Mathematics*, 27, 249-266

Nardi, E. 1990. *Relations Between the Didactics and the Epistemology of Mathematics (Sxeseis Didaktikis kai Epistemologias ton Mathimatikon)*, Unpublished Essay, Department of Mathematics, Aristotle University of Thessaloniki

Nardi, E. 1992. 'Conceptions of Advanced Mathematical Concepts and Some Cognitive Phenomena Relating to Them. Difficulties, Obstacles and Errors. The Exemplary Case of Function.', in A. Gagatsis (ed.) *Didactics of Mathematics*, Thessaloniki, Greece: Erasmus Publication ICP-92-G-2011/11, 361-389

Nardi, E. 1994. '*A Study of the Novice Mathematician's Conceptual and Reasoning Difficulties in the Encounter With Mathematical Abstraction: A Synoptic Account*', unpublished paper for the transfer to DPhil status. Department of Educational Studies, Oxford University.

Nardi, E. 1995. *The Novice Mathematician's Encounter With Mathematical Abstraction: The Case of Accumulation Point*. European Research Conference on the Psychology of Mathematics Education., Osnabrück, Germany, 54-57

Nardi, E. to appear. 'The Novice Mathematician's Encounter with Mathematical Abstraction: A Concept Image of Spanning Sets in Vectorial Analysis', *Educacion Matematica (Grupo Editorial Iberoamerica, Mexico)*

Neumann, J. V. 1983. 'The Formalist Foundations of Mathematics', in P. Benacerraf and H. Putnam (ed.) *Philosophy of Mathematics: Selected Readings*, Cambridge: Cambridge University Press, 61-65

O'Brien, T. 1972. 'Logical thinking in adolescents.', *Educational Studies in Mathematics*, 4, 401-428

O'Brien, T. 1973. 'Logical thinking in college students.', *Educational Studies in Mathematics*, 5, 71-79

Orton, A. 1983a. 'Students' understanding of differentation.', *Educational Studies in Mathematics*, 14, 235-250

Orton, A. 1983b. 'Students' understanding of integration.', *Educational Studies in Mathematics*, 14, 1-18

Perpenet, J. & Groen, W. 1993. 'A Hint is Not Always a Help', *Educational Studies in Mathematics*, 25(4), 307-321

Piaget, J. 1970. *Genetic Epistemology*, New York: Columbia University Press

Piaget, J. 1975. *The Development of Thought: The Equilibration of Cognitive Structures*, Oxford: Basil Blackwell

Piaget, J. & Inhelder, B. 1963. *The Child's Conception of Space*, London: Routledge & Kegan Paul

Pimm, D. 1987. *Speaking Mathematically*, London: Routledge and Kegan Paul

Pinxten, R. 1987. *Evolutionary Epistemology*, Dordrecht: D. Reidel Publishing Company

Pinxten, R., Dooren, I. v. & Harvey, F. 1983. *The Anthropology of Space*, Philadelphia: University of Pennsylvania Press

Poincaré, H. 1946. *The Foundations of Science*, Lancaster, Pennsylvania: The Science Press

Pólya, G. 1962. 'The Teaching of Mathematics and the Biogenetic Law', in I. J. Good (ed.) *The Scientist Speculates*, London: Heinemann, 352-356

Popper, K. 1959. *The Logic of Scientific Discovery*, London: Hutchinson

Porteous, K. 1990. 'What Do Children Really Believe?', *Educational Studies in Mathematics*, 21(6), 589-598

Presmeg, N. 1986. 'Visualisation and mathematical giftedness.', *Educational Studies in Mathematics*, 17, 297-311

Reid, D. A. 1995. *Advanced Mathematical Thinking*. PME19, Recife, Brazil.

Robert, A. 1982. 'L'acquisition de la notion de convergence des suites numeriques dans l'enseignement superieur.', *Recherches en Didactique des Mathematiques*, 3(3), 307-341

Robert, A. 1991. 'Research in Teaching and Learning Mathematics at an Advanced Level', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 127-139

Rosamond, F. A. 1994. 'The Role of Emotions: Expert and Novice Mathematical Problem-Solving', in J. J. Kaput and E. Dubinsky (ed.) *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results*, Washington D.C.: MAA Notes No 33: The Mathematical Association of America, 81-94

Schoenfeld, A. H. 1987. 'Confessions of an accidental theorist.', *For the Learning of Mathematics*, 7(1), 30-38

Schoenfeld, A. H. & Hermann, D. J. 1982. 'Problem Perception and Knowledge Structure in Expert and Novice Mathematical Problem Solvers', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8(5), 484-494

Sekiguchi, Y. 1991. *Social dimensions of proof in presentation: from an ethnographic in a high school geometry classroom*. PME16, Durham NH, Vol.2, 314-321

Selden, A. & Selden, J. 1987. *Errors and Misconceptions in College Level Theorem Proving*. 2nd International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Cornell University, 456-471

Selden, A. & Selden, J. 1993. 'Collegiate mathematics education research: what would that be like?', *The College Mathematics Journal*, 24(5), 431-445

Selden, J., Selden, A. & Mason, A. 1994. 'Even Good Calculus Students Can't Solve Non-Routine Problems', in J. J. Kaput and E. Dubinsky (ed.) *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results*, Washington D.C.: MAA Notes No 33: The Mathematical Association of America, 19-26

Sexton, R. R. 1988. *Problem Solving Processes Used by Sixth-Grade Students in Solving Routine Mathematics Word Problems*, The University of North Carolina at Chapel Hill

Sfard, A. 1989. 'Transition From Operational to Structural Conception: The Notion of Function Revisited', in (ed.) *Proceedings of PME13, Volume 3*, Paris: 151-158

Sfard, A. 1991. 'On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin.', *Educational Studies in Mathematics*, 22, 1-36

Sfard, A. 1994. 'Reification as the birth of metaphor.', *For the Learning of Mathematics*, 14(1), 44-55

Sfard, A. & Linchevski, L. 1994. 'The Gains and the Pitfalls of Reification: The Case of Algebra', *Educational Studies in Mathematics*, 26, 191-228

Sierpinska, A. 1985. 'Obstacles epistemologiques relatifs a la notion de limite.', *Recherches en Didactique des Mathematiques*, 6(1), 5-67

Sierpinska, A. 1987. 'Humanities students and epistemological obstacles related to limits.', *Educational Studies in Mathematics*, 18, 371-397

Sierpinska, A. 1994. *Understanding in Mathematics*, London / Washington D.C.: The Falmer Press

Sierpinska, A., Kilpatrick, J., Balacheff, N., Howson, A. G., Sfard, A. & Steinbring, H. 1993. 'What is Research in Mathematics Education, and What Are Its Results?', *Journal for research in Mathematics Education*, 24(3), 274-278

Sierpinska, A. & Viwegier, M. 1989. *How & when attitudes towards mathematics & infinity become constituted into obstacles in students?* PME13, Paris, Vol.163-170

Strauss, A. 1990. *Qualitative Analysis for Social Scientists*, Cambridge: Cambridge University Press

Tall, D. 1980. 'The notion of infinite measuring number and its relevance in the intuition of infinity.', *Educational Studies in Mathematics*, 11, 271-284

Tall, D. 1981. 'Intuitions of Infinity', *Mathematics in School*, 10(3), 30-33

Tall, D. 1989. 'The nature of mathematical proof', *Mathematics Teaching*, 127, 28-32

Tall, D. 1991a. *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers

Tall, D. 1991b. 'The Psychology of Advanced Mathematical Thinking', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 3-21

Tall, D. 1991c. 'Reflections', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 251-259

Tall, D. 1992. 'The transition to advanced mathematical thinking: functions, limits, infinity and proof.', in D. A. Grouws (ed.) *Handbook of research on mathematics teaching and learning*, New York: Macmillan, 495-511

Tall, D. & Vinner, S. 1981. 'Concept image and concept definition in mathematics with particular reference to limits and continuity.', *Educational Studies in Mathematics*, 12, 151-169

Tall, D. O. & Schwarzenberger, R. L. E. 1978. 'Conflicts in the learning of real numbers and limits.', *Mathematics Teaching*, 82, 44-49

Tirosh, D. 1991. 'The Role of Students' Intuition of Infinity in Teaching the Cantorian Theory', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 3-21

Tucker, A. 1993. 'The Growing Importance of Linear Algebra in Undergraduate Mathematics', *The College Mathematics Journal*, 24(1), 3-9

Vergnaud, G. 1990. 'Epistemology and psychology of mathematics education.', in J. Kilpatrick and P. Nesher (ed.) *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education*, Cambridge: Cambridge University Press, 14-30

Villiers, M. d. 1991. *Pupils' need for conviction and explanation within the context of Geometry*. PME15, Assisi, Vol.1, 255-263

Vinner, S. 1976. 'The naive concept of definition in mathematics.', *Educational Studies in Mathematics*, 7, 413-429

Vinner, S. 1983. 'Concept definition, concept image and the notion of function.', *International Journal of Mathematics Education in Science and Technology*, 14(3), 293-305

Vinner, S. 1991. 'The Role of Definitions in the Teaching and Learning of Mathematics', in D. Tall (ed.) *Advanced Mathematical Thinking*, Dordrecht / Boston / London: Kluwer Academic Publishers, 65-81

Vinner, S. & Dreyfus, T. 1989. 'Images and definitions for the concept of function.', *Journal of Research in Mathematics Education*, 20(4), 356-366

Vygotsky, L. S. 1978. *Mind in Society: the Development of Higher Psychological Processes*, Cambridge, MA: Harvard University Press

Vygotsky, L. S. 1986. *Thought and Language*, Cambridge, MA: MIT Press

Wilder, R. L. 1981. *Mathematics as a Cultural System*, Oxford: Pergamon Press

Williams, A. 1980. 'Brief Reports: High School Students' Understanding of Proof', *Journal for Research in Mathematics Education*, 11, 239-241

Wittgenstein, L. 1978. *Remarks on the Foundations of Mathematics*, Oxford: Basil Blackwell

Yerushalmy, M., Gordon, M. & Chazan, D. 1990. 'Mathematical Problem Posing: Implications for facilitating student inquiry in the classroom', *Instructional Science*, 19(3), 219-245

Zazkis, R. 1992. 'Theorem-out-of-Action: Formal vs Naive Knowledge in Solving a Graphic Programming Problem', *Journal of Mathematical Behaviour*, 11(2), 179-192

The Novice Mathematician's Encounter With *Mathematical Abstraction:* Tensions in *Concept-Image Construction* and *Formalisation*

Elena Nardi

Appendices
to the thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Linacre College

Trinity Term 1996

Introduction to the Appendices

This booklet contains the Appendices for Chapters 3, 6, 7, 8 and 9. Appendix for Chapter 3 contains a sample of the Pilot Study data (Appendix 3A) and the Chart of Incidents (Appendix 3B) from the Data Analysis of the Pilot Study.

Appendices for Chapters 6-9 are structured as follows:

- Appendix A contains the mathematical questions mentioned in the Episodes. The coding of the questions is given in the Interlude.

- Appendix B contains the answers to the questions in Appendix A, relevant figures, theorems, proofs and in general it provides the mathematical background of the Episodes. The items of Appendix B are numbered according to the numbering of the Episodes. So, for instance, fig.5 in Appendix 6B is the figure mentioned in the Episode of Section (v) in Chapter 6.

- Appendix C contains the Extracts as referred to in the Sections of Chapters 6-9. The Extracts are pieces of data and are descriptions of the Episodes (either in dialogue form or summarised reconstructions of the dialogues). Their numbering follows the numbering of the Episodes in Chapters 6-9: so, for instance, Extract 7.3 is the Episode discussed in Section (iii) of Chapter 7. The Extracts are printed in *italic* format except where the participants' utterances are directly reproduced. I note that, in order to avoid identification, the names of the participants have been altered.

Appendices for Chapter 3

<u>TUESDAY 8 JUNE 1993</u>

14:00-14:30. <u>Lona-Jay</u>

sequences and series exam questions. They are dealing with the convergence of $\Sigma x^n/n!$
L: ...let's use the ratio test.

they write the expression $x^{n+1}/(n+1)!/x^n/n!$
she: what should I write here?
J:(hesitant)...the | | of that.
she: yes, so I can be sure it's positive...
...so the convergence is proved

The next part of the question implies the use of the term-by-term-differentiation theorem

she: this is a theorem that comes up every year in the exam in some form.

By the way this exam question gives them the opportunity to decide the differentiability of e^x only by stating ('precisely') known theorems.
[it is the kind of formulation of an exam question that gives me the impression that the person who set it had a particular solution in mind and having that they tried to pose the question without 'betraying' its secret]
J: ...so $e^x=\Sigma$...J stops and thinks from where to sum...
L says from 1..
she: yes, you can try to differentiate the first term and see what happens to it.
Then with the help of L she writes down $(e^x)'=\Sigma\ ...=e^x$ and they go on with the next part of the question where an f_c is given and they have to prove that its derivative is zero.Silence.
she: you can do that can't you? Just differentiate!
[did they expect something trickier? why didn't they answer such a simple question?]
....she writes it down..
[she writes also a query for later...see below]
she: so if its derivative is zero then the function is...
[typical way of her teaching: unfinished sentences, waiting for the students to complete them]
L:...constant
she: and how do you find which constant is it?
L:...let's try a couple of values for x..
she: and which value is obvious to try?
L:...zero
they try it and then she says: 'it is a constant so you only need to evaluate at one point'
Now she writes down e^x
she: what do you know about it?
[in an hysteron-proteron way she solves the query previously established -it was about positive and negative x]
she: ...not many people would be bothered about it

Appendix for Chapter 3: B

CHART OF INCIDENTS

TUTOR ITEMS

CODE	CONTENT
T.PHIL.TEACH.	T views on mathematics teaching as expressed directly or indirectly during tutorial: *hysteron-proteron* structure of mathematics lessons p1. p7. p25. takes responsibility for student difficulty, p11 display of vulnerability, p21.p28 list of tests-example-apply.
T.FORM.INSTR.PART.	Particular instances that illustrate form of instruction: adjusts pace of teaching to individual needs p1, asks them to remember things from past terms p2 and p14, effective use of 'reverse' questions p6-twice. recognising theorems trick p11, 13, 17, 19, nice twist of student suggestion p11 and p13, optimism p13 and p23, connections-coherence p14. constructs from what they know p14-twice and p21. coordinates p14 and p21. informal proving p18, 22, 23, trial-and-error p23. on fear of delta-epsilon and on what these mean p17 and p26, identifies source of misunderstanding p26, explaining compactness p27.
T.FORM.INSTR.GEN.	Gets carried away in monologue p1,9,21,26, artificial dialectics p5,17,23. improvises all along p11,19,26 time constraints hurry up distress p12,18,brief, dense explanations p17,25, resorts to visualisation p22. closed questioning p7,25, appreciates strong memory p27.
T.BEH.ARB.STUD.	Arbitrary treatment of students' responses: ignores them and goes on her own way p2,3,13,14,15,17,18, gives logically incomplete answers-equal intervals query p3 and 15-twice. leaves student originality 'unexploited' p3,16.
T.BEH.ARB.MATH.	Use of tricks or arguments that appear 'out of the blue'
T.BEH.AFF.RESP.	Affective response to students: humour, neutral in mistakes, makes fun of weaknesses (topical or memory) p2,8, irony p13, wrong solution joke p18. snaps answer to reserve contempt p24.
T.BEH.AFF.ENHT	Affectively formative pieces of advice: 'do on with what you start' p14, setting good example p15, injecting confidence p15. compliments students p18, demystification of theorem p23.
T.SLOG.MATH.	Metamathematical slogans: 'Analysis is about taking limits' p3. proving from first principles p19. 'get the feel of the function' p22.
T.SLOG.RECIPE	Rules of thumb: diagonalisation and eigenvalues p3,6, checking 1-1 p10, proving uniqueness p14, events in probability p19, list of tests on convergence p21,23, how to remember the integral test p21, outlines use of theorem p22, uses of the integral test p23, on additive sigmas p25.
T.LANG.USE.AMB.	Ambiguous use of language: 'lot more work for me' p4, 'straightforward proof' p21.
T.LANG.USE.COLL.	Use of coloquialisms: to serve individual needs p4, to describe technique ('bare hands', 'naive', 'sophisticated' p11, 14, 'alternative', 'polished proof' p14, 'solve this without trouble' p18, 'start from scratch' p19, 'sandwiched' p21, 'stick it to a rectangle' p22, 'nasty' p24, 'things from g(x)...' p26.
T.EX.	Discussion of exam conventions: 'short answers' p6, 'looks like one you can make' p7, on ambiguity p8, negotiating marks p15, exam proof proper p16, on Give/Find p18, 'theorem that comes up every year' p25, 'stating precisely known theorems' p25, uneven measures of rigour p25
T.META.TEACH.	Improved handling of double sigmas p17, improved on $f: R^n \to R$ p18, problem on notation p22.
T.REGRESSION	On $f:R^n \to R$ p18, unnecessary conservatism p21, spoils nice handling p21. anti-constructivism p23.

STUDENT ITEMS

CODE	CONTENT
S.ORIG.MATH.	Original, peculiar or brilliant -not necesarily 'correct'- ideas of students: triangles suggested instead of rectangles in the definition of the Riemann integral p3, trapezia suggested instead of rectangles in the definition of the Riemann integral p4. leap to $\Sigma f(x_i)\xi_i$ p9. f is injective hence d_n is unique p11. on ifinity in sigma p14.
S.LANG.USE	Students' use of language: integral as the 'other way around' of differentiation, asking for 'clear' proof p22.
S.DIFF.TOP.	Students' topical difficulties: (f=0 => f(x)=?) and (f(x)=f(x)+?) and f(x-1)=f(x)-1 p6, vector space of field p7. $f:R^n \to R$ f(x) looks like? p7, specificity of f p10. on free variable p12, on the greatest lower boundary p14, conditioning in Probability Theory p15. factorising x^n-y^n p15. on absolute value p17, on the remainder in the Taylor series p18. on rankA where A is a matrix being the maximum number p13, on the integral test p21, $e_n \to f(x)$ p21, on delta-epsilon definitions p25, on set-family-union and compactness p26. f=c and two values in order to find c p25,27.
S.DIF.LOG.	Students' logical difficulties: applying theorems without checking conditions p7,22,23,27. use of induction dangerous p14. If P is true for all n≥k then it is true for all n p17, resort to memory p23,25.
S.DIFF.SYMB.	Students' symbolic difficulties: i,j notation p9. A-λ instead of A-λI confusion where A is a matrix, x^n and x^r confused p11. double sigmas p17. switching from n to k notation p17.
S.EX.	Discussion of exam conventions: interpretation of exam question p7. confusion p13. use of part i in part ii in exam question p12. question sloppily worded p12. on expected length of answers p13. ambiguity in the formulation of an exam question p19.
S.BEH.AFF.RESP.	Students' response to tutor's mathematical or other behaviour from an affective point of view: 'nasty trick' p10. tacit acceptance of the tutor's arbitrary, unjustified answers p15. coldness p17. dead zone p19, 23. student frustration coming from the tutor's failure to answer question.
S.INTUITION	lim=1 p11,14. d_n decreasing p14. definition of Accumulation Point p22.
S.METAMATH.	Students' metamathematical queries: 'how legitimate is to feel a limit and then prove it?' p16. they like slogans p23. expectations higher from problems requiring long calculations p24. [juxtapose with] frustration from simple relations found in p12. they don't do the simple thing p25.

MISCELLANEOUS ITEMS

CODE	CONTENT
BEAUTY	Instances of fruitful interaction: on the definition of the Riemann integral p4. co-operation p9.
BR-REG	Breaking the rules of minimal-participant observation p10.
SURPR-MATH.	Unnecessary complex formulation of a problem p12, different approach to the integral of 1/xlogx
T-PERC.	On infinity p19.
COMM-OBS	Comments on way of observation p23, 26.

Appendices for Chapter 6

Appendix 6A

CD2.1 - CD2.6

1. i. Show that if x is irrational and a and b are rational then $ax + b$ is irrational, unless $a = 0$.

ii. Prove that between any two real numbers there is an irrational number.

2. If $x_1, x_2, \ldots, x_n$ are real numbers show that

$$|x_1 + x_2 + \cdots + x_n| \leq |x_1| + |x_2| + \cdots + |x_n|$$

3. Given $0 < \theta < 1$ define $\Theta = \{\theta^n : n \in N\}$.

(i) Prove Θ has an infimum and that $\inf \Theta \geq 0$.

(ii) (To show that $\inf \Theta = 0$ assume $\inf \Theta > 0$ and obtain a contradiction.) Assuming $\inf \Theta > 0$ deduce there exists $n^* \in N$ such that $\theta^{n^*} < \dfrac{1}{\theta} \inf \Theta$. Consider θ^{n^*+1} to obtain a contradiction. Deduce $\inf \Theta = 0$.

4. i. If $0 < \theta < 1$ prove that given any $\epsilon > 0$ there exists $n_\epsilon \in N$ such that $\theta^{n_\epsilon} < \epsilon$.

ii. If $\eta > 1$ prove that η^n can be made arbitrarily large by appropriate choice of $n \in N$.

[The results of the previous question may be assumed.]

5. (i) If S and T have suprema prove $\sup(S \cup T) = \max\{\sup S, \sup T\}$.

(ii) Is it true that if S and T have suprema then $\sup(S \cap T) = \min\{\sup S, \sup T\}$? Justify your answer.

6. For $S, T \subset R$ and $k \in R$ define

$$\begin{aligned} kS &= \{ka : a \in S\}, \\ S + T &= \{a + b : a \in S, b \in T\}. \end{aligned}$$

If S and T have suprema prove

i. that if $k < 0$ then kS has an infimum and $\inf kS = k \sup S$.

ii. $S + T$ has a supremum and $\sup(S + T) = \sup S + \sup T$.

CD3.3

i. If x is a real number greater than or equal to zero, show there exists a natural number $\operatorname{int} x$ and a real number $\operatorname{frac} x$, $0 \leq \operatorname{frac} x < 1$ such that $x = \operatorname{int} x + \operatorname{frac} x$.

ii. If r is a natural number and $r > 1$ show that for a natural number t there exists natural numbers $n, a_0, a_1, a_2, \ldots, a_n$, $0 \leq a_i < r$ such that

$$t = a_0 + a_1 r + a_2 r^2 + \cdots + a_n r^n.$$

[Hint: if the proposition is false then there is a smallest natural number which cannot be expressed in the required form.]

iii. Show that the set of accumulation points of $\{s/r^n : n, s \in \mathcal{N}, 0 \leq s < r^n\}$ is $[0, 1]$.

iv. In the number base r, the number

$$a_n a_{n-1} \ldots a_0 \cdot a_{-1} a_{-2} \ldots a_{-m} = \sum_{i=-m}^{n} a_i r^i,$$

where $0 \leq a_i < r$. Show that there is a number in this number base arbitrary close to any real number. (We usually work in base 10 and computers in base 2)

Appendix 6B

Suppose false then $n \le a$
$\Rightarrow$ N is bounded above
$\Rightarrow$ (Compl.Ax) $\exists$ sup$N \in \Re$
$n \le a \; \forall n \in N$
$\varepsilon > 0 \; \exists n_\varepsilon \in N \; a - \varepsilon < n_\varepsilon$

let $\varepsilon = 1$ then $a - 1 < n_1$
$a < n_1 + 1$

fig.1a Jack's proof for the ArchPr

N has a supremum
$(\exists a \in \Re)$ $(n \le a, \forall n \in N)$
$\forall \varepsilon > 0 \; \exists n \in N \; a - \varepsilon < n$
In particular there exists $n_1 \in N$ such that $a - 1 < n_1$. But then $a < n_1 + 1$.

Equivalent to the ArchPr:
If $x, y \in \Re$ and $x > 0$ then $(\exists n \in N)$ $(nx > y)$
$(\forall \varepsilon \in \Re)$ $\varepsilon > 0 \Rightarrow (\exists n \in N)$ $(1/n < \varepsilon)$

fig.1b The tutor's proof for the ArchPr

$\exists x$ which is irrational
we can find a rational s
st $a/x < s < b/x$,
a, b rational

fig.2a Ben's proof for CD2.1ii

Suppose that $a, b \in \Re$ and $a < b$.
Suppose also that all real numbers x such that $a < x < b$ are rational.
We prove that $a < 2a+b\,/3 < a+2b\,/3 < b$.
By our assumption $2a+b\,/3$, $a+2b\,/3 \in Q$

$2a+b\,/3 < 2a+b\,/3 + b-a\,/3\sqrt{2} < a+2b\,/3$ is irrational by part (i) and lies between a and b.

fig.2b The tutor's proof for CD2.1ii

Seeking to prove $|x+y| \le |x| + |y|$

$|x+y| \quad -(x+y) \quad x+y$

$|x| = x \quad x \quad -x$

$|y| = y \quad y$

...

Inductive hypothesis: $|x_1 + \ldots + x_k| \le |x_1| + \ldots + |x_k|$

$|x_1 + \ldots + x_k + x_{k+1}| \le |x_1 + \ldots + x_k| + |x_{k+1}| \le |x_1| + \ldots + |x_k| + |x_{k+1}|$

fig.3 The tutor's proof for CD2.2

given η define $S = \{\eta^n : n \in N\}$
$m \ge \eta^n$
suppose that
$\exists n \in N$ st $n \ge \eta^n \quad \forall n \in N$
$n \in N$, $n+1 \in N$ and $(n+1)^n \in S$

fig.5 Andrew's proof for CD2.4ii

$S \cup T$ must be non-empty and bounded above in order to have a supremum. It is non-empty because S and T are. Also

If $x \in S \cup T$ then $x \in S$ or $x \in T$. So $x \le \sup S$ or $x \le \sup T$

that is in any case $x \le \max\{\sup S, \sup T\}$. So $\max\{\sup S, \sup T\}$ is an upper bound for $S \cup T$.

[he continues and proves it is the lowest upper bound]

fig.6a The tutor's proof for CD2.5i
[reconstruction from recording]

$(x \in \Re)\ (x <_{\neq} 0)$
By the Archimedean Property
$\exists n \in N$ such $x\text{-}1 < n \le x$
Therefore
$0 \le x\text{-}n < 1$
$0 \le \varepsilon < 1$ where $\varepsilon = x\text{-}n$.
Thus $x = n+1$ where $\varepsilon = x\text{-}n$.

fig.8a Jack's proof for CD3.3i

let $a = \sup S \cup T = \sup T$
then a is an upper bound of $S \cup T$ therefore $a > x$ for all x in $S \cup T$
if $a \notin S \cup T$ then $a < \sup S$
if $x \in S \cup T$ then $x \in S$ or $x \in T$ and if $x \in S$ then contradiction

fig.6b Alan's proof for CD2.5i
[reconstruction from recording]

[Cornelia has proved that ksupS is a lower bound]

let $b > ks$, where $s = \sup S$
then $b/k < s$. Therefore $\exists a \in S$: $a > b/k$. Then $ka < b$ and $ka \in kS$.

fig.7 The tutor's proof for CD2.5i
[reconstruction from recording]

Let r be a natural number >1. Show that for any $t \in N$ there exists $n \in N$ and there exist $a_0, \ldots, a_n \in N$ such that $0 \le a_s < r$ (for all s) and $t = a_0 + a_1 r + \ldots + a_n r^n$.

fig.8b The tutor's rephrasing of CD3.3ii

Appendix 6C

Extract 6.1

Context: *This is the beginning of the tutorial. In CD2.1 students Jack and Andrew, says the tutor, have presented 'quite sophisticated arguments in different ways' using the Archimedean Property (ArchPr — see The Episode for two formulations of the Archimedean Property). In the following, the students and the tutor discuss its content, various formulations and proof.*

The Episode:

Formulation and Meaning of the ArchPr

The tutor asks what the ArchPr says.

A1: For any number in the reals there is always a number that is bigger than that.

T1: Well, if I take any number x, then $x+1$ is bigger.

J1: I think you can get as close to x as you like. There's a number... greater than that but you can make it as close to x as... as you like.

T2. The tutor protests: if, for instance their meaning of 'close' *means* ε, *then he suggests taking* $x+\varepsilon/2$. *What they are giving him, he says, is* 'so trivial that is not worth giving it a name'.

A2: Find epsilon you choose just greater than nought and a number that exists. Say you've got a number x and you're choosing an ε, there's always a number which is between x and $x+\varepsilon$... I can't remember.

The tutor asks them to look in their notes. They say their ideas came from various books. He asks them to be more critical when trying to remember things. Jack reads from his notes 'the ArchPr for the natural numbers' *which consists of two equivalent statements. The second of these usually takes the name of ArchPr, says the tutor, while writing it on the b/b (see fig.1b for the first statement):*

$$(\forall a \in \Re)\ (\exists n \in N)\ (a<n)$$

He then asks them what the statement would say if it were false. Andrew says 'that there is an upper bound for the reals'. *The tutor says he does not agree and Andrew changes to* 'for the natural numbers'. *The tutor then briefly talks about the non-Archimedean number-systems in which this holds. On the real line, he stresses, however large are the real numbers you are looking at,* 'wherever you look there is a natural number' *and* 'that is the property of being archimedean'. *He then asks for a proof of the ArchPr.*

Jack suggests using contradiction and the Completeness Axiom and since, as Jack says, it is 'easier if it's on the board' *he presents his suggestion in writing*

J2 *[while he is writing what is in fig1a]*:...*N* is bounded above... by *a*... and also *N* is not empty. So, by the Completeness Axiom, em,... there exists a supremum,... of *N* in the real numbers... so there is a least upper bound for the natural numbers and...?

Jack sounds hesitant. The tutor asks him what he wants to do with the supremum and Jack whispers 'this is strange' *and stops. Then Andrew says:* 'If you take something from it, then there's got to be something which is between that and your supremum. So if you say that alpha is the supremum, there's got to be some *n* which is greater than a-1'. *Jack repeats the definition of the supremum and writes it on the b/b* ('if the supremum equals α then $n \leq \alpha \ \forall n \in N$. Then given any $\varepsilon > 0$, there exists $n_\varepsilon \in N$ such that $\alpha - \varepsilon < n_\varepsilon$') *and stops. The tutor then suggests Jack follows Andrew's suggestion and puts $\varepsilon = 1$. Jack does and then spends a few seconds looking silently at the b/b. Then he says:* 'However this [n_1+1, see fig1a] is in the natural numbers so there is a contradiction'.

The tutor says that he agrees but also that he wishes to comment upon Jack's writing style which was 'presumably copied from the board'. *Andrew points out that this is what the lecturer wrote. The tutor describes this writing style as* 'b/b technique' *which aims at conciseness and at avoiding long sentences. He then comments upon:*

- *Jack's $\exists$ sup$N \in \Re$.. The tutor suggests two ways of writing: either*

'N has a supremum'

or

'$\exists\ \alpha \in \Re$ such that $n \leq \alpha\ \forall n \in N$ and $\forall \varepsilon > 0\ \exists n \in N$ such that $\alpha - \varepsilon < n$'.

The supremum of N is something defined by the notation, he stresses, a function defined on non empty subsets of $\Re$. Writing supN is either nonsense or a specific real number. And it doesn't make any sense to write for instance $\exists\ 27 \in \Re$, he concludes.

• *Jack's use of* n_ε. *Jack says he does not* 'understand this notation'. *The tutor replies that the lecturer was trying to* 'produce a bit of inflection' *and emphasises that* 'you specify epsilon first and then you are choosing n, therefore your choice will have to depend upon epsilon. But that doesn't mean there is a function.' *Andrew then asks whether* 'we shouldn't use this notation with the subscript'. *The tutor thinks they could but it* 'would seem perverse to do so'. *For specific values* $\varepsilon=1$ *it is fine to write* n_1. *The tutor then writes the second statement on the b/b (fig.1b). Jack suggests taking* $\varepsilon=1/\phi$ *and with Andrew they rephrase the property in terms of* ϕ. *The tutor says he agrees but he also stresses that* 'the natural thing is to define ϕ to be $1/\varepsilon$' and thus 'Go from the known to the unknown'. *The second statement of the ArchPr* 'is just a different angle of the ArchPr for the natural numbers' *says Andrew. The tutor calls the second statement a corollary from the ArchPr and points out that* 'it should be called the ArchPr on the reals because it is the embedding of the natural numbers in the reals'.

Extract 6.2

Context: *This is the second half of the tutorial to students Cathy, George and Ben. In the first half they have been doing Linear Algebra which the tutor claims is worth less tutoring time than the more problematic Analysis: Cathy has not delivered a draft with her work to the tutor. When he asks her if she has done any work she replies* 'kind of'. *The tutor turns to CD2.1ii and points out that George and Ben have given different solutions: George's solution is longer than Ben's. The tutor invites Ben to present his solution on the b/b* 'in two lines'.

The Episode:

Ben writes his proof for CD2.1ii on the b/b (fig.2a). The tutor disagrees with the student's claim that a and b are rational numbers. George then intervenes in order to outline his solution: you can always find a rational number between two other rationals and then do this for reals. The tutor agrees that this idea can be the basis for a proof and stresses that part of George's problem in his writing was that he hadn't 'written down the data we are starting from'.

G1: Yeah, because I was trying to prove that there was a rational between any two real numbers... when your real numbers are divided by an irrational... it was a fact that there is a rational between two real numbers... you said it was a fact that there is a rational between two real numbers, I wrote that 'suppose...' and you said it wasn't a supposition, it's a fact... so could I just have written...?

T1: Oh, that's different. You see what you've written is not what you meant.

G2: Oh,... I see yeah...

T2: You see what you've written is precisely what you just said: there exist two real numbers and a rational between them. Sure they do. I can prove that by exhibiting two of them, like zero and 1.

G3: You still want to prove there that this is true for any *a* and *b*.

T3: You want to look for them all. You want to quantify...

The.tutor then explains the problem with Ben's proof: in CD2.1ii they need to start from two arbitrary real numbers. In part i the numbers for which the statement is proved are rational. So part i, at least at this stage, cannot be used. The tutor then presents a proof by contradiction (fig.2b) of which in the end he appears rather critical: it is, he says, a 'bad way to present it because it is not constructive, but it is quicker'. 'But does that prove that there is an irrational between **any** real numbers?'

asks Cathy. The tutor stresses that a and b are arbitrary numbers. She then wonders about part i of the question: 'it seems to assume that irrational numbers exist'. 'They do', *the tutor says, and √2 is one of them, as they have* 'all accepted during his presentation'. *He then concludes that there are other irrational numbers too but it is sufficient to point at the existence of one. Then it is legitimate to assume the existence of rational numbers.*

Extract 6.3

Context: *This Episode is from the same tutorial as Extract 6.1 — where the tutor and the students discussed CD2.1. Now they turn to CD2.2.*

The Episode:

CD2.2 was proved by the lecturer for n=2, say Andrew and Jack: he squared both sides and 'showed triangles'. *The picture with triangles and the squaring, replies the tutor, is the* 'genetic intuition' *for understanding the statement. The principle here however is to* 'develop all the properties of the real numbers from a minimum number of axioms', *he continues. Andrew asks whether the squaring and the picture* 'is not the right thing to do' *and the tutor stresses* 'it's the right way to understand it and remember it, but it's not the right way to prove it'. 'Pictures and triangles have no part' *in a proof where the aim is to develop logically a property for n numbers from what is known on two numbers, he concludes. He then suggests proving the triangle inequality for n=2, namely that* $|x+y|\leq|x|+|y|$, *by taking cases for x and y (fig.3). By the definition of* | |, *this leads to eight possibilities for x and y.*

Andrew sounds hesitant about the 'need' *to write explicitly the eight cases. The tutor starts writing out the table (fig.3) commenting that even though it is* ' painful to write down eight possilbilities' *it is* 'on the other hand a back-to-basics proof'. *He then claims that the eight cases can be reduced to two. Andrew suggests* 'taking the absolute values off' *in the cases x and y are both positive or negative and putting the appropriate sign. The tutor says he agrees. Andrew starts outlining the rest of the cases. The tutor then says he just realised that the least number of cases that could be written down is four. Then by symmetry the eight possibilities would be covered, he adds. Then Andrew has a comment on the triangle inequality.*

A1: Would that inequality do? Because $|x+y|$ is either equal to $|x|+|y|$, or equal to $|x|-|y|$ or a couple more permutations, er, combinations...

T1: Well, the modulus of this is equal to the modulus of that, isn't it *[he points at the various cases for* $|x+y|$*]*?

A2: Yes, it's more equal to than... In fact we seem to generalise to less than or equal to when it is equal to something or in fact to something else. It can't be really in between...

T2: You're discarding a lot of information by merely writing that $|x+y|\leq|x|+|y|$ but...

J1: It is true.

A3: But it's not continuous in the fact that you take a couple of values that range from equal to depending on the value of x and y and keep going down as far as you like.

T3: Well, first of all it's true. Secondly I mean what about the geometric picture?

A4: I don't know... I can't... where is the crossover... is this a geometric instance? It can't take any value...

T4: Well, what about complex numbers? Is this also true?

A5: I thought so...

T5: Indeed it's true of vectors and there the fact is in these other veins you are not throwing a lot out and this is only the special case of real numbers.

The tutor then returns to the proof: looking at the various options for x and y, they deduce that the triangle inequality is true for n=2. Now, says the tutor, they have to generalize to n numbers, by induction. He criticises Jack for starting his inductive proof from n=1 because the triangle inequality 'is a basic assertion for two. It then has some information in it. For one it doesn't'. *He then writes down the inductive hypothesis for n=k and proves the statement for n=k+1 by using the proved statement for n=2 and the associative law for the real numbers. He concludes that he did this in detail* ' only as a matter of style of presentation, but style matters'. *Andrew then has a question.*

A6: In the question there it said that the hypothesis or the statement they were talking about is for $n \geq 2$ because as Jack was saying we can show for one it is true. Does this invalidate the proof?

T6: Well, it doesn't say $n \geq 2$...

A7: You can start by proving that this is true for 1 and...*[the tutor nods]* But you would still have to show that P_2 is true because that helps in the second part.

T7: Well, yes, I am thinking of P_2 as being both what you are calling P_2 and the triangle inequality or the basic fact that $|x+y| \leq |x|+|y|$.

Both students nod. The tutor then suggests they look at CD2.4 (Extract 6.5).

Extract 6.4

Context: *In the discussion preceding this Episode the tutor suggests to student Kelle that in the Probability course he needs to learn the definitions and the formulas better and to* 'work out' *some of them instead of trying to reconstruct them sheerly by memory. In contrast to the demand for* 'mind-blowing' *rigour in Continuity Differentiability, in Probability, she continues, they use a lot of previously unjustified theory. In this sense Probability is easier, she adds. She says that he is quite good at concrete probabilistic questions. She requests of him to prepare a list of* 'specific things' *to ask because in a tutorial he should not expect* 'the whole theory' *to be done for him.* 'This is what the lectures are for' *she adds. The tutor then comments on the ambiguity underlying CD2.1 as to what the students are allowed to assume and in CD2.2 she is criticising his incomplete approach to Mathematical Induction — he has proved the triangle inequality for two real numbers and written* 'and so on by Mathematical Induction'. *Then they turn to CD2.3.*

The Episode:

The tutor comments upon Kelle's written solution: 'you've been pretty much stating what the question asks you to prove'. *So, she asks, how does one prove that a set has an infimum?* 'It has to be bounded below', *replies Kelle. So Θ is bounded below by what? asks the tutor.* 'By zero', *he replies. So inf$\Theta \geq 0$, deduces the tutor. And how do you define the infimum? she asks.* 'The lowest... It is the greatest lower bound', *replies Kelle. So since zero is a lower bound, the greatest of the lower bounds, that is the infimum, must be greater or equal to zero, she concludes. Then Kelle explains what he tried to do.*

K1: We prove that... proving something has an infimum, so I applied that by creating a set,... we first prove that something has a supremum, so this would have a supremum the set of lower bounds and we have to show that the supremum of this would be equal to the infimum of that.

The tutor then outlines her approach: 'turning around' *the Completeness Axiom* ('every bounded above set has a sup' *) leads to* ' every bounded below set has an inf'. *Kelle then asks:*

K2: I was wondering whether I should prove that something which is bounded above has a sup.

The tutor explains that Completeness means that every bounded set has an infimum and a supremum; also that Completeness can be assumed. So he had the 'right idea', *she concludes, but his formal writing was not adequate. She turns to his draft and comments:*

T1:... but how do you know, well you don't know, that alpha is of the form theta to the *k*? *[in his drafts Kelle has called the infΘ, alpha]*

K3: Because this set is defined as theta to the *n*. So, any member of the set would be of the form theta to the *n* and if *n* is *k*..

T2: Yes, but is the inf necessarily of the form theta to the *k*? And does the inf actually have to be in the set?

K4: No, unless it's a max... minimum.

The tutor then stresses that the infimum or the supremum of a set is not necessarily contained in the set (for instance the set {1/n, n∈N} has 0 as its inf and 0 is not an element of the set). She then presents the Approximation Lemma and applies it on CD2.3. Once she has defined n Kelle seems to have grasped her idea for the proof:*

K5: This is the greatest lower bound but this is less than... So this should be... I mean not should be but this got to be the greatest lower bound since ... this is lower than it and it's also in the set.

Thus contradiction is reached and they move on to CD2.4.

Extract 6.5

Context: *This Episode consists of two extracts from the two tutorials mentioned in Extracts 6.1 and 6.4. It is about CD2.4. Kelle's Extract refers to CD2.4i and ii and Jack' s and Andrew's to CD2.4ii.*

The Episode: Extract 6.5i

According to the tutor, Kelle has not proved 'what they want'. *What they want, says the tutor, is to prove CD2.4i for 'as small epsilon as you can get' via CD2.3 and the Approximation Lemma. This is equivalent to proving that $\theta^n \to 0$, she adds. She then asks him whether sequences have been mentioned in the lectures. Kelle says hesitantly yes and, when the tutor asks him how he will prove that $\theta^n \to 0$, he responds reluctantly, referring to sequences, that he doesn't* 'think they are meant to read anything more'. *The tutor then returns to the terminology of the question and asks what is it they have to prove.*

K1: We have to prove that there exists an *n*,... that there exists an epsilon...

T1:... not that there exists an epsilon, no...

K2:... there exists an *n* in...

T2. The tutor shows then on paper that what they want to prove is that starting with any 'kind of slice', *any epsilon, all θ^n will 'eventually lie in between' $-\varepsilon$ and $+\varepsilon$. So, how will he write that? she asks.*

K3: Er,... you can find an element such that if *m* equals *N* this will be less than epsilon, eh for $m>N$ θ^n will be less than...

T3: Yes! We are getting there. Now what is this epsilon, is it some particular epsilon or... what?

K4: It can be any... You can choose..You can find an epsilon such that...

T4. The tutor explains that for every epsilon they are given they should find an n_ε such that $\forall n>n_\varepsilon$ $\theta^n<\varepsilon$. This is something he intuitively knows about but it is imperative he proves for all ε. She then asks Kelle why.

K5: Because it might be greater than that before that, beyond that n_ε tries to be smaller for all the rest of the numbers.

T5: Why?

K6: Because if it's converging they should have been higher...

The tutor says she is not pleased that Kelle is using the convergence of θ^n in his justification because this is what they are trying to prove. She repeats the statement and stresses that it is important to see that because $\theta^{n_\varepsilon}<\varepsilon$ and the sequence is decreasing $\theta^n<\varepsilon$ will be true $\forall n>n_\varepsilon$. 'It's so obvious that it goes to nought' *remarks Kelle. She says that she agrees but they* 'still have to prove it'.

The tutor then turns to CD2.4ii: if $\eta>1$ then what about $1/\eta$? Kelle replies it will be less than 1. The tutor asks him to use part i in order to show that η^n will then be arbitrarily large. Kelle suggests 'inverting', *she nods and writes down the symmetrically inverted statement.*

The Episode: Extract 6.5ii

The tutor comments on the students' drafts: CD2.4i was fine in their drafts but CD2.4ii was a 'disaster'. *The students look surprised and the tutor invites Andrew to the b/b to* 'defend his writing'. *He starts by defining* 'the group of all powers η^n'. *The tutor stops him and asks what he means by a* 'group'. *Andrew then corrects* 'group' *to* 'set of all powers η^n'. *So $S=\{\eta^n, n\in N\}$ and by contradiction Andrew wants to prove that* 'we can find some m which is an upper bound of η^n, in which case we can show that we can't have... it's got to be less than this specific value...'. *The tutor says he is confused.*

A1: If we can find some m for which that's true, this is an upper bound of S, η^n cannot take any value, cannot be large enough... so I suppose that an upper bound for S exists and then we show it can't exist.

T1: OK, so we're going to do it with contradiction.

A2: OK. So... no... Suppose for some m in N... No. Suppose m is an upper bound of S and m is a natural number...

The tutor repeats what Andrew has said so far (assume that there exists $m\in N$ an upper bound for S, that is $\eta^n<m$ $\forall n\in N$ and reach a contradiction) and says that Andrew needs to stress that the existence of m is an assumption and that he is trying to 'demolish it'. *Andrew says he is sorry and writes the above on the b/b (see fig.5).*

A3: Now if n is an element of N, then n+1... has... $(n+1)^n$ is in S... because n+1 is a natural number...

The tutor interrupts and complains he does not understand. Andrew explains that η has been chosen as a natural number. He stops and then he realizes it has not. The tutor points out it can be any number >1. Then Andrew responds:

A4: But eta can take any value, can't we let it be...?

Jack laughs and the tutor stresses that η was given from the beginning and he cannot change it. 'Then I guess I resign' *says Andrew and asks whether in any case the idea of contradiction is* 'along the right lines'. *The tutor asks him to justify his belief by showing some strategy.* 'We really want to find an element of S which is greater than m. Thus m is not an upper bound of S. That's the strategy'. *Then Jack has an idea: since we are given eta and we assume the existence of m we might reach contradiction by trying to prove that η^m is the greatest number in S. The tutor says he disagrees: even if we assume that m is an upper bound for S, η^m is not necessarily the largest number in S, he stresses. Andrew suggests proving that* 'you can actually get as close as you want to η^m but you can't actually get there'. *The tutor reminds them that whenever they have an idea they should be ready to confront his criticisms and suggests two ways for the proof. The first in based on the ArchPr and on using CD2.4i and on defining θ as $1/\eta$. In his presentation the tutor repeats his comment from Extract 6.1 on defining unknown things by means of known things. So θ is defined as $1/\eta$, he stresses. He also corrects Andrew's use of $\leq$ instead of $<$ in the definition of θ.*

In his second suggestion, the tutor says that $\eta>1$ implies that $\zeta:=\eta-1>0$ (Jack says that too). The tutor then asks what can be done with η^n. The students are silent. Then Jack suggests expanding it. The tutor says that he does not disagree but that he had something else in mind: he reminds them they are trying to prove that η^n is getting arbitrarily large. Andrew suggests factorization. The tutor then decides to employ Jack's suggestion. Jack dictates the binomial expansion for $(1+\zeta)^n$ and the tutor says it is strictly greater than $1+n\zeta$, since as Andrew remarks, all terms are positive. But $1+n\zeta$ gets apparently as large as n can be which completes the proof. The tutor reminds them that the inequality $(1+\zeta)^n>1+n\zeta$ was proved last week and so they can use it without employing 'anything as sophisticated as the binomial theorem'. *This sparks off the following discussion:*

A5: Well, if you do something like that would you have to say right in the beginning. Fine, I'll prove this by induction, then by this inequality that was question...

T2: Well, if you've done it before can't you refer back to it? We always go back to theorems and...

A6: Yes, it depends... but in the exams you don't know that we've done it.

T3: I know you've done it! As long as you're clear. You see your job is to produce proofs which you know really are proofs and which are clear to the person you are trying to communicate with. That he can understand these proofs. That's what you have to do.

J1: When we came here they said that we have to wipe out all knowledge of math, out of our mind... and now we start assuming things that we learned since we've been here.

T4: That's the idea, yes. Two weeks ago you knew absolutely nothing. Last week you proved by induction this inequality. So this week...

A7: So we can have the data... all kinds of uses...

The tutor says he agrees: if they have any doubts, in this case the proof via the ArchPr is a 'more direct' *one.*

Extract 6.6

Context: *Extracts 6.6i and ii come from the two individual sessions with students Alan and Connie, students of the same tutor. The reason I present them together is that they both refer to CD2.5 and illustrate the two novices' different but characteristic difficulties with the notions of sup and inf. In Alan's session the tutor is demonstrating what he calls a* 'more attractive' *way to do CD2.3&4 and an alternative-to-Alan's way to CD2.1ii. The latter is also how Connie's tutorial starts. In this it turns out that Connie needs to be reminded of the Archimedean Property, as well as be given a pictorial representation of the proof. After CD2.1ii, in both sessions the students inquire about CD2.5, which is where the Extracts presented here start.*

The Episode: Extract 6.6i

The tutor says he does not approve of Alan's proof (I have reconstructed Alan's proof in fig.6b). He then explains that Alan is interpreting a statement for a particular x in $S\cup T$ ($x\in S\cup T$ means $x\in S$ or $x\in T$, stresses the tutor) as a statement for all $x\in S$. It is, the tutor says, similar to saying that because $f(x)\geq g(x)$ for some x in the domain of f and g, hence $f\geq g$.

The tutor and Alan then agree that CD2.5ii is not true. Alan cannot recall the counterexample he used. The tutor offers a counterexample ($S=\{1,2\}$, $T=\{1,3\}$) and asks Alan to think where exactly he realised that CD2.5ii was not true.

A1: It's true that sup*S* is an upper bound for the intesection but it's just saying that... it breaks down when I prove that it is the least upper bound... because if there is an element in between...

T1: It would look awfully like this again!

A2: Oh, it's the same... probably... so sup*S* is 2...

T2: Hmm... so alpha could go one and a half from there...*[Alan nods]* It's the problem about that you are considering two alternatives for all *x* as opposed to it holding for a particular *x*... I'm not certain...

A3: I'm saying... that *x* here... I think it's this... it's quite difficult to say...

The tutor says there is a 'shakier' *counterexample which requires a bit of* 'trickery'*: choosing S and T such that $S\cap T=\emptyset$ then $S\cap T$ has no supremum. He however suggested the one given above because he thought of it as* 'slightly more convincing'.

The Episode: Extract 6.6ii

Connie says she has problems with CD2.5i so, at her request, the tutor presents CD2.5i (see fig.6a). To have a supremum the union must be non-empty and bounded above. To prove the latter the tutor says he is 'forking out' the cases for $x \in S \cup T$: if $x \in S$ then $x \leq supS$. Similarly if $x \in T$ then $x \leq supT$. Tutor is then interrupted by Connie who points at his writing (reconstructed from the recording in fig.6a) and says about x:

C1: Well, one of them... *x* is in the bit where they...

T1: No, no *x* is somewhere in the union.

C2: Well, then isn't it in both of them?

T2. The tutor gives the definition of $S \cup T$, as the set of all elements in S and T, and, returning to his proof, continues his sentence: that in either case, whether $x \in S$ or $x \in T$, $x \leq supS$ or supT. Connie then interupts again:

C3: So you don't need to consider both?

T3: No, I don't need to. If either of this works then it's enough for me. So in either case $x \leq$max of the two sups.

C4: That's it?

No, he replies, this only shows that $S \cup T$ is bounded above and, since it is non-empty, that it has a supremum. The tutor completes the presentation (see fig.6a) while Connie is nodding. For part ii she says that 'On a second thought I think I had to find a counterexample'. *Then they go on with CD2.6.*

Extract 6.7

Context: *The tutor is demonstrating* 'a more attractive way', *as he says, of CD2.3 and CD2.4 and discusses with student Cornelia the use of counterexamples in CD2.5. Then they turn to CD2.6.*

The Episode:

In CD2.6i Cornelia says she showed that infS exists but she could not prove that it is equal to ksupS. She also showed that ksupS is a lower bound for kS. Therefore what she could not prove is that ksupS is the greatest lower bound of kS. The tutor suggests 'following your nose through the definitions' *(fig.7): let b> ks, where s=supS, and prove b is not a lower bound for kS, namely prove that there exists an x in kS such that x<b. So what happens if we divide b>ks by k?*

C1: You've got *b/k* is greater than *s* in...

T1: No, remember *k* is negative.

C2: It's less than...

T2: Now follow your nose, what do you know about *[b/k]*?

C3: It's got to be the greatest...

T3: No, come on. Sup*S* is the least upper bound for *S*. So what can you say for *b/k* if that's the least upper bound of *S*?

C4: It's hard work...

T4: Yeah, it is. Here is the least upper bound and here is a number smaller than it. What can you say about it?

C5: That must be in the set.

'No, no, not necessarily!' *exclaims the tutor and explains: since S can be a* 'dotty kind of set', *all you can say is that b/k is not an upper bound of S. Therefore there exists an a in S with a>b/k. Cornelia then points out that multiplying through with k gives that ak<b and ak is the x we were looking for.*

The tutor concludes that this is an example of what he means by 'following your nose through the definition', *that is* 'Apply the definition at each different stage'. *In this case, he adds, among all the things one could say about b/k, the useful observation is that b/k is not an upper bound of S. He then concludes: she could have picked up the* *right thing to say if she had focussed on her goal to prove that b cannot be a lower bound for kS.*

Extract 6.8

Context: *The tutor comments on the students' (Jack and Andrew) proof for CD3.2: it was fine but they have been reproducing on paper what is only acceptable as an example of b/b writing technique. Jack asks whether the tutor's proof is a* 'general method' *for dealing with suprema and the tutor describes the Completeness Axiom as a* 'machine, as a function, on sets in which you plug in sets and provided they are bounded above and non-empty, they have a sup'. *Once the existence of sup is confirmed, he continues, they then have to identify the exact number that has the supremum properties. The tutor then stresses that CD3 is a problem sheet that needs special attention and they turn to CD3.3.*

In CD3.3i Andrew, says the tutor, took for granted a breaking down of the natural numbers which had not so far been proved and parts of which are proved in CD3.3ii. Andrew says he thought he could assume this breaking down because it seemed natural to him. The tutor stesses that when they think of things like these they should be more critical: what would be the point of asking them to prove CD3.3ii if they could take it for granted in CD3.3i? The tutor then asks Jack to present his proof for CD3.3i on the b/b (fig.8a).

The Episode:

Jack presents his proof for CD3.3i on the b/b (fig8.a). The tutor interrupts him in order to ask how

$$x-1<n\leq x$$

comes from the Archimedean Property.

J1: It's not quite [coming from the ArchPr] because for any natural number there is a... em,... for any number there is a natural number that follows and...

The tutor agrees that 'this is what the ArchPr says'. *Jack is hesitant (*'So there's also... So if x is any real number then...'*). The tutor points out that there is nothing wrong with what he wrote but he is* 'taking two steps in one' *and that* 'when you are saying 'by the ArchPr' you're slightly misleading your reader or your listener'. *Jack looks still hesitant and the tutor points out it's* 'a little more' *than the ArchPr. Andrew then says that $n\leq x$ means* 'that the natural numbers are not bounded by the real numbers'. *The tutor nods in agreement with Andrew but is still leaning towards Jack. He then stresses:*

T1: So there is a natural number greater than x-1 and then what do you do? *[Silence]* There are natural numbers greater than x-1.

J2: So I'm taking the least.

The tutor says he agrees. Andrew wants to know why they need the least nmber greater than x-1. The tutor points out that the least number is still bigger than x-1 and Jack says that 'you can't take any natural number bigger than that. You have to take the next one along'. *He completes his writing om the b/b (fig.8a). The tutor says he thinks it is an excellent proof but points out that Jack did not need to introduce the name ε, since the question provides the name fracx. He then concludes: the point in this question is not merely prove the existence of these numbers but also to manufacture them.*

For CD3.3ii the tutor suggests rephrasing the long sentence in the problem sheet that has been causing problems with its complexly quantified formulation: quantifiers is, he says, a great source of difficulty at this stage (see fig.8b for the rephrasing suggested by the tutor). Jack then says this is a question about number bases and asks whether 10 has been an historically arbitrary choice of a base. The tutor then talks about the history of base-10 numbers. He then says he approves of the use of Mathematical Induction in their proof even though it was not 'very efficient'. *There is a quicker way, he suggests.* 'Do you go straight to the general case?' *asks Jack. The tutor suggests: note that by CD3.3i* $int(t/r)=t_1$ *is a natural number. By the inductive hypothesis this can be written as in CD3.3ii. Then* $t=rt_1+rt_0$ *is the analysis of t, requested in the question. He then suggests talking about Accumulation Points which is a crucial concept in CD3.3ii.*

Appendices for Chapter 7

Appendix 7A

CD4.1

Which of the the following functions have a limit at 0

i.	$\dfrac{x}{\lvert x\rvert}$,	$x \neq 0$,	ii.	$\dfrac{x^2}{\lvert x\rvert}$,	$x \neq 0$,
iii.	$e^{-1/x}$,	$x \neq 0$,	iv.	$e^{-1/x}$,	$x > 0$,
v.	$x \sin 1/x$,	$x \neq 0$?			

Justify your answers.

CD5.1

Define int$x$, $x \in R$ to be the greatest integer n such that $n \leq x$. Determine the points of continuity of

i.	int$x$,	$x \in R$	ii.	x int$x$,	$x \in R$
iii.	int$(\sin x)$,	$x \in R$	iv.	int$(1/x)$,	$x \in R \setminus \{0\}$.

CD7.1

Define the derivative of a function.

Prove that the derivative of the function $f(x) = x$, $x \in R$ is given by $f'(x) = 1$, $x \in R$ and the derivative of $g(x) = 1/x$, $x \in R \setminus \{0\}$ is given by $g'(x) = -1/x^2$, $x \in R \setminus \{0\}$.

Prove that the derivative of $h(x) = x^n$, $n \in Z$, $x \in R \setminus \{0\}$ is given by $h'(x) = nx^{n-1}$, $x \in R \setminus \{0\}$.

CD7.2

State the chain rule for differentiation of a function of a function.

Find the first and second derivatives of

$$f(x) = \begin{cases} e^x & \text{when } x < 0 \\ x+1 & \text{when } x \geq 0 \end{cases}, \quad g(x) = \begin{cases} 0 & \text{when } x < 0 \\ e^{-1/x} & \text{when } x \geq 0 \end{cases},$$

when they exist.

[You may assume that e^x is given by the series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ and is its own derivative]

B6

For each integer $n = 1, 2, 3, \ldots$, the function $f_n : \mathbb{R} \to \mathbb{R}$ is defined by

$$f_n(0) = 0 \quad \text{and} \quad f_n(x) = x^n \cos(1/x^2) \text{ if } x \neq 0.$$

Determine for which n

(i) f_n is continuous at 0;

(ii) f_n is differentiable at 0;

(iii) f_n has a derivative which is continuous at 0.

B7

(a) Find the number of real roots of the equation $x^5 - 5x + 10 = 0$.

(b) Find the image of the function $f : (0, \infty) \to \mathbb{R}$ defined by

$$f(x) = x^{-1}e^x, \quad x > 0.$$

Find also an interval I such that f has a strictly decreasing continuous inverse f^{-1} which maps $\operatorname{Im} f$ onto I.

B10

Suppose that the real-valued function f has a continuous $(n-1)$th derivative on the real interval $[a, a+h]$, and an nth derivative on $(a, a+h)$ (where $n \geq 2$). Taylor's Theorem states that there then exist real functions P and R and a real number θ, $0 < \theta < 1$, such that $P(h)$ is a polynomial in h of degree $n-1$ and

$$f(a+h) = P(h) + R(h, \theta).$$

Write down formulas for $P(h)$ and $R(h, \theta)$.

(a) Find θ explicitly when $f(x) = x^3$ and $n = 2$.

(b) The real-valued function g is defined and has a second derivative on $[0, 2]$. Also $|g(x)| \leq 1$ and $|g''(x)| \leq 1$ for all x in $[0, 2]$. By considering the Taylor expansion about x, or otherwise, prove that $|g'(x)| \leq 2\frac{1}{2}$ for all x in $[0, 2]$.

SS7.1

(From an old Mods paper.) Evaluate the following infinite sums, giving reasons for your answers:

$$(i) \quad \sum_{r=1}^{\infty} \frac{1}{r(r+k)}, \quad \text{where } k \text{ is an integer, } k \geq 1,$$

$$(ii) \quad \sum_{r=1}^{\infty} \frac{1}{r(2r+1)}, \qquad (iii) \quad \sum_{r=-\infty}^{\infty} e^{-|x+r|} \quad (0 \leq x \leq 1), \qquad (iv) \quad \sum_{r=1}^{\infty} r^2/3^r.$$

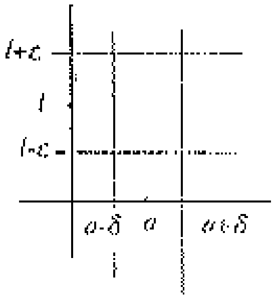


fig.1a

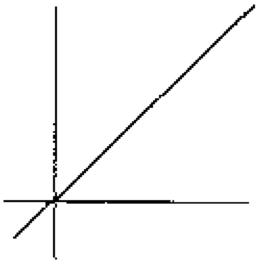


fig.1b $f(x)=x$

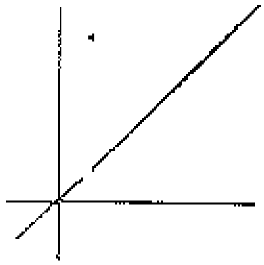
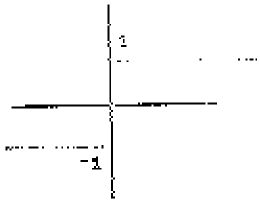


fig.1c $f(x)= \begin{cases} x, & x\neq 1 \\ 27, & x=1 \end{cases}$



$f(x)= x/|x|,\ x\neq 0$

fig.1d

The tutor's proof.

Let $l\in \Re$. We want to prove that l is not a limit of $f(x)$ if $x\to 0$.

Try: $\varepsilon := 1/2$. Whatever positive value δ has, take

$$x_0 := \begin{cases} 1/2\delta, & \text{if } l\leq 0 \\ -1/2\delta, & \text{if } l>0. \end{cases}$$

Then $|x_0-0|<\delta$ but nevertheless $|f(x)-l|>\varepsilon$.

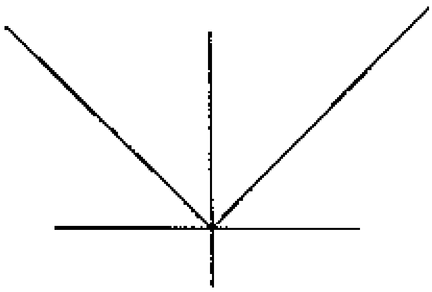
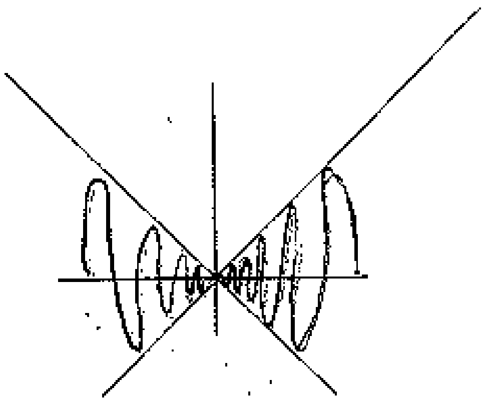


fig.1e $f(x)=x^2/|x|$



Suppose that $\varepsilon>0$. Take $\delta:=\varepsilon$

If $|x|<\delta$ then certainly $|x\sin 1/x|<\varepsilon$.

Hence $\lim_{x\to 0} x\sin 1/x=0$.

fig.1f $f(x) = x\sin 1/x$

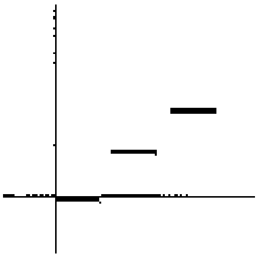


fig2a (i) int$x$

fig2a (ii) $x$int$x$

f is continuous at x_0 if $\lim_{x\to x_0} f(x)=f(x_0)$, that is,
$\forall \varepsilon>0\ \exists \delta>0$: $0<|x-x_0|<\delta \Rightarrow |f(x)-f(x_0)|<\varepsilon$

fig2b

(i) $f(x)=$ int$x$

$x_0 \notin \mathbf{Z}$. Then $x_0 \in (n, n+1)$ for some $n\in N$. Then $f(x_0)=\text{int}x_0=n$.
Then $\forall x\in(n, n+1)$ $f(x)=\text{int}x=n$ and for $\delta=\varepsilon$ $\lim_{x\to x_0} f(x)=f(x_0)$,
that is $\lim_{x\to x_0}\text{int}x=n$ for $x_0\in(n, n+1)$.

$x_0\in \mathbf{Z}$. Then $x_0=n$ for a $n\in N$. Then $f(x_0)=n$.
For x_0-1/k, $f(x_0-1/k)=n-1$. So $|f(x_0-1/k) - f(x_0)| =1$
which is not $<\varepsilon$, $\forall\varepsilon>0$. For instance, if we had chosen $\varepsilon=1/2$ for $\delta=1/k$ the inequality would not hold.

So int$x$ is continuous everywhere except at the integers.

(ii) From (i) we know that int$x$ is continuous everywhere except at the integers. We also know that x is continuous everywhere. Therefore the product $x$int$x$ is continuous everywhere except at the integers.

(iii) Continuous except where sin$x$ is an integer, that is 0 or 1.

(iv) Similarly to (ii), for the composition int$(1/x)$: continuous everywhere except where $1/x$ is integer.

fig2c

$f_r = \{$ $1,\ 1/r+1 < x < 1/r$; 0, elsewhere

where $r \in N$: $\lim \sum f_r \neq \sum \lim f_r$

fig.3

The *Fourier Series* of a function f is
$1/2a_0 + {}_{k=1}\sum^{\infty} (a_k \cos kx + b_k \sin kx)$

where
$a_k = 1/\pi \ {}_0\int^{2\pi} f(x) \cos kx dx$
and
$b_k = 1/\pi \ {}_0\int^{2\pi} f(x) \sin kx dx$

If f satisfies *Dirichlet's Condition*, then it converges to
$1/2\ (f(x+) + f(x-)$
which, in case f is continuous, is equal to $f(x)$.

fig.4

(i) f_n is continuous at zero $\forall n \in N$ because
$|x^n \cos 1/x^2| < |x^n| \to 0$ as $x \to 0$, that is $\lim_{x\to 0} f_n(x) = f_n(0)$.

(ii) f_n is differentiable at zero $\forall n \geq 2$, because
$\lim_{x\to 0} f_n(x) - f_n(0) / x - 0 = \lim_{x\to 0} x^{n-1} \cos 1/x^2 = 0 = f'_n(0)$.
For $n=1$ $f'_1(x) = \cos 1/x^2$ oscillates.

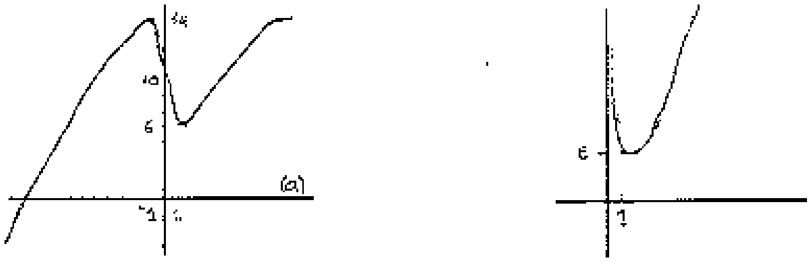
(iii) $f'_n(x) = nx^{n-1} \cos 1/x^2 - 2x^{n-3} \sin 1/x^2$

For $n \geq 3$ $\lim_{x\to 0} f'_n(x) = f'_n(0)$, therefore f'_n is continuous at zero.

For $n=2$ $f'_2(x) = 2x \cos 1/x^2 - 2/x \sin 1/x^2$, which tends to infinity.

For $n=3$ $f'_3(x) = 2x^2 \cos 1/x^2 - 2 \sin 1/x^2$, which oscillates.

fig.5 B6



The Intermediate Value Theorem.

If a real function f is continuous on an interval $[a, b]$ then it takes all the values between $f(a)$ and $f(b)$.

The Inverse Function Theorem

If f is continuously differentiable and strictly monotonic on an interval, then it possesses an inverse function on this interval.

(a) Consider $f(x) = x^5 - 5x+10$. Then $f'(x) = 5x^4-5 = 5(x^2+1)(x-1)(x+1)$, therefore the critical points of f are -1 and +1. $f''(x) = 20x^3$, that is $f''(1) = 20$ and $f''(-1) = -20$. So f has a minimum at 1 and a maximum at -1 so, by the *Intermediate Value Theorem*, it must have one root for some $x<-1$.

(b) $f(x) = e^x/x$, $x>0$. Then $f'(x) = e^x/x\,(1-1/x)$ which has a root at $x = 1$. $f''(x) = e^x/x\,(2/x^2 - 2/x+1)$, therefore $f''(1) = e > 0$. So f has a minimum at 1. Because f is continuous Imf is an interval and given that f is not bounded above ($\lim_{x\to\infty} f(x) = \infty$), Im$f = (e, +\infty)$. Also, because f is strictly decreasing in $I = (0,1)$, by the *Inverse Function Theorem*, we can define f^{-1} which maps Imf onto I.

fig.6 B7

Taylor's Theorem

If $f: \Re\to\Re$ has its $(n+1)$-th derivative on an interval $[a, a+h]$ then

$$f(a+h) = f(a) + hf'(a) + 1/2\, h^2 f''(a) + ... + 1/n!\, h^n f^{(n)}(a) + R(h, \theta),$$

$$\text{where } R(h, \theta) = 1/(n+1)!\, h^{n+1} f^{(n+1)}(\theta)$$

(a) For $f(x) = x^3$, $[a, a+h] = [0, 1]$ and $n+1 = 2$, it is $\theta = 1/3$

(b) Let $x\in[0,2]$. Then take the Taylor expansion of g in $[x, x+h]$ where h is such that $[x, x+h] \subseteq [0,2]$. This gives $|g'(x)| \le 2/h + h/2$ which attains its maximum at 2 1/2 for $h=1$ or -1. Then $[x, x+h] \subseteq [0,2]$ if we choose $h=1$ for $0\le x<1$ and $h=-1$ for $2\ge x>1$. So $\forall x\in[0,2]$ we can choose x+h such that the Taylor expansion of g in [x, x+h] gives $|g'(x)| \le 2\ 1/2$.

fig.7 B10

Appendix 7C

Extract 7.1

Context: *It is the beginning of the tutorial. The tutor and students George and Cathy work on CD4.1. The tutor reads aloud CD4.1 and asks for a definition of limit.*

The Episode:

The tutor asks the students to give a definition of limit. Cathy sounds hesitant and asks the tutor whether he wants it 'properly defined'. *The tutor says yes and Cathy then turns to George. The tutor leaves for a few minutes and asks them to decide on a* 'beautiful definition' *of what a limit is. After a couple of minutes of whispering George writes on the b/b:* *Formal Definition of Limit I*

$$\lim_{x \to a} f(x) = L \text{ given } \varepsilon > 0$$
$$|f(x) - L| < \varepsilon$$
(1)

The tutor stresses 'it cannot always be $<\varepsilon$' *and George then adds below (1)*

$$\text{there exists } \delta(a, \varepsilon)$$
$$\text{whenever } 0 < |x - a| < \delta.$$
(2)

The tutor then asks for 'the picture of this new caption'. *George replies:* *Intuition and Limit*

G1: When you...as your...x approaches a where the limit is then...approaches that point...then the difference between that $f(x)$ and the limit is...you could...it is smaller than any epsilon...well not any epsilon...epsilon you...

The tutor then turns to Cathy and asks for her 'intuitive concept'.

C1: Well, I think...I mean it doesn't really give me one. I mean I don't understand what delta is supposed to be. Is it a number or is it a function or what?

T1: Number, number.

C2: Number. In that interval?

T2: What interval?

C3: a comma epsilon.

T3: Is that the notation the lecturer has used?

The students nod. The tutor sounds quite surprised with Cathy's interpretation and 'coming back to what George said' *he explains that all they need to say is that whenever x approaches a, f(x) approaches L. F(x), he explains,* 'doesn't bounce off to all sorts of different directions'. *To* 'pin down' *the notion of* 'approaches', *or* 'close to', *we use ε. Depending on how accurate we want to be, how close we want to get, we take a small ε. Then* 'the smaller the epsilon is the harder is to get that close'. *Cathy then asks* 'What does delta depend on?'.

Formal Definition of Limit II

The tutor turns to the b/b and rewrites the definition:

$$L = \lim_{x \to a} f(x) \text{ if for any } \varepsilon > 0$$
$$\text{there exists } \delta > 0 \text{ such that whenever}$$
$$0 < |x-a| < \delta,$$
$$|f(x) - L| < \varepsilon$$
$$(3)$$

stressing that he wishes to delete $\delta(a, \varepsilon)$ which has 'given rise to concern'.

He also presents the definition in a drawing (fig.1a). Then Cathy asks:

The Meaning of the Formal Definition of Limit

C4: Well, what would happen if it didn't exist, this delta? I mean the way it's said kind of...I mean if there wasn't the limit you wouldn't be able to do that.

The tutor says that her question is about the case where L is not the limit and asks them to turn to CD4.1. The question is about, he says, whether there exist a number L such that $\lim_{x \to 0} f = L$. Namely, such that (3) on the b/b is true. Cathy asks him to repeat and he does. Discussion then is as follows:

C5: So say I am given an epsilon and we want *f(x)-L* to be this epsilon and you...then there would be a delta for which *f(x)-L* is always less than epsilon.

G2: It's that you have to find the delta to...

T4: Hold on. This is the definition...

C6: So it could never be true that if *L* existed, then you couldn't find a delta for any given epsilon...I mean...

T5: Well, you are trying to deny a definition, aren't you?

C7: Deny a definition?

The tutor stresses that this is the definition of continuity: 'If we cannot prove that this condition holds then we have one of two possibilities: either the function doesn't have a limit at L or we're just being unable to prove it'. *George then is concerned about the dependence of δ on ε and the tutor explains that dependence here is taken in the general sense (that is given an ε beforehand, one should be able to find a δ) and does not suggest the existence of a function between δ and ε. He suggests looking at an example. Cathy then adds about the definition of a limit:*

C8: Is it alright saying that when x is close enough to a, then $f(x)$ is close enough to L?

The tutor enthusiastically agrees and asks them to 'prove the conjecture that if $f(x)=x$ then $\lim_{x\to 1}f=1$'. *The students are silent and the tutor draws fig.1b. He then asks them to identify a δ for $\varepsilon=10^{-472}$. Cathy hesitantly suggests $\delta=\varepsilon$. The tutor substitutes $\delta=\varepsilon$. It works. George then points out that* 'it normally works to choose your delta in terms of that epsilon' *and Cathy that* 'then you also choose your delta so that...you...depending on the function'. *The tutor adds that if he had chosen a more complicated function, the choice of delta would be probably a more complex process. Then Cathy asks:* *The Mechanism of the Definition. $f(x)=x$*

C9: Why should we take limits? I mean if we put 1 in there then we see what value the function takes and...

The tutor turns to CD4.1i and explains that the question is not to define the limit of the function at zero but to prove whether it exists or not; not to find out whether the limit is equal to the value of the function at zero. For instance if for the function in fig.1c the $\lim_{x\to 1}f$ exists but it is not equal to $f(1)$. The tutor points out that the students have a 'prejudice for' *continuous functions.* *CD4.1i*

Cathy then turns to CD4.1i and asks whether the limit is one. The tutor stresses that what they care about here is to find the limit and not to find out whether the function is continuous. He asks the students to simplify the definition of the function in CD4.1i and Cathy presents it as the piecewise function

$$f(x)= x/|x| = \begin{cases} 1, & \text{if } x>0 \\ -1, & \text{if } x<0 \end{cases}$$

The tutor draws fig.1d and says it is unfortunate that their first question is a hard one. They are silent and he asks them to see intuitively that there is a different limit from the left than from the right and so the limit does not exist. The students are still silent and the tutor eventually presents a proof (fig.1d). He writes down the full negation and asks them what ε he should take. After a long pause Cathy points at the b/b, at the beginning of the tutor's proof and asks whether 'this is in terms of epsilon'. *The tutor insists that the picture suggests an ε and, after another long pause, he* *proposes ε=1/2 and completes the presentation (fig.1d). Cathy in the end asks:* 'does *L* have to be -1 and 1?'. *The tutor stresses it is* 'crucial that it doesn't have to be'. *Cathy then asks if the proof works for L>1.*

G3: No, because they would be too far away from anything they would know...

The tutor adds that it is even easier then to define δ and stresses that the proof has to cover every L∈ℜ. George then points out that the proof also works for ε=1 and the tutor replies that in fact he had been a bit too cautious when he chose ε=1/2. The students look quite preoccupied. The tutor emphasises that this was a hard question to start with because 'proving that something does not have a limit is always one stage more complicated than proving it does have a limit'.

Similarly to CD4.1i, they simplify the definition of $x^2/|x|$ *to* CD4.1ii

$$f(x)= x^2/|x| = \begin{cases} x, & \text{if } x>0 \\ -x, & \text{if } x<0 \end{cases}$$

and draw its graph (fig.1e). The tutor then asks Cathy to prove that $\lim_{x\to 1} f=0$. *Cathy responds:*

C10: You should prove that $f(x)$ gets to 0, as x gets to 0. And you want me to prove the limit of that?

T6: What do you mean? That's what we've been doing!

C11: Are we?

The tutor says that this is what they do and Cathy then suggests they 'want to choose a delta'. *She sounds puzzled about choosing δ. She looks at the definition of limit and rewrites it in terms of CD4.1ii:*

[Suppose] ε>0. Let δ=ε
[then] so if 0<|x|<δ
then |f(x)-L| = |x| <δ=ε

(4)

The tutor corrects some of her writing [words in brackets above] and George suggests δ=ε noticing that δ depends on the function. The tutor mentions various cases of dependence between f and δ. In one of his examples he refers to oscillation.

George then mentions CD4.1v. The tutor suggests they find out about the existence of the limit in CD4.1v and again asks for a picture of the function. George says that sinx is between -1 and +1 but in general the two students do not contribute to the construction of the drawing. As a result the tutor stresses that 'it is something they could have thought of from their A-levels'. *He then completes the fig.1f and asks for their guess about the limit. Cathy suggests zero and he agrees. Given* $\varepsilon>0$ *what* δ *then could be? Cathy asks whether* δ *should be less than* ε *and suggests* π/ε. *The tutor completes the proof (fig.1f) saying that it is much simpler to choose* $\delta=\varepsilon$. CD4.1v

The tutorial concludes with the tutor's comments upon another question from the same problem sheet which involved the evaluation of some limits. The tutor says that George 'had the right ideas but didn't express them correctly'. *Cathy says that she didn't have a lot of ideas and the tutor reminds her that she* 'did when she was at school' *and that she shouldn't forget about what she learned at school. She then asks about the algebra of limits:* 'but isn't it imprecise?'. *The tutor says he wants to make clear it is not. He then recommends that, in order to find limits, the students should use inequalities that would help them simplify algebraic expressions.* *Guessing Limits, A-Level Approach and the Algebra of Limits*

Extract 7.2

Context: *In the beginning of the tutorial the tutor stresses that student Kelle has* 'started doing better' *with δ–ε definitions. Still there are problems with his formal writing. Also in his written solutions he has used the theorem*

$$\lim (f+g) = \lim f + \lim g$$

which they haven't proved yet. The tutor proves it by contradiction. Then they turn to CD5.1 which Kelle could not do.

The Episode:

CD5.1i: the tutor suggests 'drawing a picture' *of intx (fig.2a). Intx is* 'discontinuous at integers' *says Kelle. She agrees and asks him to prove that*

$$\lim_{x \to x_0} f(x) = f(x_0) \text{ or the various values of } x_0$$

First for $x_0 \notin Z$. Then Kelle replies:

K1:We have to define an epsilon...define a delta such that *x* minus the limit point..*f*(*x*) minus the limit point is less than any epsilon.

The tutor explains that for $x_0 \in (n, n+1)$, according to the definition of continuity (fig.2b) and the definition of intx, $f(x_0)=n$. What is then $f(x)$? she asks. 'Nought', *says Kelle but changes to $f(x)=n$ when he looks at the drawing, where n is the nearest, smaller integer than x. The tutor then asks him to exploit the fact that*

$$\forall x \in (n, n+1)\, f(x)=n$$

and Kelle claims that 'x has to be less than 1 and greater than zero'. *To the tutor's objection he changes his mind to* 'between *n* and *n*+1'. *The tutor completes the proof (fig.2c) for $x_0 \neq n$, $n \in Z$.*

K2: Can we then say that f is continuous except the integers?

'That's what we proved' *replies the tutor. She then asks him to prove discontinuity at $x_0=n$ (fig.2c) and Kelle suggests using* 'side limits'. *She nods with approval but* 'prefers using the definition': 'this way is slightly better because you're not quoting big theorems'. *She recommends starting proofs always* 'first with intuition and pictures and then with definitions and theorems' *(fig. 2c).*

The tutor stresses that CD5.1i will help them prove ii-iv. In CD5.1ii Kelle looks at the graph of intx and notices that intx=0 for x∈[-1, +1]. She agrees, draws a picture of CD5.1ii (fig.2a) and claims that f will then be discontinuous at the integers apart from zero. She suggests two ways to prove this: the definition or previous results from the algebra of limits. Kelle starts dictating the definition but she interrupts him: she says 'she'd rather use the theorem on the product of limits and prove discontinuity by contradiction'. *Kelle responds* 'If we assume it is continuous then...So then xintx when *x* is *n* is continuous' *and she stresses that they have to find a contradiction in assuming that xintx is continuous. He suggests 'calling the limit of xintx l'. She interrupts him: she says she prefers doing it* 'by reasoning, rather than epsilons'. *She then proves CD5.1ii and iv (fig.2c) and asks him to think about CD5.1iii. After a few silent seconds Kelle responds as follows:*

K3: Sin*x* is strictly less than 1 for all *x*...so since...therefore intsin*x* is continuous for all *x*...

T: Well not for all *x* because int*x* is not continuous for all *x*....

K4: Except for *x*=*n*.

T: You mean *x*=*n*?

K5: Well, when sin*x* is *n*. Sorry.

The tutor completes the proof (fig.2c) and again recommends 'drawing pictures' *in order to find out about limits.*

Extract 7.3

Context: *In the beginning of the tutorial the tutor and students Jack and Andrew discussed Linear Algebra. Now they turn to CD7.*

The Episode:

In CD7.1 the tutor notes that they 'both produced imperfections which were perhaps less matters of principle than of carelessness' *and that the question requires proofs based on first principles such as the definition of derivative. Andrew then says that* 'we can't assume any of the foundations...we are going sort of backwards'. *For finding the derivative of x^n, for $n>0$, both students have used induction but the tutor suggests the use of binomial theorem because it seems much more* 'natural' *to him. Jack protests: he* 'was not happy with the idea of assuming limits' *in his inductive proof. The tutor stresses that this discussion of principles is also relevant to CD7.2 which will be discussed later. He then finds the derivative of x^n, $n>0$, via the binomial theorem and the algebra of limits. He recalls from the history of mathematics that the binomial theorem was proved to serve the purposes of this proof and outlines a use of the product rule for finding the derivative of x^n, $n<0$.*

He then turns to CD7.2 and asks the students how they both knew that the $lim_{x\to 0}(e^x-1)$, is 1. Jack replies that he used the fact given in CD7.2 ($e^x=\sum x^n/n!$, $n=0,...$) and took limits 'on both sides'. *Andrew says he did the same and also justifies inserting the limit within Σ by quoting $lim(a+b+c)=lima+limb+limc$. The tutor interrupts him and firmly reminds them that the Σ above is not a sum but a limit. Therefore $lim\Sigma$ is actually a double limit. Andrew asks whether they* 'can exchange limits'. *The tutor writes down the $lim\Sigma$ and the Σlim and asks the students whether these are the same. Andrew replies* 'not necessarily'.

The tutor asks for a counterexample and Jack attempts constructing a function ($f(x)=y/x$: he starts but stops). The tutor reminds them that limits are operations like linear transformations and they do not commute. He then tries to construct a counterexample. Andrew sounds confused: 'So are we looking for an example where this is the case?'. *The tutor stresses that he is going to show* 'an example where it is not'. *The tutor completes the presentation of the counterexample (fig.3) but the students are sceptical: Andrew asks about how* 'useful' *the tutor's argument is: it ruins their proof but in this case it might be true that $lim\Sigma=\Sigma lim$. Jack adds he thinks it is and that* 'you can not-justify yourself on paper as long as you can justify yourself in a tutorial'. *The tutor comments upon Jack's words:* 'Well, this undermines the whole...' *and presents a proof for CD7.2 in which they do not need to consider the theorem on the exchange of limits: the theorem they've been* 'using instinctively here, without justifying it' *will be mentioned and proved next term even though they have been using it as a fact in other courses (Analytical and Numerical Methods and Differential Equations).* 'It's the job of this course to question it', *concludes the tutor.* 'Next term it will be reinstated with all the full rigour that mathematics is capable of'.

Extract 7.4

Context: *In the beginning of the tutorial the tutor and students Camille and Frances discuss Camille's queries from Series and Sequences, Vector Analysis and Topology. Finally Camille has another query from the introduction in the lectures to the notion of the Fourier Series of a function.*

The Episode:

C1: We've been doing Fourier Series and I don't really understand. I find difficult to understand this lecturer. She never said what a Fourier series is meant to be, she just started 'equation...'

The tutor says that 'the idea is that you are going to write your function in terms of a series involving trigonometric functions, sines and cosines'. *She further explains that the function needs to be periodic and despite the use of trigonometric functions intervals such as* $[-\pi,\pi]$ *or* $[0, 2\pi]$, *with an appropriate change of variable, any interval* $[a,b]$ *of periodicity is acceptable. Camille then asks if the lecturer's definition of periodicity* ($f(x)=f(x+a)$) *and her own* ($f(x)=f(x+ka)$) *are the same. They are, the tutor replies, as long as the k in Camille's definition is the smallest possible. The tutor defines the Fourier series of a continuous function (fig.4) and Camille interrupts in order to ask whether continuous means pointwise continuous. The tutor replies it does and continues:*

T1:...so you've got a continuous f on $[0,2\pi]$ and you can extend it periodically if you need to...so you've got $f(x)$ 'twiddles' [~], you mustn't put =, there are theorems that state when it's =...em...

C2: That's just means very close...

T2: No, this is just saying 'this is the Fourier series of this function'...

C3: Oh, so this is equivalent.

T3: No, 'this is the Fourier series of this function'...and then we define it. So we have half of a_0 plus the sum from 1 to ∞ $a_k\cos kx$ plus$b_k\sin kx$)...

C4: I didn't mean similar...I meant equals...

T4: No, it's not equal. Because sometimes you see you want to write down...it's not always the case that the Fourier series converges to the function, it's not always true that the Fourier series converges. But you want to be able to write it down. So you just put a ~ so that you can work it out and not worry whether it converges to it or not.

C5: *[turning to Frances who has been silent all this time]* Do you see this?

Frances nods and the tutor laughs. With Camille dictating she writes down a_k and b_k (fig.4). She then notes that this sum converges to $1/2[f(x+0)+f(x-0)]$. Camille then says 'But it doesn't really converge' *but the tutor reassures her that* 'no, it converges properly' *and moreover that if f is continuous then the Fourier Series converges to f because then $f(x+0)=f(x-0)=f(x)$. Camille is sceptical and then she asks whether $f(x+0)$ is what she knows as $f(x_+)$, that is the limit of f from the right. The tutor says that these are equivalent notations.*

Camille then turns to an exercise they were presented with in the lectures: Find the Fourier Series of f where

$$f(x)= \begin{cases} 1, & 0<x\leq\pi \\ -1, & -\pi\leq x\leq 0 \end{cases}$$

She is asking why the lecturer omitted calculating $1/2\ a_0$ and a_k. The tutor explains that the lecturer missed these out because they are equal to zero ('if you integrate that over $(-\pi, \pi)$ you are going to get the answer zero because the area there is minus the area there taking account of signs, I mean of s-i-g-n-s'). *Camille sounds fed up with the lecturer and frustrated.* 'Is she a friend of yours?' *she asks the tutor. Slightly embarassed the tutor replies* 'we are colleagues I think is what one says!'. *Silence follows and then the tutor explains that, in case the function is odd, that is $f(x)=-f(-x)$, $a_k=0$. Camille then says* 'So if you have a function from zero to π, you can extend it in two ways' *and the tutor agrees that* 'you can extend it either way, either with an even extension or with an odd extension'. *Camille then recalls that the lecturer mentioned* that 'you can extend it in three ways'. *She shows her notes to the tutor and asks for the third way. The tutor looks at them, frowns and ponders for a few seconds. She then exclaims:* 'Oh, plus, the ordinary, the Fourier series on $[0, a]$ without doing an extension'. *Then Camille asks:*

C6: So is the Fourier series an approximation of f?

T5: I wouldn't like to say that. I mean it was...*[pause]*...mmm...in most cases you will be looking at it's not an approximation, you'll have an equality there. And also in what sense would you say it was an approximation? *[Camille hums and haws]* I mean if you're actually trying to find nice and easy functions which approximate a given function you wouldn't do that...em,...you go for polynomials of Chebyshev polynomials or something like that which if you are interested in calculating integrals they are giving you much better approximation because the errors are so...it's...don't think about this as being an approximation. Em,...I mean what you are interested in are those functions...I mean you are interested in the Fourier series for a particular class of functions...I mean there are functions for which the Fourier Series actually converges to the function. Em,..now...there are various ways of approaching that problem. One of the easiest is to say whether you've got a function that is piecewise continuous and differentiable and that's not...I would need to think about it for a moment. I mean in other words we are sort of skipping things here a bit because you are doing this clearly from the point of view of using it, so you really need to know the minimum to actually do things. Because the actual theory of it gets quite complicated. And we do a bit...no you don't do this...maybe next year...

A short discussion follows on the course structure for Years 2 and 3 and the tutorial closes.

Extract 7.5

Context: *This is the beginning of the tutorial for students Camille and Eleanor. They are discussing B6.*

The Episode:

The tutor reads B6 (fig.5 illustrates the solution used in this tutorial) and asks Eleanor for her response in part i. Eleanor found that n must be ≥2. The tutor asks her why but Eleanor cannot reply. Then Camille claims that n is strictly >2 but then she explains that

$$\forall n \in N\text{-}\{0\}\ \lim f_n(x)=0 \text{ as } x \to 0$$
$$\text{and because } f_n(0)=0$$
$$f_n \text{ is continuous } \forall n>0.$$

The limit is zero, concludes Camille, because cosine is bounded by -1 and 1 and $x^n \to 0$, *as* $x \to 0$ *for* $\forall n \geq 1$. *For part ii Eleanor found that n must be ≥2 and the tutor agrees and asks Eleanor to justify. Eleanor dictates*

$$\text{the limit of } (f_n(x)-f_n(0))/x \text{ as } x \to 0.$$

Then Camille asks how the tutor decided that 'it is n≥2 straightaway' *and the tutor replies* 'because Eleanor knows the answer and I know the answer!'. *She then asks Camille what is the reason for looking at this limit. Camille replies that* 'the limit of that is the derivative at zero'. *She then dictates the derivative and notes that this limit is zero* $\forall n \geq 2$. *Camille then asks:*

C1: You just want it to be differentiable at zero, you don't want it to be continuous...? So it doesn't matter if...it doesn't need to tend to zero...

T1: Well, it will do. I mean it just will do because we know it does. I mean we don't need it. But that's what happens.

C2: Always?

T2: Yes, because we've just proved it. I mean we are using the fact that we know the limit of this from there. As long as we make sure that this exponent is at least one.

C3: And the derivative at zero is it always continuous at zero?

T3: No. That's another question that we've got to investigate for part iii. Er,...so certainly f_n at zero will exist if $n \geq 2$ and its limit at zero will be equal to...?

C4: Zero.

T4: Yes, and we've also got to check...to show that if $n=1$ it doesn't exist. You've got to cover all the cases. For $n=1$ what will it be equal to?

C5 and E1: $\cos 1/x^2$.

T5: And we know what's gonna happen to that?

E2: Keeps going to infinity.

T6: Does not tend to a limit.

For part iii the tutor notes that they know that n must be at least 2 from part ii. So, she says, they are looking for what happens when $n \geq 2$. Silence. Eleanor suggests looking at the derivative and then 'doing a limit like what we did before'. *Camille dictates*

$$f'_n(x) = nx^{n-1}\cos 1/x^2 + 2x^{n-3}\sin 1/x^2, \ \forall n \geq 2$$

and the tutor observes that 'in fact this is true for $\forall n \in N$'. *The tutor asks about x and Camille says it has to be >0. The tutor asks them to* 'think more generally' *and Eleanor says ≠0. They also know that $f'_n(0)=0$ $\forall n \geq 2$, the tutor adds. So the question is...?she asks. Camille finishes the tutor's sentence:* '...for which n to have the limit of this derivative to be zero'. *The tutor accepts that and then Camille says that it is so $\forall n \geq 4$.*

C6: Because this term *[points at the first term of the derivative]* will be at least x or x^2,...it will be a power of x and therefore the limit of that is zero because the limit of the power is zero.

Camille adds that the second term of the derivative tends to zero for $n \geq 4$. She then asks what would happen if $x>1$. The tutor replies that $x \to 0$ so this is not an issue here. So they know that the derivative is continuous at zero for $n \geq 4$. They still have to exclude the cases for $n=2$ and $n=3$. The tutor writes them down and asks.

T7: So what can we say about these?

C7: One tends to...this one is OK...because we just don't know where $1/x$ is going...the second is more...

T8: Is it? *[turning to Eleanor who is mumbling something]* What did you say, Eleanor, it's going to what?

C8: *[after a pause]* But it's not quite vast...

T9: I didn't quite hear what Eleanor said!

C9: Oh, yeah...Going to infinity? *[Eleanor nods]*

T10: But it's not quite...it doesn't tend to...I know what you mean but...the absolute value tends to infinity.

The tutor explains how the function oscillates and notes that the first part is 'well-behaved and you cannot quite cancel the bad behaviour of the second part of $f_2(x)$. If they both behave badly they might somehow cancel out the bad behaviour...but that's not gonna happen here'. *So here for $f'_2(x)$ the first part tends to zero, as x tends to zero, and there is no limit for the second. And the other one, even though it exists, it's not continuous at zero, for n equal to 3. She closes by writing the two limits down.*

Extract 7.6

Context: *This is the beginning of the tutorial for students Abidul and Frances. They are discussing B7.*

The Episode:

The tutor asks the students how they 'decided there was just one root' *in part a (fig.6).*

A1: I found the stationary points and there were two, so I said after the minimum the graph has to go up and after the maximum it has to go down and so it has to cut the graph somewhere because it tends to infinity...

The tutor nods in approval and asks Abidul what theorem she used in order to determine that because f goes to infinity as x tends to -∞, f must have a zero. The student is hesitant and whispers 'the Rolle's'. *The tutor waits for a few seconds and then repeats the question to Frances. Frances replies it was the IVT and the tutor asks what does the Intermediate Value Theorem say.*

F1: *[after a pause]* If *f* is continuous...then. And if *f* of...if *b* is greater than *a*, then *f*(*b*) is going to be...

A2: *F*(*a*) is less than 1 over...Then *c*...

T1: *[laughing]* I think between you, you've got a correct version. So we've got a *c* somewhere...where?

A3: In *a*,*b*. And it is *f*(*c*)= lamda...

The tutor completes the outline of the proof by pointing out that the IVT guarantees that f takes all the values between its maximum and -∞ and therefore it has to cut the x axis at one point.

In part b (fig.6) both students found that f has a min at 1, to the left of which it is strictly decreasing and to the right is strictly increasing. Therefore Imf=f((0,∞))=[e, +∞). They both have drawn the graph but cannot answer why Imf=[e, +∞). Abidul mumbles that she used the second derivative and continuity and that the function is monotone increasing. The tutor prompts them by saying that, from IVT they know that Imf must be an interval because f is continuous. The other thing they need to know, she continues, is that f is unbounded above. The tutor points at the graphs they drew and notices that they have drawn f as an unbounded function. She asks them why they 'sent off' *the graph like this. Frances whispers that she drew f like that* 'because as x gets very small...$1/x$...'. *The tutor nods in agreement then stops Frances and completes the argument:* $1/x \to \infty$ *and* e^x *multiplies it with something big.*

About the second part of b, Abidul says she didn't 'really know what to do' *and Frances suggests* $I=(0,1)$. *The tutor nods in approval. After a pause she responds to the tutor's request for an explanation with:*

F2: Because you can tell from the graph! *[the tutor asks why again and a pause follows]* If...you reflect it on y...

The tutor asks her what theorem she used in order to guarantee the existence of the inverse function. The students are silent and the tutor states the Inverse Function Theorem and outlines its use for the particular function in question.

Extract 7.7

Context: *The tutorial began with Topology. Now the tutor and students Cary and Beth turn to B10.*

The Episode:

Beth dictates the Taylor expansion (fig.7). For part a of the question the tutor suggests they substitute x^3 for f(x) and 2 for n. Both students calculate and find $\theta=1/3$.

In part b the tutor stresses that what the two inequalities say is that g cannot increase too fast. She is 'surprised' *the students have been unable to complete the question because* 'there is nothing much you can do here'. *She suggests fixing an x and taking the Taylor expansion around x. The students are silent and she asks:* 'Now $g(x)$ is going to be equal to...?' *Beth starts dictating* '$g(x)$=...' *but the tutor interrupts to remind her that* 'we've got to decide what we are going to do with a and h'. '$g(x)$ is in [0,2]?' *asks Beth and the tutor then points out that, in order to obtain an expression for g'(x), they have to take the expansion around x+h where x is in [0,2]. Beth dictates. She confuses f and g. The tutor corrects her and she completes correctly. The tutor points at the term they are interested in, g'(x), and she brings this term on the left hand side of the equality. With Beth's observation that the g and g'' terms in || are <1 they get to*

$$|g'(x)| \leq 2/h + h/2$$

where h is such that $x+h \in [0,2]$

given that $x \in [0,2]$.

This, says Beth, is less than 2 1/2 and this is attained for $|h|=1$. The tutor agrees and asks for a proof. Beth remains silent. Cary suggests differentiating $2/h+h/2$ in order to prove that its maximum is 2 1/2. The tutor agrees but she also wants them to find out about an algebraic way of doing this. She tries using the fact that $a^2+b^2 \geq 2ab$ but fails. She then accepts Cary's way and returns to their original question: via differentiation they prove that the maximum value of $2/h + h/2$ is 2 1/2 and attained for $|h|=1$. Now the tutor wants them to find out if it is always possible, that is for all x in [0,2] to choose h with $|h|=1$ so that x+h is in [0,2]. For x=1 it is OK, observes Beth. What about the other x in [0,2]? asks the tutor. For x<1? she asks. Beth suggests taking h=1 and then she suggests that for x>1 we can take h=-1. The tutor agrees and concludes that

$$\forall x \in [0,2], \text{ for either } h=1 \text{ or } h=-1,$$
$$|g'(x)| \leq 2\ 1/2.$$

Extract 7.8

Context: *This is the beginning of the tutorial for students Cathy and Cliff. They are discussing SS7.*

The Episode:

The session begins with SS7. Cliff had problems with SS7.1iv and the tutor promises to come back to it once they have worked on i and ii. So he invites Cliff to present SS7.1i. Cliff 'splits up' *$\Sigma 1/r(r+k)$ as $\Sigma(1/kr - 1/k(r+k))$ and subsequently calculates the infinite sum. The tutor agrees but suggests the more* 'formally acceptable' *way of doing the same not on the infinite sum, but the finite sums and then taking the limit. SS7.1ii was similar.*

The tutor then asks Cathy to outline what she did in SS7.1iii: she broke the $(-\infty)$-$(+\infty)$ sum in two: $(-\infty)$-0 and 0-$(+\infty)$. Then she removed || and calculated the two infinite sums. The tutor agrees and asks Cathy to present SS7.1iv (left column of the following table). The tutor agrees and illustrates an alternative way (right column of the following table).

Cathy's Way	*The Tutor's Way*
${}_1\Sigma^\infty r^2/3^r = {}_1\Sigma^\infty r^2-1+1/3^r$ $= {}_1\Sigma^\infty (r-1)(r+1)/3^r \quad S_\infty=1/2$ $= {}_1\Sigma^\infty (r+1)(r-1)/3^r+1/2$ $= {}_0\Sigma^\infty (r+2)r/3^{r+1}+1/2$ $= 1/3\ {}_1\Sigma^\infty r(r+2)/3^r + 1/2$ $= 1/3\ {}_1\Sigma^\infty r^2/3^r + 1/3\ {}_1\Sigma^\infty 2r/3^r+ 1/2$ $2/3\ {}_1\Sigma^\infty r^2/3^r= {}_1\Sigma^\infty 2r/3^{r+1} + 1/2$ $= {}_1\Sigma^\infty 3r/3^{r+1} - r/3^{r+1} +1/2$ $= {}_1\Sigma^\infty r/3^r - r/3^{r+1} + 1/2$ $= (1/3-1/9+2/9-2/27+3/27...) +1/2$ $= (1/3+1/9+1/27+1/81...) = 1/2$ $+1/2$ so $\Sigma r^2/3^r = 3/2$	*Note that if* $f(x) = {}_1\Sigma^\infty x^r = 1/1-x$, then $f'(x) = {}_1\Sigma^\infty rx^{r-1} = 1/(1-x)^2$ and $f''(x) = {}_1\Sigma^\infty r(r-1)x^{r-2}$ Then by writing f'' in terms of f' and f, and for x=1/3, it turns out that $\Sigma r^2/3^r = 3/2$.

The same technique, continues the tutor, which allows us to differentiate an infinite sum term by term applies to another question from the same sheet, on power series.

Appendices for Chapter 8

Appendix 8A

LA5.23

Let $S \subseteq V$, $T \subseteq V$. Show that

(i) $\langle S\rangle = \langle S \cup \{0\}\rangle$

(ii) if $T \subseteq S$, then $\langle T\rangle \leq \langle S\rangle$

(iii) $\langle\langle S\rangle\rangle = \langle S\rangle$

(iv) if $T \subseteq S$ then $\langle T\rangle \leq \langle S\rangle$

(v) $\langle S\rangle + \langle T\rangle = \langle S \cup T\rangle$

Show also that $\langle S \cap T\rangle = \langle S\rangle \cap \langle T\rangle$ is not true in general.

LA5.24

Show that rowspace $(PA) \leq$ rowspace (A) for P an $r \times m$ matrix and A an $m \times n$ matrix.
Deduce that if B can be obtained from A by row operations then rowspace (B) = rowspace (A).
Verify that the non-zero rows of an echelon form matrix are linearly independent.

LA6.26

Compute the dimensions of the following subspaces of $M_n(\Re)$:

(i) the set of diagonal matrices
(ii) the set of symmetric matrices
(iii) the set of antisymmetric matrices
(iv) the set of matrices with trace zero.
[the trace of (a_{ij}) is defined as $\sum_{i=1}^{n} a_{ii}$, i.e. the sum of the diagonal elements.

LA6.29

(a) Let $X = \{(\alpha, \beta, \gamma, \delta)$: $\alpha+\beta+\gamma=\alpha-\delta=0\}$, $Y = \{(\alpha, \beta, \gamma, \delta)$: $\beta+\gamma=\delta=0\}$.
Find a basis for $X+Y$ which contains a basis for X and a basis for Y.
(b) Let X and Y be subspaces of dimension n-1 of a vector space of dimension $n \geq 2$.
Prove that is $X \neq Y$ then $X \cap Y$ has dimension n-2.
What is the geometrical significance of this in $\Re^3$?

LA7.35

Let T: $V \rightarrow V$ be a linear transformation of a finite dimensional vector space V into itself.
Show that $\mathrm{Im}T^2 \leq \mathrm{Im}T$ and that $\ker T \leq \ker T^2$. Prove the equivalence of the following:

(a) $V = \ker T \oplus \mathrm{Im}T$

(b) $\ker T = \ker T^2$

(c) $\mathrm{Im}T = \mathrm{Im}T^2$

[Recall that $V = \ker T \oplus \mathrm{Im}T$ if $V = \ker T + \mathrm{Im}T$ and $\mathrm{Im}T \cap \ker T = \{0\}$]
[You may find it helpful to derive the equation

$$\dim \mathrm{Im}T = \dim \mathrm{Im}T^2 + \dim(\mathrm{Im}T \cap \ker T)$$

by applying the Rank and Nullity Theorem in the case where T acts on $\mathrm{Im}T$].

B3

The real vector spaces V and W have bases $\mathcal{E} = \{e_1, e_2, e_3\}$ and $\mathcal{F} = \{f_1, f_2, f_3\}$ respectively. The linear mapping $T : V \to W$ has matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 2 \end{pmatrix}$$

with respect to $\mathcal{E}$ and $\mathcal{F}$. Let $\mathcal{E}' = \{e_1', e_2', e_3'\}$, where $e_1' = e_1 - e_3$, $e_2' = e_2 + e_3$ and $e_3' = 2e_1 - 2e_2 - e_3$. Prove that $\mathcal{E}'$ is a basis of V, and find the matrix of T with respect to $\mathcal{E}'$ and $\mathcal{F}$. Find a basis $\mathcal{F}'$ of W such that T has matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with respect to $\mathcal{E}'$ and $\mathcal{F}'$.

Appendix 8B

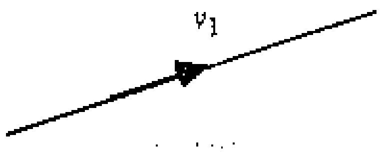


fig.1a

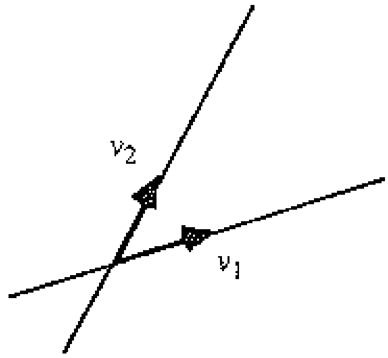


fig.1b

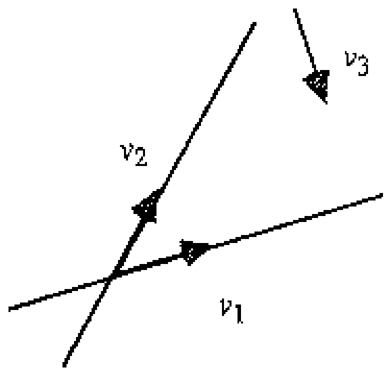


fig.1c

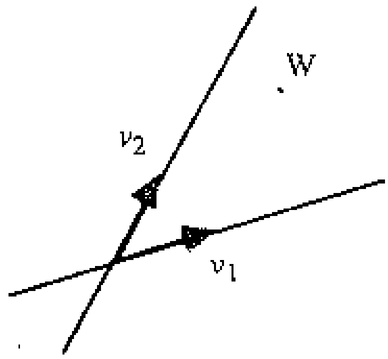


fig.1d

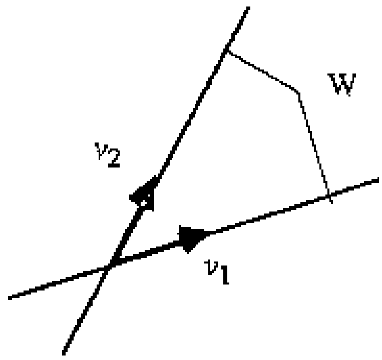


fig.1e

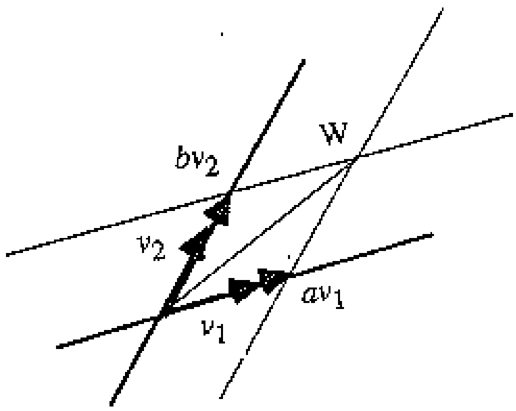


fig.1f

Let $X = \{x_{ij}\}$, where $i,j = 1, ..., n$, be a matrix of Trace zero: $Tr(X) = 0$. Then

$$X = \sum_{i\neq j} x_{ij}E_{ij} + \sum_{i=1}^{n-1} x_{ii}(E_{ii} - E_{nn}).$$

Matrices E_{ij} and $E_{ii}-E_{nn}$ in the above are all of Trace zero, they span the set of X and their number is

$$(n^2-n)+(n-1) = n^2-1.$$

Therefore the dimension of the subspace of matrices of Trace zero is n^2-1.

fig.4 LA6.26i

Let $u_1, ..., u_p$ be a basis for $X\cap Y$. This can be extended to

a basis for X: $u_1, ..., u_p, x_1, ..., x_q$

and to

a basis for Y: $u_1, ..., u_p, y_1, ..., y_r$.

The set $\{u_1, ..., u_p, x_1, ..., x_q, y_1, ..., y_r\}$ spans $X+Y$. It is also linearly independent because:

Suppose that

$$a_1u_1+ ...+ a_pu_p+ b_1x_1+...+ b_qx_q+ c_1y_1+...+ c_ry_r = 0.$$

Then:

$$a_1u_1+ ...+ a_pu_p+ b_1x_1+... + b_qx_q = - c_1y_1-...- c_ry_r$$

Left Hand Side Vector belongs to X and Right Hand Side Vector belongs to Y. Therefore the vector above belongs to $X\cap Y$ and there exist $d_1, ..., d_p$ such that

$$d_1u_1+ ...+ d_pu_p = a_1u_1+ ...+ a_pu_p+ b_1x_1+ ... + b_qx_q = - c_1y_1-...- c_ry_r$$

The above relationships imply that all *a*s, *b*s, *c*s and *d*s are equal to zero.

Therefore $\{u_1, ..., u_p, x_1, ..., x_q, y_1, ..., y_r\}$ is a basis for $X+Y$.

fig.5 LA6.29i (partly)

$\text{Im}T$, $\text{Im}T^2$, $\ker T$ and $\ker T^2$ are subspaces of V. Therefore all I need to show is:

$$\text{Im}T^2 \subseteq \text{Im}T,$$
$$\ker T \subseteq \ker T^2$$

Let $x \in \ker T$. Then $Tx = 0$ and because $T0 = 0$, $T(Tx) = T^2x = 0$, hence $\ker T \subseteq \ker T^2$.

Let $x \in \text{Im}T^2$. Then $\exists u \in V$: $T^2u = v$, that is $T(Tu) = v$, that is I found an element in V, Tu, which is mapped on v through T, therefore $v \in \text{Im}T$ and so $\text{Im}T^2 \subseteq \text{Im}T$.

To prove that a $\Leftrightarrow$ b $\Leftrightarrow$ c, I prove b $\Leftrightarrow$ c and a $\Leftrightarrow$ b.

<u>b $\Rightarrow$ c</u> From the Rank and Nullity Theorem I know that

$$\dim V = \dim\text{Im}T + \dim\ker T$$
$$\dim V = \dim\text{Im}T^2 + \dim\ker T^2$$

Then, from $\ker T \subseteq \ker T^2$, $\dim\text{Im}T = \dim\text{Im}T^2$ and from $\text{Im}T^2 \subseteq \text{Im}T$, I deduce $\text{Im}T^2 = \text{Im}T$. So b $\Rightarrow$ c. Symmetrically I can prove c $\Rightarrow$ b.

<u>b $\Rightarrow$ a</u> To prove that $V = \text{Im}T \oplus \ker T$, I need to prove that

$$\text{Im}T \cap \ker T = \{0\} \text{ and } \text{Im}T + \ker T = V.$$

In fact all I need to prove is that $\text{Im}T \cap \ker T = \{0\}$ because:

if $\text{Im}T \cap \ker T = \{0\}$ then $\dim(\text{Im}T + \ker T) = \dim \text{Im}T + \dim\ker T = \dim V$

therefore, since $\text{Im}T + \ker T \subseteq V$, $\text{Im}T + \ker T = V$.

So let $v \in \text{Im}T \cap \ker T$, $v \neq 0$. Then $v \in \ker T$ therefore $Tv = 0$.
Also $v \in \text{Im}T$ therefore $\exists w \in V$: $Tw = v \neq 0$. Then $T^2w = Tv = 0$, therefore $w \in \ker T^2$.
But I know that $w \notin \ker T$ because $Tw \neq 0$ and this contradicts my assumption (b), $\ker T = \ker T^2$.
Hence $\text{Im}T \cap \ker T = \{0\}$.

<u>a $\Rightarrow$ b</u> I know that $\ker T \subseteq \ker T^2$. So all I need to prove is that $\ker T^2 \subseteq \ker T$.

Let $v \in \ker T^2$. Then $Tv = 0$ and, if I call $Tv = w$, $Tw = 0$. Then $w \in \ker T$ but from (a) this implies that $w = 0$. Then $w = Tv = 0$, therefore $v \in \ker T$.

fig.7 LA7.35

To prove that $\mathcal{E}'$ is a basis of V I need to prove that $\mathcal{E}'$ is linearly independent. Indeed

$$\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0,$$

when e'_i are substituted as combinations of e_i and given the linear independence of e_i, implies

$$\lambda_1 = \lambda_2 = \lambda_3 = 0.$$

To find the matrix $A = \{a_{ij}\}$, where i, j = 1, 2, 3, of T with respect to $\mathcal{E}'$ and F, I construct the three equations

$$Te'_i = \sum_{j=1}^{3} a_{ij} f_j \ , \ i = 1, 2, 3.$$

I substitute Te'_i as combinations of Te_i and then substitute Te_i as combinations of f_j. As a result I get

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & 2 \\ -1 & 3 & 2 \end{pmatrix}$$

Finally to find a basis $F = \{f_j\}$, j=1, 2, 3, of W such that T has matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

I calculate Te'_i as combinations of Te_i and substitute Te_i as combinations of f_j. Then I call

$$Te'_1 = f_1 \text{ which gives } f_1 = f_2 + f_3$$
$$Te'_2 = f_2 \text{ which gives } f_2 = f_1 + 2f_2 + 3f_3$$

Because $Te'_3 = 0$ and the zero vector cannot be an element of a basis, I can choose as f_3 any vector W so that $\{f_1, f_2, f_3\}$ is linearly independent, for instance f_1.

fig.8 B3

Appendix 8C

Extract 8.1

Context: *Four pairs of female students are taught in four consecutive tutorials about the span of a set:*

Tutorial   *1: Abidul and Eleanor*
2: Patricia and Frances
3: Camille and Cleo
4: Beth and Cary

The mathematical structure in the four sessions is as follows:

definition of the span of a set
constructing the span of
$\{v_1\}$,
$\{v_1, v_2\}$,
$\{v_1, v_2, v_3\}$,
where v_i *are vectors in a vector space V.*

In the following I cite extracts from Tutorials 1, 2 and 4, which were almost identical in structure as well as a summary of Tutorial 3 which was structurally different because of one of the students' interventions.

Discusssion of the concept of Span starts with the students' questions about its meaning in all four sessions. In Tutorial 3 the tutor, who in the meantime was surprised by the unanimous unease with the concept, finds out that the definition of Span of a set given in the lectures was

> *the span of* $S=\{v_1, v_2, \ldots, v_k\} \subseteq V$ *is the smallest subspace of V which contains S.*

The definition used in these tutorials is

> *the span of* $S=\{v_1, v_2, \ldots, v_k\} \subseteq V$ *is the subspace of V that contains the linear combinations of* $v_1, v_2, \ldots, v_k$

and I emphasise that none of the eight students had heard of it before. The tutor explains that in order to construct the span of $S=\{v_1, \ldots, v_k\}$ *we take all the multiple scalars of* v_i *and the vectors coming out of adding those. She also warns the students that there is a 'change of gear' in the course which now becomes more abstract.*

The Episode:

Tutorials 1, 2 and 4

Constructing the span of $\{v_1\}$. *The tutor draws fig.1a and asks the students to find the span of* $S=\{v_1\}$. *She distinguishes between the cases where* $v_1=0$ *and* $v_1\neq 0$. *The students all point out that if* $v_1=0$ *then* $<S>=\{0\}$. *In Tutorial 4 however Beth claims that* $<S>$ *is...*

B1: ...the empty set.

T: What do you mean by 'empty set'?

B2: Zero.

The tutor repeats clearly to Beth that if $v_1=0$ *then* $<S>=\{0\}$. *Then she explains that if* $v_1\neq 0$ *then* $<S>=<v_1>=\{av_1: a\in\Re\}$ *and asks the students to demonstrate on paper (fig1a) what is the role of a: students Abidul, Frances and Beth respond as follows:*

A1: You multiply by the multitude of v_1 and *[in case* $a<0$*]* you're going that way *[opposite direction to* v_1*]*.	F1: It will be the distance a times v_1 for av_1 and *[in case* $a<0$*]* you have to turn round beyond zero.	B3: *[shows with her fingers the positive and negative directions of* av_1*] and says that all* av_1 *will be* 'on a line'.

Constructing the span of $\{v_1, v_2\}$. *For* $<v_1, v_2>$ *the tutor adds a vector* v_2 *to fig.1a (fig. 1b) and distinguishes between* $v_2\notin <v_1>$ *and* $v_2\in <v_1>$. *In the case where* $v_2\notin <v_1>$ *discussion in the three tutorials is as follows:*

A2: So it's going to be the plane defined by v_1 and v_2. B1: See how far it goes to this direction and how far to this...*[shows with her fingers on the paper something like a translation in order to show* 'how far' *she wants to go]* *The tutor draws the parallels but the girls do not know how to continue. Eleanor whispers something that sounds like* 'finding $b_1$and b_2', *the coefficients of v_1 and v_2. Abidul adds:* A3: It's b_1 times the multitude of v_1... and b_2 times the multitude of v_2.	P1: The sums of the form $a_1v_1+a_2v_2$ will belong to $<v_1, v_2>$. Tutor: Pictorially? P2: You complete the parallelogram... *The tutor explains that $<v_1, v_2>$ will be the whole plane and asks* 'Can I get to the whole plane this way? By using v_1 and v_2?'. *And* 'if I had a vector w here how would you find what scalar multiples of v_1 and v_2 I should use to write it in terms of v_1 and v_2?'. *Silence.* P3: That would be a minus av_2... *[asked to show on paper Frances and Patricia bend towards the drawing and point with their fingers].*	B4. *Beth, while saying that the span of v_1 is one line and the span of v_2 is* 'another line', *is very surprised to hear that $<v_1, v_2>$ is the whole plane:* 'Is it the whole thing?' *she asks.* *The tutor draws a vector w and asks for the coefficients b_1 and b_2 in the linear combination $w=b_1v_1+b_2v_2$.* C1: We just have to take components. Tutor: *[makes a wry face at hearing the word 'components']* Although I've drawn this on the plane, I might have matrices or whatever else. *With the tutor's closed questioning Cary outlines the Parallelogram Rule.*

In Tutorials 1 and 2 the tutor introduces the Parallelogram Rule.

Asked about the case $v_2 \in <v_1>$ students Abidul, Patricia and Beth respond as follows:

A4: You're just gonna get the same line again. It's... the sum will be a times v_1 plus again... I mean it's gonna be the same line.	P4: v_2 is the spanning of v_1... *The tutor corrects: the line on which v_1 lies is the span of v_1. Then Patricia asks:* P5: Would it just be around v_1? You wouldn't have the plane...	B4: It would be along the line.

Since the students are hesitant about how to prove their claims the tutor then, using the linear dependence of v_1 and v_2, proves that $\langle v_1, v_2\rangle = \langle v_1\rangle$.

Constructing the span of $\{v_1, v_2, v_3\}$. *Asked about $\langle v_1, v_2, v_3\rangle$ (fig.1c) in Tutorials 1 and 2 the students distinguish between the cases where $v_3 \in \langle v_1, v_2\rangle$ and $v_3 \notin \langle v_1, v_2\rangle$. In the latter case* 'It's going to give us the three dimensional space' *and in the former* 'v_3 wouldn't matter' *point out students Abidul and Frances in tutorials 1 and 2 respectively. In tutorial 4 Beth refers to the former exclusively and only when the tutor draws her attention to the latter she notices that then there is another dimension coming in.*

Tutorial 3

Student Camille starts the conversation on Span by confusing the term Span with the word 'scope'. *Given the* 'more practical definition' *of Span of a set used in these tutorials, Camille notes that $0 \in \langle S\rangle$ because $\langle S\rangle$ is* 'the sum of these vectors, with any coefficients and the coefficients can be zero'. *Also that* '$\langle S\rangle$ is bigger than S'. *So, she concludes,* $\langle S\rangle$ 'is all the combinations of directions you can get from those directions'. *The tutor agrees that* 'this is a way of looking at it' *and notes that $\langle S\rangle$ is a subspace of V. Then Camille asks:*

C1: How do you prove that this is the smallest subspace... I understood how it contains S but how do you prove it's the smallest?

The tutor explains how $\langle S\rangle$ can be constructed by 'building it up' *from the elements of S, then taking all their multiples and adding these up. Cleo then asks:* 'then $\langle S\rangle$ is a subset of V?'. *Camille explains:*

C2: It's something more than that, it's a subspace. What's the diff... a subspace is something smaller that has the same properties like the properties on addition is that it... The difference between a subset and a subspace is that a subset can be for instance the empty set and a subspace has to have enough elements to keep all that...

In response to Camille's questions the tutor talks about $<v_1>$ *and draws it as a line (fig.1a) on paper. She then asks about* $<v_1, v_2>$. *Camille responds:* 'it's the whole... *plan [in French this means the 'plane']...* what-you-call-it provided that v_2 is not on that line'. *The tutor then points out that if we pick a point W on the plane we can write* $w=av_1+bv_2$ *(fig.1d). How can we then find a and b? she asks. The dialogue is as follows:*

C3: By taking the projection of the point to the axes...

T1: What do you mean by the projection?

C4:..the perpendicular, by dropping...

T2: Drawing perpendiculars, is that right?

C5: Mm.*[meaning yes]* Then measuring...

The tutor then turns to Cleo: does she agree? Cleo asks the tutor to repeat the question as well as what they are trying to do. The tutor repeats the question and concludes by reminding Cleo that Camille 'suggested dropping perpendiculars'. 'Don't you say "take projections" ?' *asks Camille. The tutor (due to a phonecall) leaves them for a few minutes to discuss among themselves and when she returns Camille says that they want to take the projections and measure the distance between O and the projected points. Asked by the tutor whether she agrees Cleo nods and the tutor draws two vectors (fig.1b) and asks them how we add these up. They both know the Parallelogram Rule. The tutor asks them if now they* 'revise what they suggested'. *They are silent and the tutor decides to apply the suggestion and* 'drop perpendiculars' *(fig. 1e). At the sight of fig.1e Camille changes her mind:*

C6: I made a mistake. I thought w was a point in the space. I didn't think...

Everybody is silent for a few seconds. Then the tutor asks how we add v_1 *and* v_2. *Silence continues.* 'Draw the parallel', *replies Camille. She then suggests they remove the parallels to the two lines until they meet W and when one parallel meets the first line it is* av_1 *and the other it's* bv_2 *(fig.1f). To the tutor's request Camille points at where* av_1 *and* bv_2 *are on the drawing. For the latter she says* 'minus bv_2' *which she explains with* 'because it's on the left'. *When the tutor asks her to show* 'plus bv_2' *Camille changes her mind and dismisses* 'minus'.

Subsequently the case for $v_2 \in \langle v_1 \rangle$ *is discussed and then the tutor asks about* $\langle v_1, v_2, v_3 \rangle$. *Camille notes:*

C7: So when you take *k* then you have *k* dimensions.

T3 Yes, once you make sure that the one you add up is not in the span of the others.

C8: After 3 I can not imagine it. It's getting confusing.

This part of Tutorial 3 closes with the tutor explaining the convenience of working in n dimensions. In the end Camille repeats the definition of $\langle S \rangle$ *used in this tutorial (as shown in the Context) in order to* 'make sure she understood it correctly'.

Extract 8.2

Context: *Same college, tutor and students as Extract 8.1. The tutor explains that a subset S of a vector space V is a subspace of V (S<V) if S is a vector space and that a common way to prove that S<V is to use the Subspace Test (namely prove that 0∈S and that addition and scalar multiplication are closed in S). The students cannot recall the SubspaceTest at first and the tutor suggests two examples: prove that the n×n symmetric matrices and the matrices of Trace zero are subspaces of $M_n(\mathfrak{R})$. The third example is to*

prove that $U=\{f:\mathfrak{R}\to\mathfrak{R}, f(0)=f(1)\}$ is a subspace of $\mathfrak{R}^{\mathfrak{R}}$, the set of all real functions from $\mathfrak{R}$ to $\mathfrak{R}$.

The students look as if they are not familiar with $\mathfrak{R}^{\mathfrak{R}}$.

The Episode:

Tutorials 1, 2 and 4

Note: In the following I have summarised Tutorials 1, 2 and 4 and highlighted some instances in each one of them.

In Tutorial 2, even though it turns out that the students do not know what $\mathfrak{R}^{\mathfrak{R}}$ is and have not realised that $U\subseteq\mathfrak{R}^{\mathfrak{R}}$, Abidul mechanically suggests checking out the two conditions of the Subspace Test. The tutor realises their unease with $\mathfrak{R}^{\mathfrak{R}}$ and asks what is the zero element of $\mathfrak{R}^{\mathfrak{R}}$. She reminds them that the zero element is an element of U and she asks them what is the property it has to satisfy. Silence follows. Eleanor says about the zero element of $\mathfrak{R}^{\mathfrak{R}}$ that 'It will stay the same'. *The tutor disagrees and switches to talking about a general vector space. To their silence the tutor writes down $a+0=0+a=a$ and asks them to think in terms of the zero element z in U and $\mathfrak{R}^{\mathfrak{R}}$. Abidul says $z(x)=x$ and Eleanor says* 'nought'. *The tutor agrees with Eleanor that it is the function of value 0 everywhere and stresses that it is a function they have been dealing with* 'for ages' *in Analysis.*

In Tutorial 4, Beth remembers that the zero element of $\mathfrak{R}^{\mathfrak{R}}$ is the zero function when the tutor lists the axioms that a vector space has to satisfy (one of them is that there must be an element, called zero element, such that $a+0=0+a=a$, $\forall a\in\mathfrak{R}^{\mathfrak{R}}$).

In Tutorial 1 the students remain silent for longer and they cannot recall the property that the zero element of $\Re^{\Re}$ must satisfy. The tutor switches to asking the same question in a general vector space and Patricia responds that in a general vector space 'if you add zero to any vector you end up with the same vector'. *The tutor asks them to apply that in U but, since the students remain silent, she eventually defines the zero function (z: $\Re \rightarrow \Re$ z(x)=0) and stresses that* 'It is the same function from Analysis in a slightly more abstract context'.

In proving closure, the second condition of the Subspace Test, the students suggest evaluating af+bg, for f, g∈U and a, b∈$\Re$ at 0 and 1 and proving that the two values are equal. In some cases the tutor needs to prompt the students who seem to confuse f with f(x). The sessions close with the tutor repeating the conditions of the Subspace Test.

The Episode: Tutorial 3

Camille is asking about $\Re^{\Re}$:

C1: Is it the same as $\Re^2$?

The tutor defines $\Re^{\Re}$, $2^{\Re}$ and $\Re^2$ (starting from the definition of A^B as the set of functions from the set B to the set A). Camille asks whether these are transposes. These are the mappings from one set to the other, replies the tutor, and are vector spaces over addition of functions and over scalar multiplication, both pointwise. She then asks the students what is the zero in $\Re^{\Re}$ but Camille is still trying to understand what $\Re^{\Re}$ consists of.

C2: It's a mapping from $\Re$ to $\Re$... and each element of $\Re$ has a correspondent... with the mapping... the graph... it's a mapping from $\Re$ to $\Re$ so it's...

T1: No, no, no. Each mapping is a subset of $\Re$x$\Re$. It's not...

C3: And *U* is a subset...

T2: Yes. No. No, you are not looking at the individual *f*. Yes, it's true that *f* is a subset of $\Re^2$. That's true but I'm not looking at the individual *f*. I'm lookiing at the set of all *fs*.*[The tutor returns to the question about the zero element of $\Re^{\Re}$]* What's your mapping from $\Re$ to $\Re$ with that property? It's not in there. *U* is not a subset of $\Re^2$.

C4: Do you have an example of a function *f* that isn't in $\Re^2$?

T3: Yeah, I mean any of them... if you like you could have... cos$2\pi x$... this is a function that... I mean to say that something belongs to $\mathfrak{R}^2$... what you are saying is that f belongs to $\mathfrak{R}^2$. I mean that f belongs to $\mathfrak{R}^2$. But it doesn't because that... to say that f belongs to $\mathfrak{R}^2$ is to say that f is an ordered pair. It's a set of ordered pairs.

C5: $(x, f(x))$... but isn't f an ordered pair?

T4: No, f doesn't belong to $\mathfrak{R}^2$. f is a subset of $\mathfrak{R}^2$. And it belongs to the set $\mathfrak{R}^{\mathfrak{R}}$. It belongs to a set of functions.

C6: It's very hard to imagine that... a set is usually a set of elements or matrices...

T5: Ah,... yes, that makes it harder. But I mean you could do it with equations.

The tutor stresses that all the things they are talking about here are casual things in Analysis and that their trouble is with the vector space context.

C7: So $\mathfrak{R}$ is a subset of $\mathfrak{R}^{\mathfrak{R}}$...

T6: The elements are...

C8: How about $\mathfrak{R}^2 \rightarrow \mathfrak{R}$... is it a subset of this?

The tutor repeats the definition of $\mathfrak{R}^{\mathfrak{R}}$ and draws the paralled with A^B, the set of functions between sets A and B, where A has k elements and B has m. She stresses that $\mathfrak{R}^{\mathfrak{R}}$ contains the functions they 've been dealing with usually in Analysis and that it is the vector space context used in the example that makes things look more complicated. She then repeats the question about the zero vector in $\mathfrak{R}^{\mathfrak{R}}$: it has to be a function in $\mathfrak{R}^{\mathfrak{R}}$ that satisfies some kind of property. Then Camille asks:

C9: Is the zero vector a function?

T7: Zero vector is a function because all of them are functions here... all the elements...

C10: They are not vectors?

The tutor repeats the definition of the zero element of a vector space and explains that in this case this element is a function that has the properties of the zero element. The students listen and Camille says:

C11: So we are not looking for the zero vector anymore but for the zero function.

The tutor accepts that z(x)=0 is the zero element of $\Re^{\Re}$ *and advises the students to think as if they are in Analysis. So* $z \in U$*, concludes the tutor. Camille quickly explains that* $z \in U$ *because* $z(0)=0=z(1)$*. She also checks out the condition for closure in U. The proof that* $U < \Re^{\Re}$ *is thus completed.*

Extract 8.3

Context: *Student Connie has difficulty with* 'bases and spanning sets'. *Her confusion, she says,* 'became bigger' *in this morning's session — the group tutorial that this tutor teaches to the eight students of this college on the morning of their individual tutorial day. I note that in that tutorial the tutor suggested they replace the expression* $\sum a_i v_i$ *from the definition of*

> *the span of a set* $S=\{v_1, v_2,\dots, v_k\}\subseteq V$ *as* $<S>= \{\sum a_i v_i \,/\, a_i\in F\}$ *where F is the field over which the vector space V is defined*

with the expression 'linear combinations of'. *Connie had used this notation and the tutor recommends caution because* $\sum$ *implies the* 'infinite sum'. *In the Linear Algebra course they will only need the finite sum of vectors to express the elements of an infinite set. The span of a finite set can be an infinite set. For example, on the plane the span of* $S=\{(0,1)\}$*, which is a finite set, is the x-axis. More generally if* $S<V$ *(S is a subspace of a vector space V) then* $<S>=S$*.*

The Episode:

The tutor has just said that if $S<V$ *then* $<S>=S$*. A rather puzzled Connie asks hesitantly if* 'this is because of the axioms'. *Specifically because of closure under addition and scalar multiplication, the tutor adds. He stresses that in the Linear Algebra course of this year they are restricted in the cases where* $S\subseteq V$ *and S is finite. He then asks for her intuitive idea on* $<\{(1,0), (1,1)\}>$*.*

C1:...lamda (1,0)...

T1: What about geometrically?

C2: *x*-axis...

T2: What else?

C3: And the reals... does this become... constant if you take this as zero...?

T3: Yes, but what about geometrically if you do that? What vectors in the plane do you take?

C4: You take any...

The tutor says he agrees with C4: the span of these two vectors is the whole plane and this is the smallest number of vectors that span the plane. These two vectors *form a basis for the plane and there is a theorem according to which all bases have the same number of vectors. This number we call the dimension of the space.*

Connie then asks him to complete the proof that he started in the morning tutorial that

for a set S⊆V where V is a vector space <<S>>=<S>.

The tutor presents the proof (in brief: <S>≤V and <T>=T when T≤V). She then asks him to explain why in LA5.24

rowspace (PA)≤rowspace(A)
for P an rxm matrix and A an mxn matrix.

The tutor starts by defining rowspace(A)=<{r_i}> as the span of the set of columns r_i of matrix A. Connie interrupts to ask:

C5: Is the spanning set the set of all... all the... same property about image... it's all about the values that thing can take?

The tutor advises Connie not to think in terms of THE spanning set since for a subspace there are various spanning sets. He suggests using the expression 'a set being spanned by S'.

C6: So that means it spans something... so it's the values that it takes...

The tutor defines <S> as the set of all the finite linear combinations of the elements of S and reminds her of the plane example. She then asks:

C7: Why are they [*spanning sets*] useful?

Spanning sets are not particularly interesting, he replies; bases are, because they provide with a finite expression of the elements in a vector space. Also a dimension is a characteristic of the vector space and there are important theorems involving the dimension of a vector space. Connie interrupts the tutor in order to return to LA5.24.

C8: Basis is like... from the question here... when we have a spanning set... you use row operations to find a basis in there... and one of them disappears... so is the basis the smallest set of... the smallest...

The tutor repeats that she should avoid expressions such as THE basis of a vector space. He then points at various bases of the plane.

Connie then asks how LA5.24 can be used to prove LA5.25, as requested in the sheet. The tutor explains that by reducing the matrix of the vectors that span U to echelon form, which is what is suggested in LA5.24, we can find the minimum number of vectors that span U. These vectors will form a basis of U and here the reduction to echelon form leads us to a minimum number of 2. 'So is this the basis?' *wonders Connie (C9) and the tutor repeats that this is one basis for U.*

C10: So is this a *2D* basis?... I can't understand intuitively what it is.

The tutor reminds her that as they have always said in the Geometry lessons the plane is two-dimensional and the space is three-dimensional. That need to express vectors in space with three coordinates, or two on the plane, is an intuitive idea of a basis.

C11: And also *x* and y and *r*-thetas are also bases?

T4: *r*-thetas are another story. That's polar coordinates. Think about different bases of this type.

C12: So it's just a way of describing the axes.

T5: Right. If you like... it's a way of describing coordinates.

C13: So this is kind of a funny coordinate system.

The tutor agrees and repeats that there is a variety of bases for a vector space and that the technique he outlined above provided one of them.

Extract 8.4

Context: *In LA6.26 several vector spaces are given. The question is about finding their dimensions. Student Andrew suggests finding a basis of the vector space and counting the number of vectors in it: this number is the dimension of the vector space. The tutor agrees with him.*

The Episode:

Student Jack found out correctly all the dimensions, except one, but what the tutor says he wishes to discuss here is their writing. Andrew's presentation was 'a bit insecure', *says the tutor, and he used more complicated notation than was needed. He then reminds the students of a set of matrices he mentioned earlier in the term which serves as good notation for LA6.26:*

$$E_{pq}=(x_{ij})$$

$$\text{where } x_{ij}=\begin{cases} 0, \text{ if } i\neq p \text{ or } j\neq q \\ 1, \text{ if } i=p \text{ and } j=q \end{cases} = \Delta_{ip}\Delta_{qj}$$

where Δ *is Kronecker's delta*

Jack then offers E_{ii}, $i=1,\dots n$, *as a basis for the set of diagonal matrices. The tutor says he agrees and asks for a basis for the set of symmetric matrices. Andrew protests that he is* 'not quite sure what happens' *and asks* 'What does the definition of E_{pq} mean?'. *Jack explains* 'it's all zero except the *pq*-th entry' *and the tutor apologetically rewrites* x_{ij} *as*

$$x_{ij}=\begin{cases} 1, \text{ if } i=p \text{ and } j=q \\ 0, \text{ otherwise} \end{cases}$$

He also draws E_{pq} *on the b/b as a matrix with all its entries equal to zero except the pq-th entry which is 1. For the symmetric matrices Jack starts manufacturing a basis by including* E_{ii}. *Andrew suggests* 'taking the sum of E_{pq}' *but the tutor does not look pleased. Andrew explains that he wants to 'exploit the zeros and ones in* E_{pq}' *in order to* 'write down all matrices' *but Jack notes this will not happen by* 'adding them'. *The tutor reminds them of the requirements for a basis (span and linear independence) and Jack suggests* E_{ii} *and* $E_{ij}+E_{ji}$ *for* $i<j$, *for i and j from 1 to n. These are* $n+n(n-1)/2= n(n+1)/2$. *For the antisymmetric matrices Jack suggests the same set of matrices except the* E_{ii}, *hence the dimension is* $n(n-1)/2$. *His suggestions are accepted.*

The set of matrices with Trace zero seems to be the most problematic of all (see fig.4 for the solution presented in the tutorial): Jack suggests E_{ij} such that i≠j and he stops in order to think further. Andrew suggests taking these 'together with E_{ii} from 1 to n-1'. *The tutor disagrees: what is $Tr(E_{ii})$? Andrew replies it is zero but with the tutor's reminding of what the trace is he changes his mind to $Tr(E_{ii})=1$. A few silent seconds follow. Then Jack whispers:* 'It's one less because... ' *but Andrew is still struggling with $Tr(E_{ii})=1$.*

A: Why is it one?

T: Because you told me E_{ii}.

A: Yes, but I was thinking of E_{ii} from 1 to n-1.

Jack repeats the definition of E_{ij} and the tutor stresses that E_{ii} do not belong to the vector space of matrices with trace zero. To their silence about how the diagonal elements of the matrices with trace zero will be expressed, the tutor suggests $E_{ii}-E_{nn}$ where $i=1,\dots, n-1$ are matrices with trace zero. These with the matrices E_{ij} for i≠j will form a basis. Span and linear independence are obvious but to Andrew's request the tutor presents them on the b/b. The tutor concludes by outlining the method for finding a basis: 'the principle is to exhibit a set of elements of the set that span the set and are linearly independent'.

Extract 8.5

Context: *The Extract presented here follows Extract 8.4 where the notions of a basis and the dimension of a vector space were exemplified in finding the dimension of the vector spaces given in LA6.26. Moving on to LA6.29 the tutor stresses that behind the proof for this question lies an important theorem about finding bases for a vector space V, two subspaces X and Y and the sum X+Y. To his question whether they can recall the theorem he is talking about, students Jack and Andrew protest about the absence of examples in both Continuity and Differentiability and Linear Algebra courses (*'we've never found a basis for anything'*). The tutor insists on his question about recalling the theorem and Jack dictates*

> $dim(X+Y)=dimX+dimY-dim(X\cap Y)$, *where X and Y are subspaces of a vector space V,*

which he remembers, he says, from his reading, not the lectures. Silence follows the tutor's request for the proof and he suggests applying the theorem on LA6.29ii.

The Episode:

Part a: The tutor reads LA6.29ii and asks what they know about X+Y. Jack replies that it 'has a greater dimension' *but is then silent. The tutor asks them to consider the information they have: X and Y are of dimension n-1 in a space of dimension n. They are still silent. The tutor prompts them:* 'Jack was right about "greater" dimension. It is strictly so'. *Silence. What does it mean* 'bigger than *n*-1'? *Silence. They both whisper it could be* 'anything' *and the tutor reminds them that X+Y cannot be of dimension greater than n. Andrew then deduces that dim(X+Y)=n and the tutor says that then X+Y=V. By replacing the known dimensions in the above formula, $dim(X\cap Y)=n-2$. The tutor says that* 'neither of them looks terribly pleased'. *They* 'look as if they do not accept it', *he says. Jack replies:* 'No, I accept it. It's just based on a formula we haven't proved'. *In a while they will, promises the tutor.*

Part b: First the tutor draws their attention to 'a mistake there that everybody makes' *including Andrew: given a vector space Z of dimension n and two subspaces X and Y of dimension n-1, Andrew considered a basis for Z, and then produced* 'the' *basis for X by removing one vector and produced* 'the' *basis for Y by removing another. The tutor proves (employing the counterexample of $\Re^{27}$ and $\Re^{26}$) that a basis for a vector space does not necessarily contain a basis for a subspace; on the contrary it is true that any linearly independent subset of elements of Z can be extended to a basis for Z.*

Also he stresses that phrasing such as '*the* basis of a subspace' *is misleading because a basis is not unique.*

Part c: The tutor then suggests trying to construct a basis for X+Y that contains a basis for X and a basis for Y. Jack suggests 'We choose a basis for *X* and *Y* first' *but the tutor asks him to* 'forget about *Y*' *for now and start from the idea that we can extend any linearly independent set to a basis for Z. Andrew wonders how starting from a basis for X will allow them to find a basis for Z that contains both a basis for X and a basis for Y. Jack whispers a suggestion: start from a basis for X and then consider the intersection. The tutor agrees: they consider a basis for X∩Y and extend to a basis for X and to a basis for Y (fig.5). Then Jack claims that* $\{u_1, \ldots, u_p, x_1, \ldots, x_q, y_1, \ldots, y_r\}$ *will be a basis for X+Y. The tutor agrees and they prove that this set is linearly independent.*

Once linear independence is proved Jack notes that this actually was a 'quite straightforward' *proof for the formula. The tutor agrees and adds that this not only verifies the formula but provides with a method to extend linearly independent sets to bases for the vector space. Then Jack wonders:* 'Can we extend it for more than two subspaces?'. *Andrew is sceptical: he is concerned whether 'the situation cannot be treated the same way. He suggests considering X to be the sum of two subspaces A and B and check out the consequences. Jack replies that his question was of a more general nature: what happens if we take three subspaces A, B and C instead of two. Andrew asks him to* 'show me you can do it then'. *The tutor thinks it is an interesting question and asks them about the dimension of A+B+C. Andrew thinks it will be* 'the three dimensions minus the dimensions of the intersections'. *The tutor responds that Andrew's suggestion sounds reasonable but reminds them that A, B and C are subspaces, not subsets and therefore they* 'behave more complexly'. *So the* 'analogue' *might not work so perfectly. He nevertheless invites them to experiment with, for instance,* $\Re^{27}$ *and requests they move on to other problems.*

Extract 8.6

Context: *Same college, tutor and students as Extract 8.1 (in the order 2, 1, 3, 4). The mathematical problem in all sessions is the following:*

Given two vector spaces V and W and a linear mapping T: V→W, find matrix A for T.

The tutor suggests the following method:

> *Consider a basis $\{e_i\}$ for V and a basis $\{fj\}$ for W.*
> *Express $T(e_i)$ as combinations of fj.*
> *The coefficients of these combinations form the columns of matrix A.*

The Episode:

Considering a Basis for $P_3(\Re)$. As an application of the tutor's method (see Context), the students and the tutor find the matrix for

$$T: P_3(\Re) \to P_3(\Re) \text{ where } Tp(x)=p(x+1).$$

All agree that, because V=W, it can be $e_i=fj$ and that $\dim P_3(\Re)=4$. In Tutorials 2 and 3 the students give the 'usual' *basis, as the tutor calls it, $\{e_i, i=1,...,4\}= \{1, x, x^2, x^3\}$ immediately. In tutorials 1 and 4 the students seem to have difficulties. In the following, I present the discussion that the tutor's request to the students to give the* 'usual basis' *in Tutorials 1 and 4:*

P1: Is it...1+x....oh no... *The tutor asks for* 'simpler' *polynomials. To her question* 'how $p(x)$ looks like' *Frances replies* 'a x cubed plus b x square plus c x... plus d'. *So $p(x)$ is a linear combination of?* P2: x^3, x^2, x and the constant. T: Well... ? P3: Just one.	*Beth gives* '1 0 0 0, 0 1 0 0... '. *The tutor reminds her that they are talking about polynomials, not matrices. To her silence the tutor asks* 'how p(x) looks like' *and Beth replies* 'a x cubed plus b x square plus c x... plus d' *where a,b,c and d are real numbers. So $p(x)$ is a linear combination of?* B1: Er,... to the s from 1 to 3... *The tutor insists that they are* 'looking for some simple looking polynomials. Nice simple things' *and repeats her question about what* 'simple polynomials' *is $p(x)$ made out of.* B2: Mmm... it's just made by linear ones... T: What do you mean linear ones? B3: It's a product of three... *The tutor realises that Beth is talking about factorisation and she stresses that multiplication of polynomials is not linear. Asking Beth to* 'take a step back' *she repeats that they are trying to write $p(x)$ as a linear combination of four other polynomials $d_1,..., d_4$ which form a basis. She asks Beth what is a linear combination of $d_1,...,d_4$?* B4: $ad_1+bd_2+cd_3+dd_4$... *The tutor says she does not approve of this notation for the coefficients but moves on and asks Beth to compare* $ad_1+bd_2+cd_3+dd_4$ *with* ax^3+bx^2+cx+d *Beth then dictates the* 'usual basis' *$\{1, x, x^2, x^3\}$ for $P_3(\Re)$.*

<u>Express T(e_i) as combinations of e_i.</u> The tutor asks the students to calculate T(1). The discussion in the four tutorials is as follows:

P4: Two. *The tutor says that she knows what Patricia has done. They laugh and the tutor brings their attention to the definition of T. Patricia says it is x+1 and the tutor wonders what happens if there is no x to map. To their silence she notes that* 'in fact there is nothing to do'. *Patricia exclaims* 'ah!' *and says that then T(1) is* 'just one'.	*The students' initial silence is followed by Abidul's suggestion* ('1+*x*') *for which she can however give no reason. The tutor repeats that we have to replace x with x+1 and asks what* 'problem they have with it'. A1: *x* doesn't appear. ... One?	C1: Two. *The tutor disagrees and asks what* 'problem they have with it'. 'Take your polynomial and wherever *x* appears, replace it with *x*+1' *she suggests.* Cleo: Just one.	B5: Polynomial 2. *Asked what she means she replies:* B6: Because we are adding one. *The tutor repeats the definition of T and notes that to add 1 you first have to have an x.* B7: Is it one?

Subsequently the tutor prompts the students with closed questions so that they give the coefficients of T(1) and thus construct the first column of A ((1, 0, 0, 0)). The students immediately follow with the other three columns. In the following, the discussion is from Tutorial 3.

While calculating Cleo asks what would happen if instead of x they had 1+x. The tutor says that Cleo has made 'a good point' *and asks the students whether 1, 1+x, $(1+x)^2$, $(1+x)^3$ would form a basis too. Camille nods and asks if matrix A would then be the diagonal matrix. The tutor nods and says that generally different bases produce different matrices. Camille asks in what cases the matrix is the diagonal matrix. The tutor replies that there are ways of choosing the basis so that the matrix is the diagonal one. Camille asks whether this is because we choose T(d_i) as the basis. The tutor recommends cautious application of the method dictated by the proof of the Rank and Nullity Theorem: Camille recalls that you consider a basis for kerT and extend it to a basis for the domain of T. Then the images of the elements of this basis will form a basis for ImT. Cleo then asks* 'What's this Im?', *the tutor gives the definition of Image and closes by explaining that all the matrices of a mapping produced from different bases are of the same rank.*

Extract 8.7

Context: *The tutor says that students Cathy and George generally did not have many problems with LA7. Particularly in LA7.35 Cathy, he adds, offered a* 'beautiful solution'. *Cathy says she is not very contented with the amount of 'rigour' she used. The tutor invites her to the b/b and asks her to present her proof fully.*

The Episode:

Cathy writes on the b/b her definition of ImT, where T is a linear transformation of a vector space V on V(V 'transforms itself' *says Cathy).The tutor sounds discontented with her definition (Note: unfortunately I only have the modified and final version of Cathy's writing and not what she initially wrote on the b/b) because she has been using the same letter for the elements of ImT and V.* 'They all come from V' *says Cathy, but the tutor explains that this might lead to confusing situations such as Tx=x where x is meant differently. He also stresses that what she has been saying is not what she has been writing:* 'for v in V' *usually means* 'for all v in V'. *He then defines ImT while Cathy is writing the definition on the b/b:*

$$ImT=\{w\in V: w=T(v) \text{ for some } v\in V\}$$

She then says that her solution ' might be wrong now' *and defines ImT^2 on the b/b as*

$$ImT^2=\{T(w) \text{ for } w\in V: T(w)=T^2(v) \text{ for some } w\in V\}$$

The tutor then asks for her notion of T^2. ' Isn't T square applying T twice?' *she says. He suggests thinking of T^2 as just one transformation of V on V. He then turns to her definition of T^2:* 'If you read it in English what does it mean?' *he asks. Silence. To him her definition* 'doesn't make much sense'. *She changes her writing to*

$$ImT^2=\{s\in V: s=T^2v \text{ for some } v\in V\}$$

with which he agrees. She then claims that 'ImT^2 has less vectors than ImT' *but when asked to prove her claim she says she is not sure;* 'ker$T$ is easier to do formally' *she claims. The tutor asks her to prove $kerT\leq kerT^2$ first, if she likes. Cathy then defines kerT as*

$$kerT=\{v\in V: T(v)=0\}$$

and says:

C1: If T sends some v to zero, then T^2 will send the same vector to zero, so it can never be less but it could be more... Is that formal enough?

The tutor says it is ' beautifully literary' *and explains that the reason that all they need to do is prove that $kerT \subseteq kerT^2$ is because both $kerT$ and $kerT^2$ are subspaces of V. He then proves that $kerT \subseteq kerT^2$ and the students similarly prove that $ImT^2 \subseteq ImT$ (by taking a vector in ImT^2 and proving that it belongs to ImT).*

Now they turn to the part of LA7.35 which is about the equivalence of propositions a, b and c. Cathy in her writing had suggested proving that $a \Rightarrow b \Rightarrow c \Rightarrow a$. The tutor says her suggestion is sufficient and the student nods in agreement. George says he proved that $a \Leftrightarrow c$ but could not decide how to go on and stopped at $a \Rightarrow b$. The tutor recommends perseverence. He also points out that b and c look more or less the same whereas a looks different. That means, he adds, that they are probably of the same degree of difficulty and he suggests proving $b \Rightarrow c$, $c \Rightarrow b$, $b \Rightarrow a$, $a \Rightarrow b$. He then asks whether they know anything about the dimension of image and kernel. Cathy replies: the Rank and Nullity Theorem (RNT). The tutor writes

$$dimkerT + dimImT = dimV$$

and asks: 'What this can tell us about the dimension of T^2?'. *Silence follows. The tutor suggests that they* 'have to combine somehow' *the information they have about images and kernels. Cathy then rearranges the RNT in terms of dimImT and sees that b implies that $dimImT = dimImT^2$. George says that if a subspace is contained in another then the dimension of one is $\leq$ the dimension of the other. This is obvious, replies the tutor, and reminds them of $ImT^2 \subseteq ImT$ and $kerT \subseteq kerT^2$. Cathy then points out that then $ImT^2 = ImT$. By the same argument, says George, $c \Rightarrow b$.*

About $b \Rightarrow a$ Cathy says:

C2: If the intersection wasn't zero, then kerT^2 would be bigger, wouldn't it?

Silence follows. George tries to start another suggestion but the tutor stays with Cathy's idea, that is to start from $ImT \cap kerT \neq 0$, which he asks her to pursue. Cathy then says:

C3: If ImT intersection ker$T$ was not zero... then kerT^2 would be bigger than ker$T$.

The tutor says he approves and asks for a proof. Cathy hesitantly says 'It seems true but... '. *The tutor asks for a proof again and Cathy suggests taking $v \in ImT \cap kerT$, $v \neq 0$. Then $Tv = 0$. She also says that then $T^2v = 0$ but the tutor points out that this is not helpful and that it is true anyway because $kerT \subseteq kerT^2$. George then points out that* 'v is special. It is in ImT'. *The tutor asks what this means but the students* *remain silent. Then George suggests* 'apply translation T to... '. *He is interrupted by the tutor who reminds him that* $v \in ImT$ *is an* 'existential' *statement: there exists* $w \in V$ *such that* $v=Tw$. *Cathy then suggests:*

C4: Transform it again and then it is nought so they are not equal.

The tutor rephrases that to: 'then $w \in \ker T^2$ but $w \notin \ker T$. *This contradicts b and therefore Cathy's instinct was right. She also says that they have to prove* $V=\ker T+ImT$ *but the tutor points out this is trivial because if two subspaces intersect trivially then the dimension of the sum is the sum of the dimensions. So* $b \Rightarrow a$.

For $a \Rightarrow b$ *Cathy suggests:*

C5: Go backwards, not backwards as such but if it equals the direct sum and then equate this to the dimension of the brackets, so by the inequality, the intersection is zero.

The tutor says that he thinks that 'the first half of your strategy works. Second half is not clear'. *It's not* 'what he had in mind', *he adds, but he suggests pursuing Cathy's idea.*

C6: If the intersection is zero then surely $\ker T$ cannot be bigger than $\ker T^2$...

T1: Every time you say 'surely'... you give an argument-by-sweet-'n-reasonable Cathy. I mean an argument by voice is an argument by intimidation but you make me say: where is the proof? Surely I agree with you but...

She then suggests: 'Can you say let w be a vector in $\ker T^2$ but not in $\ker T$ and show that this is a contradiction?'. *He agrees but he also adds that he* 'can see no negative fact implied by her suggestion so there is no point pursuing contradiction'. *Contradiction, he explains,* 'only works well if you can use well a negative fact'. *He then presents the* 'positive' *proof, that is one which is not by contradiction, he had in mind (fig.7).*

Extract 8.8

Context: *Two pairs of female students are taught in two tutorials about finding the matrix of a mapping between two vector spaces with respect to some particular bases:*

Tutorial 1: Patricia and Cleo
Tutorial 2: Beth and Cary.

In Tutorial 1 a previous discussion made evident some of the students' difficulties with row operations. In the following, the tutor and the students discuss B3. Some of the students did some work on it but none of them completed the question. B3 (solution provided in fig.8) is about

Proving that E' is a basis of V
Finding the matrix of T with respect to bases E' and F
Finding a basis F' such that a given matrix is the matrix of T with respect to bases E' and F'.

The Episode:

Proving that E' is a basis of V. *The students suggest proving that E' is a basis of V by proving that E' spans V and is linearly independent. The tutor notes the redundancy of their thinking: they only need to prove one of these conditions. The students agree but, when asked why, none of them can answer the question. Only Patricia vaguely replies* 'you can deduce the one from the other' *and, when asked to specify how, she wonders whether it is because we already know that E' is linearly independent. The tutor disagrees and points out that, since dimV=3 and E' has three elements, according to a theorem they have done, all they need to show is either that E' spans V or is linearly independent.*

Finding the matrix of T with respect to bases E' and F. *In Tutorial 2 the students have successfully found the matrix so no discussion of this part of the problem takes place.*
In Tutorial 1 the students remain silent when the tutor asks them about the meaning of a matrix of a mapping T with respect to bases E and F. Then Cary says: 'That *Te* equals *f*. All the elements in it...' *but gives no further explanation. The tutor then asks specifically what Te_1 is going to be. The students are silent and at some point Patricia points at another, irrelevant matrix on Problem Sheet B and whispers something like* 'that's the matrix of the transformation'. *The tutor repeats that her*

question generally concerned the matrix with respect to bases E and F and, in particular, their idea of what the rule is that 'connects the map with the matrix'. *Patricia hesitantly points at the columns of the matrix and whispers:* 'The coefficients of the sums... f_1 plus f_2... '. *The tutor explains that 'the first column gives you the coefficients of Te_1' and, at her request, the students dictate the calculations correctly.*

Finding a basis F' such that the given matrix is the matrix of T with respect to bases E' and F'. In Tutorial 1 this part of the problem is not discussed. The tutor has asked them to complete the question on their own.

In Tutorial 2, the tutor and the students discuss their findings on basis F':

B1: They only have two elements.

T1: Oh, is that the proper conclusion to get to? What do you know about... ?

C1: I've got three. The third one is zero.

T2: Ah, can you put zero in a basis?

C2: *[after a pause]*... You can't.

T3: Right. Because... think about the zero vector...

C3: Oh, do... these can be written...

T4: Yes, it's not linearly independent... your three vectors will be linearly dependent... and what's the dimension going to be?

B2: Two.

T5: Ah, but what this... if you have a basis f_1, f_2, f_3...

B3: *[after a pause]*... Then you have to have three...

T6: Right,... OK, so Cary is right: you've got to find a third one. Em,... what... do you know the first two? *[Silence]* When you are constructing f' how does your reasoning go about what you should put in it?

B4: Te'_1 will be f'_1... and Te'_2 has got to be f'_2...

T7: Right. And we know that in any case that Te'_3 is... ?

B5: Zero.

T8: Zero. So we need... so what condition do we need on f'_3?

B6: *[after a pause]*... To equal zero.

T9: No. We said that it mustn't equal zero.

B7: It's got to be linearly independent with f'_1 and f'_2.

T10: Does it matter what it is?

B8: No.

The tutor stresses that 'you can make f'_3 whatever you please as long as you make it linearly independent with these two' *and Beth chooses* $f'_3 = f_1$. *The discussion closes with a check on linear independence and with the tutor's generalisation from 3x3 matrices to mxn.*

Appendices for Chapter 9

GRF5.1

If p is a prime, and H and K are subgroups of a group G, each of order p, show that $H \cap K = \{e\}$ (use Lagrange).
If F is finite, deduce that the number of elements of order p in G is a multiple of $(p-1)$.

GRF5.6

Use GRF5.1 to show that a group of order 35 must contain an element of order 5 and an element of order 7.

GRF5.8

A subgroup H of a group G is called a proper subgroup if $H \neq G$. If G is finite, prove, or give a counterexample to, each of the following:

(a) G cyclic $\Rightarrow$ every proper subgroup is cyclic,

(b) every proper subgroup is cyclic $\Rightarrow G$ cyclic,

(c) $|G|$ is prime $\Rightarrow G$ has exactly one proper subgroup,

(d) G has exactly one proper subgroup $\Rightarrow |G|$ is prime.

GRF7.3

(i) Define the left cosets of a subgroup H of a group G to be the sets $gH = \{gh: h \in H\}$.
Show that H is a normal subgroup of G if and only if $gH = Hg$ for all $g \in G$.

(ii) Deduce that if K is a subgroup of G with $[G{:}K] = 2$ then $K \nabla G$.

GRF8.5

Check that if $xy = yx$ in a group, and hcf $(o(x), o(y)) = 1$, then $o(xy) = o(x).\, o(y)$.

Appendix 9B

If $|H| = |K| = p$ and p is a prime then H and K contain, each, p elements of which p-1 are of order p. If H and K have a common element $\neq e$, then this element is of order p and it is contained in both H and K. Therefore it generates H and K. Because H and K are of the same order and are generated be the same element, $H = K$. Hence H and K either coincide or have only e as a common element.

Moreover the above implies that H and K together contain $(p-1)+(p-1) = 2p - 2 = 2(p-1)$ distinct elements of order p. Similarly, if G is finite, for every element g of order p, p-1 elements of order p are contained in G: the powers of g from 1 to p-1. Therefore the number of elements in G of order p is necessarily a multiple of $(p-1)$.

fig.1 GRF5.1

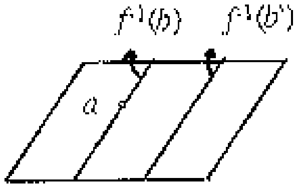


fig.2a

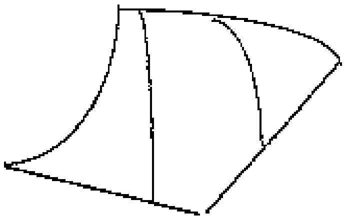


fig.2b

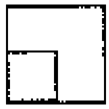


fig.2c

If $|G| = 35$, then from Lagrange, its elements are of order 1, 5, 7 or 35. In particular an element $g \neq e$ is of order 5, 7 or 35.

If there is an element g of order 35, then it generates G, that is G is cyclic and g^5 is of order 7 and g^7 is of order 5.

If there is no such element of order 35, then an element $g \neq e$ is necessarily of order 5 or 7. Not all such elements can be only of order 5 or only of order 7. Say, for instance, that there is an element g in G which is of order 5. Then, from GRF5.1 (fig.1) we know that, because 5 is a prime and G is finite, that the number of elements of G which are of order 5, must be a multiple of 5-1 = 4. But 34 is not a multiple of 4, thus G must contain elements of order 7. Similarly, G cannot contain only elements of order 7 because 34 is not a multiple of 7-1 = 6.

fig.3 GRF5.6

First Isomorphism Theorem for Groups

Let G and G' be groups and $\phi:G\to G'$ a homomorphism. Then $\ker\phi=K\nabla G$. Then $G/K \sim \operatorname{Im}\phi$.

Proof

To define G/K I need to prove that $K\nabla G$, that is $gK = Kg$ for all $g\in G$.

Let $gk_1\in gK$, where $k_1\in K$. Then $gk_1= (gk_1g^{-1})g \in Kg$ because ϕ is a homomorphism and $\phi(k_1)= e$. Therefore $gK \subseteq Kg$. Similarly $Kg \subseteq gK$. Hence $gK = Kg$.

To prove that $G/K \sim \operatorname{Im}\phi$ I need to define an isomorphism between G/K and $\operatorname{Im}\phi$.

Let

$$\psi: G/K \to \operatorname{Im}\phi$$
$$\psi(Kg) = \phi(g)$$

To prove that ψ is an isomorphism, I need to prove that it is a well-defined, 1-1, onto homomorphism between G/K and $\operatorname{Im}\phi$.

Well-defined: $Kg_1=Kg_2 \Rightarrow \psi(Kg_1) =\psi(Kg_2)$ (i.e. $\phi(g_1)=\phi(g_2)$) because $Kg_1=Kg_2$ implies that $g_1\in Kg_2$, that is there exists a $k\in K$ such that $g_1=kg_2$, and $\phi(g_2) = \phi(kg_2) = e\phi(g_2) = \phi(g_2)$.

1-1: $\psi(Kg_1) =\psi(Kg_2) \Rightarrow Kg_1=Kg_2$ because $\psi(Kg_1) =\psi(Kg_2) \Rightarrow \phi(g_1)=\phi(g_2)$ which gives that
$g_1g_2^{-1}\in K$.

onto: If $g\in \operatorname{Im}\phi$ then $g\in Kg\in K$. I assign $\psi(Kg)$ to be g therefore ψ is onto.

homomorphism: $\psi(Kg_1Kg_2) = \psi(Kg_1g_2) = \phi(g_1g_2) = \phi(g_1)\phi(g_2) = \psi(Kg_1)\psi(Kg_2)$

fig.6 The proof of the First Isomorphism Theorem for Groups

If $[G:K] = 2$ then there are two cosets of G over K. We know that one of them is K and, because the cosets cover G, the other necessarily is G-K. K and G-K are the two right and the two left cosets, that is, if $[G:K] = 2$ then its right and left cosets coincide, therefore, from the definition given in GRF7.3i, $K\nabla G$.

fig.7 GRF7.3ii

Appendix 9C

Extract 9.1

Context: *Student Connie has asked the tutor about the meaning of generator in Group Theory. She has also asked the tutor to explain Alan's (another student from the same college) solution for GRF5.1 presented in that morning's session to all the students of this college. The tutor says she lacks clarity in her conception of* 'order' *either of a group or of an element and defines the concept:*

> *n is the order of an element g in a group G if it is the least positive integer for which $g^n=e$, where e is the identity element of G.*

He then explains that

- *if G is a finite group, then n exists because* 'all powers of g would have to start repeating themselves at some point'.

- $<g> \leq G$, *the inverse of an element g^t in $<g>$ is g^{n-t} and there are exactly n elements in $<g>$.*

The Episode starts with Connie's question about $<g>$.

The Episode:

Connie is asking about $<g>$:

C1: Is this a cyclic group?

T1: Yes, a cyclic subgroup.

C2: Were we talking about them there *[in GRF5.1]*?

The tutor explains that they haven't mentioned it yet in relation to GRF5.1 and that he is still trying to clarify the meaning of order of an element or of a group for her. Another thing 'she needs to know' *is, he says, Lagrange's Theorem:*

if $H<G$ then $|H|/|G|$

because the cosets of H in G are 'distinct chunks of G of equal size'. *Also, he adds,* $|g|/|G|$ *and if ordH=p where p is a prime number then the p-1 elements of H, except e, have order p. He then turns to GRF5.1 (fig.1) and explains that first they showed that if two subgroups of order p overlap only at the identity, they together contain 2(p-1) elements of order p.*

C3: Well, I thought that H had just p elements because when you've got these groups when you take the order of the group you count the... identity?

The tutor repeats that each of the elements in H, except e, is of order p.

C4: I don't understand how an element can have an order.

The tutor repeats the definition of the order of an element and adds that 'the reason it's called order is because it's the order of the least subgroup which contains the element'.

C5: Does this mean if you take this little g times p-1 you would have the same element?

T2: That's right. If you take $g^2,\ldots,(g^2)^{p-1}$ you would take all these elements in some different order. Is that what you are saying?

C6: No, I was just saying if you take g^2 times g^{p-1}...

The tutor explains that $g^2g^{p-1}=g$ and stresses that every other element in H will be of order p. So, he concludes, if G has r subgroups like H, only coinciding at e, and H contains p-1 elements of order p then G contains r(p-1) elements of order p. He stresses that the notion of the order of an element is 'subtle but confusing'.

C7: So you don't do g times p-1 but g to the $2p$-1...

T3: I'd do this if I wanted to find the order of g^2. I would square it and cube it and so on.

C8: Ah! And you would time these together?

T4: Yes, like this is g^2... this is g^6...

C9: Are these separate things?

T5: Yes, I am looking at g^2, g^2 all squared... and you don't hit the identity until you get p again.

C10: OK...

Extract 9.2

Context: *This is the beginning of the tutorial with student Camille.*

The Episode:

Camille says she has been confused with 'what the lecturer referred to as something like $a \sim b$ when $f(a)=f(b)$'. *The tutor replies that the lecturer might have been trying to prove one of the Isomorphism Theorems for Groups:*

> *if $f: G \rightarrow G$, where G is a group, then there is a bijection between the equivalence classes of the equivalence relation defined by '$a \sim b$ when $f(a)=f(b)$' and Imf.*

The tutor draws fig.2a to illustrate that if b and b' are different then $f^{-1}(b)$ and $f^{-1}(b')$ are different. Camille looks at fig.2a and asks:

C1: Why are they all straight lines?	T1: Because I drew them this way. It doesn't mean anything. It's only the way I drew it. And you can draw little 'squigges' if you like *(fig.2b)*.
C2: And these are the equivalence classes?	

The tutor nods in agreement: the set of all elements equivalent to b is the equivalence class of b and mapping b to its equivalence class establishes a correspondence between the elements of the group and their equivalence classes. Camille seems to be sceptical and says:

C3: Yeah... but the lecturer does it in a more complicated way... he defines an equivalence class like this and then he defines g of e_a equals $f(a)$. It's very confusing because we don't know what g is.	T2: *[pointing at her correspondence]* That's his definition of g, is it?
C4: When he proves it I don't see at all a bijection.	T3: *[reassuringly]* What he's done is equivalent to what I have done except that I've phrased it slightly differently.

C5: You've shown us a bijection and he didn't really show that... he just had it onto by definition and he was looking at the 1-1.	

The tutor then says that it has to be proved that to define an equivalence class it doesn't matter which representative you take. So, she continues, in $e_a=e_{a'} \Leftrightarrow a \sim a' \Leftrightarrow f(a)=f(a')$ *one direction is about proving that f is well defined and the other one is to prove it is 1-1. Camille looks convinced and moves on to her next question: can the tutor prove that* 'there is a 1-1 correspondence between the conjugates of x and x?'. *The tutor asks Camille to define C(x) and Camille defines the centraliser of x as*

$$C(x) = \{h \in G / xh=hx\},$$

the conjugate of x as

'$g^{-1}xg$ *for some* g'.

and also points out that

if $g \in C(x)$ *then* $g^{-1}xg=x$.

So, says the tutor, the size of C(x) tells you 'how many repetitions you get if you try to write x as a conjugate'. *So, she continues, if two elements* g_1 *and* g_2 *give the same conjugate they belong to the same coset of C(x). And vice versa. This happens only when* $g_2g_1^{-1}$ *belongs to C(x). Therefore there is a 1-1 relation between cosets of the centraliser and conjugates of x and these have the same number of elements. Camille is quiet and looks sceptical. Then she asks:*

C6: What are cosets materially?	T4: What do you mean by that?
C7: If we have a group *G* and a subgroup *H* why do we bother to find the cosets?	T5: Because of results like this. They turn up naturally.
C8: Cosets are a group multiplied by an element in the big group.	T6: *[hesitantly]* Yes...it's a set...
C9: Cosets are just a moving...	T7: That's right. That's one way... you can look at it as translates of a subgroup... sort of multiplying *g* with everything in *H* and it shifts it...

C10: *[after a pause]* So if we have a square of size one and then the group G is like this...*(fig2c)*	T8: You have to be slightly careful... It is slightly... don't think about in... you're not thinking of applying it on squares, are you?
C11: So then it would have four cosets?	T9: Mmm... if the subgroup was a quarter of the size of the whole thing, yes... it would have four cosets... that's right.
C12: And the cosets are always the same size as the original.	T10: That's right. As we know they partition the group.

Extract 9.3

Context: *This is the beginning of the part of the tutorial on Groups, Rings and Fields for students Jack and Andrew. The tutor asks Jack to present his proof for GRF5.6 on the b/b.*

The Episode:

Jack suggests:
Consider $x \in G$. Then $<x> = \{e, x, x^2, ..., x^{k-1}\}$ for some $k \leq 35$. Then $|x|=1$ or 5 or 7 or 35. If $x=e$ then $|x|=1$. Say then $x \neq e$. Then $|x|=5$ or 7 or 35. If $|x|=35$ then $<x>=G$, that is G is cyclic and $|x^5|$ is of order 7 and $|x^7|$ is of order 5. He then explains what would happen if he had picked x such that $|x|=5$:

J1: So there's only two cases to look at: say I had picked an element of order 5. Then... em,... it's OK if I use a problem from the sheet before? So, *GRF5.1* shows that the number of elements... we can define a subgroup with four elements... what?... that says that the number of elements of order 5 will be a multiple of four. Because you've been defining discrete... disjoint subgroups... er,... but here actually we've got 35... except the identity we've got 34 elements... but 4 actually won't go into... doesn't divide 34. So therefore there must be... you can't just have subgroups of order 5... you must also have another subgroup... so if you had a subgroup of order 5 then you would have to have a subgroup of order 7. And alternatively, because 6 doesn't go into 34,...

T1: You seem to have the right idea but you seem to have phrased it a bit... differently. You said you choose an $x \neq e$ in G of order 5. But that doesn't discard the fact that there might be elements of order 35 in there, does it?

J2: Yeah, it could have been a cyclic group.

T2: So I mean you are right: there's 34 elements other than the identity and that, if it isn't cyclic they have orders 5 or 7. But because you choose an element of order 5 it doesn't mean that there are no elements of order 35. So what you have to say is consider the case when G is cyclic - which you did - and then take the case where G is not cyclic, not suppose that x had order 5 because x has order 5 doesn't mean that 34 must have five and seven.

J3: But I took a partition of different cases,...

The tutor stresses that Jack's taking cases was fine but his 'arguments weren't just identical to the right arguments' *for this (fig.3). What one should say, stresses the* *tutor, is: if G isn't cyclic then there are 34 elements all of order 5 or 7 because there is no element of order 35. The tutor concludes:* 'it's a minor point but try to make clear which argument you are following down'. *Jack asks Andrew if he* 'got this'. *Andrew nods. The tutor reassures them it was a minor point and moves on. They had no problem in GRF5.7 so they go to GRF5.8 (see Extract 9.5).*

Extract 9.4

Context: *The tutorial with student Connie started with a discusion of GRF5.8a and d.*

The Episode:

C1: I don't really understand what they are.

The tutor then defines the centraliser of an element x in a group as

the set C(x) of elements that commute with x:

$$C(x) = \{y \in G \,/\, xy=yx\}$$

C2: You can also swap back to *y* over there, can't you?

T1: But this is a different thing though: you can say that the centraliser of *y*, if y is an element of *G* is the set of *x* in *G* such that *xy=yx*, but is that what you mean? *[Connie nods]* It's what I'm saying here I just changed the dummies so to speak.

C3: Can you give me an example so that I can understand?

The tutor then proves that

if G is an Abelian group then, $\forall x \in G$, $C(x)=G$.

Also he proves that for $(12) \in S_3$, (13), (23) *and* $(123) \notin C((12))$. *Therefore* $C((12))=\{e\}$. *Connie interrupts him in order to ask about (12) itself: does it belong to C((12))? The tutor apologises: so, he concludes, C((12)) contains e, (12) and its inverse.*

C4: We were introduced to... We've defined conjugate as *x* times *g* times x^{-1}... I mean how do you choose *g*? *g* is an element in the group...?

T2. The tutor defines the set C_g *of conjugates of g. In an Abelian group an element g is only conjugate to itself.*

C5: So it's for every *g* belonging to *G*?

T3: For all *g* belonging to *G* you run through them all and you get what you get as a set here.

C6: *x* has several conjugates then?

The tutor nods and says that if G is Abelian then $C_x=\{x\}$. *Then he suggests they find* $C_{(12)}$ *in* S_3. *Connie is confused,* 'it is too complicated' *she says about the* *calculations involved. (Note: I omit the detailed conversation on these calculations since it would probably distract the reader from the main point of the Episode which is Connie's effort to clarify the notion of conjugate of an element). While the tutor's calculations for $C_{(12)}$ are still incomplete, Connie, who has been sceptical and quiet all this time interrupts the tutor in order to ask:*

C7: What does the conjugate actually mean? Because the inverse is actually going to send it back to itself... what are you using it for... I mean... because I am always getting confused with inverses and conjugates.

T4: Yeah, they are different things...

C8: What's you need it for? What'll do?

The tutor warns her she 'is leaping ahead' *and says that they haven't finished the example from S_3. He then explains the connection between conjugacy classes and the definition of normal subgroups. He mentions an example from S_4 as well as other uses of conjugates: for instance proving that if* $|G|=p^2$, *where p is a prime then G is Abelian, involves conjugates.*

Extract 9.5

Context: *Extract 9.5 follows Extract 9.3 with students Jack and Andrew. Similarly to other tutorials, Jack and Andrew say that the lecturer has introduced the notion of ideal set in order to prove GRF5.8a. The tutor points out* 'this is too heavy a notion for what you have to prove here' *and adds there is* 'nothing particularly difficult here' *and you do not need to reach conclusions* 'by some amazing theorem'. *He finally stresses that* 'you're almost skipping the mathematics by throwing in a notion like that' *(the notion of an ideal set). The tutor then turns to GRF5.8d. He says that he thought Jack's proof was clear and invites Andrew to* 'convince' *them of his unclear, but otherwise right, argument for GRF5.8d.*

The Episode:

Andrew comes to the b/b and talks about what he wants to prove: if |G| *is not prime then there is more than one proper subgroup. So, he adds, then we know (Note to the Reader: from GRF5.8c) that* |G| *is prime iff G has one proper subgroup. He then turns to the b/b and starts writing. At the same time he explains:*

A1: If $|G|$ is not prime, OK?, then we may write its order as pq for some... we don't need them to be 1 because... it's a very similar idea to *GRF5.1* in that... Let's choose x in G, OK? We take out the identity. OK, now since G is finite x has got some order, like we did before, and what we can say is that x, the order of x equals 1, pq or...? What's it called if it's neither pq or 1? *['proper factor' replies the tutor]* Proper factor! The idea is that we kind of look first at the easy cases when x is a generator of this group... and we can disregard this case and look at x as a factor of this. What I am trying to show is there exists at least two proper subgroups and I'll explain... Right, order cannot be one because x is not the identity. If the order of x was pq, er,... then, er,... Is there a standard notation for the order of x? Is this OK? *[$|x|$ is fine says the tutor. There are other ones but none of them is standard, eg gp<x>]* Right. So if the order of x is pq then the set generated by x^p and the set generated by x^q are both proper subgroups of G because if you generate them you go through x^p x^{2p}, x^{3p} and then x^{pq} and you go back to the beginning again. So these two subgroups...

Jack interrupts in order to ask: 'Can p equal q?' *and the tutor stresses that this wouldn't matter. However, he continues, Andrew must remember that he only has to construct one proper subgroup of G. {e} is the other one. Jack insists that he is* 'still not happy with the logic behind that'.

A2: OK, I'll go through that again. If x has order n and n divides pq then $\{e\}$ and *[$<x^n>$]*... are proper subgroups of G because I mean n, by definition this is a group.

So I am showing that if $|G|$ is not a prime then its order is always a product of two integers and we can show from this, by taking different cases as in *GRF5.6* that then it has at least two proper subgroups... in the case where $|x|$ is not 1 because x is not $\{e\}$, if $|x|$ is pq then the set generated by x^p is a subgroup so that makes two...

The tutor sounds ready to interfere (he exclaims 'Er,...!'*) and Jack reacts with* 'I do not see the point of doing this'. *The tutor says he is* 'not convinced by the last argument' *and notes that if x is of order n then* $<x^n>=\{e\}$. *Andrew looks confused. He is staring at the b/b.*

A3: Oh, you mean...? *[pause]* Oh, in that case... er,... I must be able to choose something! If... if x is of order n,... but that means, oh! We just want... what am I doing? *[still staring at the b/b]* Oh, dear! *[turns around and explains]* What we are showing is that if $|G|$ is not prime, then there exist at least two subgroups. So if sets exist which only contain one subgroup, they must be prime. We've shown before that all groups of prime order have only one proper subgroup *[pause]*. So, if n is the order of G and n is not a prime then for a factor of G there must be a proper subgroup of G of order equal to that factor.

The tutor and Jack nod approvingly and the tutor closes the discussion of this question with stressing that the converse propositions of Lagrange's Theorem are not true.

Extract 9.6

Context: *In Tutorial*

1: Eleanor
2: Cary and Beth
3: Cleo and Patricia
4: Abidul and Frances

the tutor suggests they prove the First Isomorphism Theorem for Groups:
Let G, G' be groups and $\phi: G \to G'$ a homomorphism. If $K=\ker\phi$ then: $G/K \sim Im\phi$.
Proof (fig.6): since $K \nabla G$, G/K can be defined. Then

$$\psi: G/K \to Im\phi$$
$$\psi(Kg)=\phi(g)$$

is an isomorphism, namely ψ is a well-defined, 1-1 and onto homomorphism.

In the following, I present extracts from the proving process in the four tutorials.

The Episode:

In the four tutorials the conversation starts with the tutor's request towards the students to state the theorem.

I note that by (*) I mean the expression

$$Kg_1=Kg_2 \Leftrightarrow \phi(g_1)=\phi(g_2) \quad (*)$$

<u>Tutorial 1.</u> Eleanor listens to the tutor silently and throughout the presentation of the proof she only answers the following very closed questions:

- *The tutor repeatedly requests the student to recall the theorem. Finally Eleanor hesitantly says that she has associated the theorem with a normal subgroup. The tutor insists that it is a particular normal subgroup and Eleanor remembers $\ker\phi$. She responds affirmatively to whether $\ker\phi < G$ and $\ker\phi \nabla G$. She also dictates the calculation to why $gK=Kg$.*

- *Once the tutor has illustrated verbally that the number of cosets in G/K is the same as the number of elements in $Im\phi$, she asks Eleanor to see the* 'obvious map' *between G/K and $Im\phi$. Eleanor whispers 'Kg' and the tutor introduces ψ.*

- *While proving that ψ is an isomorphism between G/K and $Im\phi$, the tutor asks what the $\Leftarrow$ of (*) proves. Eleanor replies* 'onto' *and, after the tutor explains to her that* 'onto' *is obvious from the definition, she changes her mind to* '1-1'.

- *Finally, while Eleanor dictates correctly the rule for the product of cosets ($Kgh=KgKh$) she remains silent when the tutor asks her to use it in proving that ψ is a homomorphism. Eventually, under the tutor's firm direction she does.*

<u>*Tutorial 2.*</u> *I note the following:*

- *Beth states the theorem. The tutor corrects: ϕ can be between any two groups G and G' and not necessarily on G itself.*

- *Similarly to Eleanor, in Tutorial 1, Beth also suggests mapping an element of Imϕ to its coset.*

- *In proving (*) Beth recalls (from the group tutorial in this college the day before) that two cosets Kg_1 and Kg_2 are equal iff* '$g_1g_2^{-1}$ is in the kernel'. *Also both students dictate the necessary calculations until $\phi(g_1)=\phi(g_2)$ is reached.*

- *Beth knows that the reverse direction of (*) proves that ψ is 1-1.*

- *Once well-definedness, 1-1 and onto is proved the tutor asks for the last thing to prove and discussion is as follows:*

B1: That it is closed.

T1: What do you mean by "closed"? If we are talking about an isomorphism of groups and we've proved it's a bijection...?

B2 and C: It's an isomorphism.

T2: Right. So it is a homomorphism because..? What do you want to check?

B3: ϕ of g_1g_2 is ϕ of g_1 times ϕ of g_2.

T3: Ah, we are talking about ψ.

B4: Oh, then ψ of these...

T4: Er,... ψ is...

B5: Oh, Kg_1 then and Kg_2.

Beth then dictates the calculations.

<u>*Tutorial 3.*</u> *I note the following:*

- *Patricia says that she can only recall that the theorem is related to cosets and a mapping which she says she thinks is an isomorphism. The tutor corrects: it is a homomorphism. She then asks for the conclusion of the theorem. Patricia replies:*

P1: *G*... divided by... kernel... er,...*f*...Does this *[~]* actually mean equals?

T1: Isomorphic. It means isomorphic

- *Similarly to tutorials 1 nad 2, Cleo suggests, after listening to the tutor's explanations, mapping Kg to g. The tutor asks her to reverse that (*'it is the other way around'*) to g to Kg.*

- *The tutor continues the presentation of the proof without any participation until she asks what is the meaning of $g_1g_2^{-1} \in K$ while proving (*). The students are silent She returns to the* 'basic notion' *of a kernel and asks what a kernel is:*

C1: It maps the elements at zero.

T2: Not zero!

C2: Identity.

T3. *[writes down the definition* 'properly'*: $ker\phi=\{g \in G /$. She leaves it unfinished and asks]* Such that?

C3: Maps them to the identity.

T4: Which means something is equal to the identity.

C4: *[after a pause]* g does.

T5: When you say something goes to the identity, is equal to the identity, what do you mean by that?

C5: ϕ does.

- *The tutor then returns to the proof of (*). Starting from $g_1K=g_2K$ she asks the students to prove that $\phi(g_1)=\phi(g_2)$. The students are silent. The tutor points out that two cosets $Kg_1$and Kg_2 are equal when $g_1g_2^{-1}$ is in K. The students are still silent. She returns to the basic notion of a kernel (see conversation above). In the end of that discussion Patricia points out that $\phi(g_1g_2^{-1})$ will be the identity. With gradual prompting by the tutor the students arrive at $\phi(g_1)=\phi(g_2)$. The tutor points out that $\Rightarrow$ of (*) means ψ is well- defined. What about $\Leftarrow$? she asks. The students are silent. Then they both mumble something about $\phi(g_1)=\phi(g_2)$ saying something about ϕ. The tutor stresses that their answer must be related to ψ, not ϕ. The students look lost. The tutor then repeats that they have to prove that ψ is an isomorphism. The students are still silent and the tutor asks what an isomorphism is.*

P2: It is a homomorphism..

T6: Yes? And...?

P3: It has to be 1-1...

T7: And...?

P4: A bijection...

T8: Yes...

P5: Onto.

T9: Yes and in other words which bit have we proved here? You've got a choice: you can say it's a homomorphism, it's 1-1 and it's onto *[laughter]* Which bit have we proved? *[pause]* Is it the bit that says it is a homomorphism?

P6 and C6: *[after a pause]* No.

T10: Is it the bit that says it's 1-1? *[pause]* What do you need to prove ψ is 1-1?

C7: Prove... er,... oh, it's this direction that proves it's a well-defined operation... so it's the other one that proves it's 1-1.

ψ is obviously onto, says the tutor and suggests that they prove ψ is a homomorphism. Patricia hesitantly dictates the equality they have to prove: I note that only with the tutor's prompting she recalls and uses the rule for the product of cosets. The tutor concludes that the key point about ψ is that the homomorphic property works ('The point here is that this thing works').

<u>Tutorial 4.</u> *The students cannot recall the theorem and remain silent. The tutor then tells them off:* 'Number one thing is you learn the theorems. Have you learned what the theorem says? Specially when it's something that's got a name attached to it'. *She then states the theorem.*

- *Similarly to previous tutorials Abidul suggests mapping g to Kg.*

- *Similarly to previous tutorials Abidul recalls that two cosets Kg_1 and Kg_2 are equal* 'When $g_1g_2^{-1}$ is in K' *which is true iff* '$\phi(g_1g_2^{-1})$ is one'. *Frances uses the homomorphic property to manipulate $\phi(g_1g_2^{-1})$ but* 'gets stuck' with the $^{-1}$. Abidul helps her by pointing out that ' It [ϕ] doesn't really matter, does it?' *that is $\phi(g_2^{-1})=\phi(g_2)^{-1}$. Thus the well definedness is proved. Frances points out that the $\Leftarrow$ direction of (*) shows that ψ is 1-1 and the tutor explains why ψ is obviously onto.*

- *The discussion for proving that ψ is an isomorphism is as follows:*

A1: ψ of Kg_1g_2 is ψ of Kg_1....

T1: Say that again.

A2: ψ of Kg_1, Kg_2...

T2: Right. And what did you say at the beginning?

A3: ψKg_1g_2...

T3: I mean you have to show that ψ of the product is...

Frances finishes the tutor's sentence (applies the definition of multiplying products) and the proof is completed. The tutor closes by calling this 'a very standard method of proof involving quotients'. *The problem that they have to be concerned is well-definedness. She also asks them to learn and remember the theorem because it is very important.*

Extract 9.7

Context: *This is the beginning of the Abstract Algebra part of the tutorial to student Connie who asks for some clarifications on the proof for GRF7.3 presented by student Nick in the group tutorial the same morning (fig.7).*

The Episode:

First Attempt at Explaining the Argument in GRF7.3

C1: I don't quite understand what he is... he is showing that the left cosets equal the right cosets.

The tutor says that this is not a very 'precise' *phrasing and explains that if [G:K]=2 then there are two cosets, K and G-K which are both the right and the left cosets. Connie then points at G-K and asks:*

C2: Yeah... Is that what you call quotient?

T1: No. This is a complement. This is a set-theoretic term only. I mean it in a set-theoretic sense. Yes, you are quite right to ask me that. The quotient of G over K is written like that *[G/K]*, you are thinking of that as a fraction in a rather vague, general sense... but this is like minus. But then minus has some funny connotations as well so let's use it like that which is the purely set-theoretic sense. Alright?

He then repeats that if we have two left cosets one of them will be K and the other one, since the union of cosets must be the whole group must be the rest of the elements in G. The same goes for the right cosets, he adds. Connie sighs with disappointment and he suggests 'getting into more detail'.

Explaining the Construction of Cosets

T2: Remember that the idea of cosets is that you have this big group with its elements...

C3: And you are multiplying...

T3:... and you've got a subgroup and you start having parcels which are the same as the subgroup. Forget about multiplying because this is not... well, we are trying to prove this is normal, we don't know yet...*[repeats that in case [G:K]=2 each of the two cosets has half of the elements of the group in it].*

C4: K is a subgroup then?

T4: K is a subgroup.

C5: And a coset?

T5: A coset isn't a subgroup. Only when you have a subgroup K over G, K is the only coset that is a subgroup. All the other elements don't have the identity in them. It's not possible for them to be subgroups. No chance if they don't have the identity in them. It's a partition of the whole group into subsets.

Second Attempt at Explaining the Argument in GRF7.3

C6: So you just said that you can only... that the coset of K in G is K.

T6: That's one of them. There is another which is all the other elements in the case where there is only two. In general the picture is like... if we had three in this case... but we are in a special case which is looking like that...*[draws a group and splits it in two parts, putting K and G-K in each part].*

C7: So there is two subgroups... half of G.

T7: When the index is 2 which is saying that there are two cosets yeah. The index of G is half the size of G. If you like to talk about a finite situation.

The Case of [G:K]=3 as opposed to [G:K]=2

C8: If it was three then...?

The tutor explains that the 'argument wouldn't work at all'. *There would be two more cosets, he continues, so G-K would be the union of two cosets and this wouldn't be saying anything. Of course K would still be one of the cosets. The case where there are two cosets is very convenient because then you know that K is one and the rest of the group is the other.* 'Now', *he closes,* 'that argument works for both right and left cosets you are talking about'.

C9: But if you had G divides K, no... index of G of K is 3... then... you'd have K being one...

T8: Then you would have g_1K there and g_2K there...

C10: And those two would be another case then.

T9: No, this is one coset and this is another.

Connie nods 'alright' *and the tutor adds that, since G* 'is a disjoint union of cosets' *if the index was three then the picture of G would be split in three cosets one of which is K and the others are g_1K and g_2K. Connie points at these other two parts.*

C11: And these two things aren't subgroups...?

The tutor replies that cosets are disjoint and the identity element is contained only in K so the other two cannot be subgroups. 'These are', *he says,* 'translates of subgroup... as you might say by a group element which is not the identity element'.

Explaining Again the Construction of Cosets

Connie then asks about g_1 and g_2:

C12: And is that any group..is that any element of the group *G*?

The tutor adds more cosets in his picture and explains that 'you can choose any element of the group which is not in *K* and it will give you a distinct coset. Then, once you've done that, if you want to look for another coset, you have to take any element which is not in this and not in this. And then it will give you a fresh coset'.

C13: And these are going to be in *G* because it's closed.

He nods and repeats the generation of cosets once more.

C14: So if *G* over... if the group's index is prime...

T10: Hold on! Say that a bit more carefully.

C15: If the index of the group is prime...

T11: You mean if $[G:K]$ is prime?

C16: No. *G* is...

T12: Oh, you mean the order of *G* is prime. Well, then there are no interesting subgroups.

C17: Oh it's only *G* then, isn't it?

T13: *G* and the identity subgroup.

C18: And then you have one coset?

'Well, you've got to have a subgroup before you start talking about cosets' *he replies. He then explains that if we take G as the subgroup then there is only one coset and this is G. If on the other hand we take {e} to be the subgroup then we split G in cosets that contain every single element.*

Third Attempt at Explaining the Argument in GRF7.3

Connie then returns to the case where $[G:K]=2$.

C19: Right. And so this is saying here... that the left cosets are equal to the right cosets.

'Well, we haven't quite pinned that down yet but that's the basis for it' *he replies and repeats his argument: we know that there are two cosets, that one of them is K and that cosets are disjoint and their union is G. Therefore the second coset must be G-K. The argument works for right and left cosets alike.*

Again the Case of [G:K]=3 as opposed to [G:K]=2

Connie is still concerned with the case [G:K]=3:

C20: What about this case?

T14: Well, that doesn't apply. You can't make such an argument in this case.

C21: *[after a pause]* Right. And why not?

T15: Because if you take *K* here, then the rest of the elements in the group they don't form a coset. They form a union of 2 cosets.

C22: But if you take these three things individually... these three sets... then... do they all go... these three are the cosets... the union of these two isn't...?

T16: No, each individually is, yes.

C23: And then could you say that the right cosets of these three things are equal to the left cosets?

T17: No. Because they might not look at all similar.

The tutor stresses that in this case the right and left cosets 'might not look at all similar' *and illustrates in paper how multiplying with K from right and from left could lead to two different sets gK and Kg. So, unless K is a normal subgroup, it is not possible to know whether the right and left cosets are the same: once you marked off K the other elements can be in any of the other two.*

Explaining the Construction of Cosets Again

C24: So when you have *g* times this element and *g* times this element which means that this can be a left coset... or I don't know which way it goes left or right...

T18: *gK* is a left coset.

C25: And then... does that mean you can put an element in here and times these things?

T19: What we are actually doing is picking any odd element in the group which is not in K and multiply it with everything in K putting g in front.

Fourth Attempt at Explaining the Argument in GRF7.3

C26: Right. So that says that you can have these...

T20: That's a coset. But we've proven that there's only two cosets *[repeats the picture for K and G-K and how it is the same for left and right cosets. He also illustrates the picture for [G:K]=3 and how the left and right coset partitions of G can be quite different]*.

C27: And if K is a normal subgroup?

T21: Then left and right are the same and so is the picture of them. Are you happy with that?

C28: Yeah...

Extract 9.8

Context: *The tutor and students Cathy and Cliff have been discussing Cathy's proofs in Analysis and Algebra as well as some alternatives suggested by the tutor. Cliff had problems in GRF8.5 so the tutor asks Cathy to present her proof on the b/b.*

The Episode:

Cathy, slightly reluctantly, goes to the b/b. Before starting she warns the tutor: 'Oh, OK but you may not like it!'. Does it matter?' *replies the tutor. Cathy presents her solution (abbreviated on the left part of the following table). The tutor says he accepts her argument and stresses how commutativity of x and y has made her proof possible. He then makes an alternative suggestion (abbreviated on the right part of the following table).*

Cathy's Way	*The Tutor's Way*
Let $X=o(x)$, $Y=o(y)$ *and* $t=o(xy)$. *By commutativity she shows that* $(xy)^{XY}=e$. *Then:* $XY=mt$ *for an integer m. Hence:* $XY\geq t$. *By commutativity and* $hcf(X,Y)=1$ *she shows that* $t=nXY$ *for an integer n. Hence* $t\geq XY$. *Therefore:* $t=XY$.	*He recalls that, in case of commutativity and* $hcf(o(x),o(y))=1$, $o(xy)=o(x)o(y)/\lvert\langle x\rangle\cap\langle y\rangle\rvert$ *implies* $o(xy)=o(x)o(y)$ *because then* $\langle x\rangle\cap\langle y\rangle=\{e\}$.