# INTRODUCTION

*The art of teaching is the art of assisting discovery. (Mark Van Doren, 1894-1972)*

Part of the art of teaching is the ability to make “professional judgements” about students’ learning, and to act on those judgements. However, the bases of professional judgement are often difficult to identify, let alone quantify, in order that one might make a convincing rationale for such decisions, or meet requirements of accountability. Judgements made in the course of classroom teaching are often influenced, not just by what students say, but ways in which they convey the content. There is much written about the pragmatics of language in everyday discourse – how voice, gesture, hesitations and register convey the speakers’ certainty or confidence of understanding. These rules of engagement in discourse also act in classroom exchanges, and send subtle messages to teachers, and student peers, about the understanding of subject matter.

These messages might often be implicit. Consequently, justifying “professional judgement” on matters of pedagogy becomes difficult. It is the aim of this research to make explicit aspects of student utterances that contribute to teacher judgements and to align those utterances with a model of cognitive development.

A model associating language use with conceptual development requires objective measurement of student understanding together with analysis of student talk. Because a broad-brush approach that analysed students’ mathematical understanding in general is simply not feasible, the focus of this study concentrated on students’ understanding of aspects of introductory algebra. Students were surveyed with an algebra test, and this was followed up by interviews with a sample of the students.

Survey data provided information about algebra concepts, and conceptual development demonstrated by the students, through Rasch modelling of the responses and an analysis of errors. Linguistic features were identified from interview data and analysed with respect to groups of student-ability and sets of items identified from the Rasch model.

Much of the literature on language and mathematics describes research into aspects of classroom discourse fostering greater mathematical understanding through the social construction of meaning. This and other research perspectives on writing in mathematics, syntactical understanding, and the view that mathematics might best be taught through language arts strategies, is outlined in Chapter 1.

Chapter 2 considers further research that has directly informed this particular study. This includes broadly-based research on the pragmatics of mathematical language, research using primary students' discussion of their arithmetic understanding, and studies that demonstrate an association between the language of tertiary entrance students and their mathematical success.

The argument for a language-conceptual model, in the context of introductory algebra, draws together evidence from two sources: that of students' algebra content knowledge; and, students use of language in their explication of that knowledge. Quantitative data is obtained from an objective algebra test (survey) the responses to which were Rasch-modelled. Qualitative data from interviews with students is used to interpret patterns in the Rasch model, and analysed for linguistic features that might indicate quality of understanding. Methodological approaches are described and discussed in Chapter 3.

Four subsequent chapters document results and analyses. Chapter 4 describes results from Rasch-modelling of the survey data. The model can be used to demonstrate the validity of the data collection instrument and provide the bases for subsequent analyses. Patterns of item difficulty and student ability have been used to frame the analyses of the algebra survey, and the linguistic features of the interviews.

Chapter 5 interprets the results of the algebra survey, discussing the clustering of items, and the ability grouping of students from the Rasch model. Errors made by participants on the survey are identified, and used to interpret patterns in the Rasch model. These patterns suggest conceptual links that are identified through accounts of students about their thinking when attempting to answer the items.

Chapters 6 and 7 describe and interpret results from the pragmatic analysis of the language used by students when talking about their mathematics. Chapter 6 discusses analyses of particular linguistic features identified in students' conversations.

Linguistic features studied are: the wordiness of statements (verbosity); the use of personal pronouns; the types of responses; and, the modality, and tense of the discourse. Analyses of language used are empirically based – counts of the linguistic features are used, and significance of differences established. Verbatim accounts from students are used as substantiating evidence for conclusions drawn.

The two strands of the argument, the algebra and the language, are drawn together in Chapter 7. Five contexts illustrating the association of students' language use with the quality of their understanding of algebraic concepts are described in detail. These contexts are: the role of brackets; notions of *cancelling*; concepts of *like* and *unlike terms*; the conjoining of terms; and, the meanings embodied in generalisations of students of different abilities.

Patterns of linguistic behaviour and algebraic understanding exemplified in the previous chapter are generalised in Chapter 8. A model explicitly linking language use and conceptual development in the context of introductory algebra is described.

The final chapter firstly outlines constraints on the study, and identifies limitations of the evidence together with possible effects on conclusions drawn. The results of the study are then summarised. Lastly, findings from the study, and constraint, are discussed in terms of implications for pedagogical practice and further research.

# CHAPTER 1: REVIEW OF THE LITERATURE

How does language used influence mathematical thought? What might be revealed about mathematical thought by the language used? The connections between mathematics and language are obvious – mathematics cannot proceed without its being communicated – but also subtle. Research in the field of language and mathematics, in the connections between the two and in the consequent educational implications focuses on aspects such as: language *and* mathematics; language *in* mathematics; language *for* mathematics, and language *of* mathematics. There are also the philosophical views of mathematics *as* language that influence the way the relationships between mathematics and language are seen.

The following sections discuss each of these aspects. The first section summarises the various views of mathematics *as* a language. Subsequent sections consider cognitive connections between language *and* mathematics; ways in which language *in* mathematics appears as classroom discourse; and, how natural language becomes language *for* mathematics. The final section discusses the language *of* mathematics as the mathematical register.

## MATHEMATICS AS LANGUAGE

The phrase *mathematics as language* may be considered literally or metaphorically. The literal view of mathematics as a language sees the system of signs and symbols, and associated syntax, essential for communicating the concepts, as being in important ways distinct from those of other languages. The metaphorical view of mathematics as a language attempts to describe the operation of its semiotics and syntax in terms of the natural languages of the cultures in which mathematics is conducted. The two perspectives are not mutually exclusive, but together provide a stereoscopic view of how mathematics might be communicated. Both the literal and metaphorical aspects of mathematics as language are discussed in this section.

## Mathematics as language: literally

Mathematical ideas are typically regarded as being communicated through a system of written symbols, some of which have been taken from the Latin alphabet (such as the letters used in algebraic notation), others from the Greek (particularly in geometry) and others purposefully minted (such as the signs for the arithmetic operations and the equal sign) (Pimm, 1987, pp.138-148). Written forms of mathematical ideas often bear a superficial syntactical similarity to the English verbal equivalent of the mathematical idea. Consequently, mathematics may be seen as not only having its own language, but as being a language itself. Herein lie many of the difficulties students have in developing a deep understanding of the mathematics encapsulated in the particular code that typifies the *register*, “a set of meanings that is appropriate to a particular function of language together with the words and structures which express those meanings” (Halliday, 1985).

The types of difficulties that arise when establishing a mathematical register make Bechervais’ (1992) analogy between the learning of mathematics and the learning of a foreign language compelling. In particular, he distinguished between students having a *communicative competence* (also Pimm, 1994b) in a language (or mathematics) and a *linguistic competence*. The former means that one can use the language, in a variety of forms, to communicate with a variety of audiences. Those with linguistic competence know about the structure of the language, about how its components are organised, but may be unable to use this knowledge to communicate effectively.

Bechervaise gave an example of the expression $f^n(x) = \dfrac{(x + 3^n) - 1}{3^n} = 1 + \dfrac{(x - 1)}{3^n}$ (p.3) which students may be able to “read” or “spell” but for whom it had no meaning, much as someone might be able to sound out a passage of French or German, but have no idea of what it means. The difficulties of interpreting the deep meanings of written mathematics are exacerbated by examples of what Schweiger (1994) termed *language free* communication. He gave the example of $\dfrac{d\sin x}{dx} = \cos x, \int_{\delta A} \omega = \int_{\delta A} d\omega$ (p.298) which conveys clear and distinct meaning to those with the expertise and experience. To use words to describe the mathematics would result in unnecessary complexity and opacity. Such examples serve to demonstrate that even the communicational aspects of mathematics extend beyond a language repertoire to rely solely on representational forms.

Austin and Howson (1979) disagreed with the literal view of mathematics as a language, describing mathematics as an action, a process, communicated through language. They argued that because mathematical activity can be conducted in any language, it cannot be a language. English as a language cannot be conducted in French or Chinese. It is a logical nonsense to suggest this. However, mathematical activity can be conducted in French or Chinese as well as it can be conducted in English. But is exclusivity a hallmark of different languages? Languages adopt from others new words to explain new experiences, such as the notorious *le weekend*, abhorred by French traditionalists, just as mathematics has done.

Mathematical languages often privilege words from the lexicon of the natural language in which mathematics is conducted. Mathematics adds to this lexicon words coined to describe certain concepts for which the natural language has no suitable equivalents (Ellerton & Clements, 1991). The dominant present-day mathematics has developed in a Eurasian culture, so that the introduction of these mathematical ideas and the linking of these to the mathematical concepts of cultures originating outside that region have presented problems which are political and cultural as well as language-based (Austin & Howson, 1979; Charbonneau & John-Steiner, 1988; Clarkson & Thomas, 1993; Dawe, 1995). In this case, mathematics might be considered a new, and foreign language.

### Mathematics as language: metaphorically

Pimm (1987) and Halliday (1985) used *mathematics as a language* as a metaphor that served to draw attention to the characteristics of the mathematics register and the power of natural language as used in that register. However, as Halliday (1976) and Kress and Hodge (1979) noted, registers distinguish degrees of enculturation when used as *territorial* or *status markers* rather than as accurate ways of articulating ideas (Pimm, 1987, p.109). The struggle to make clear complex ideas often results in students (and teachers) using illustrative metaphors, some of which have become part of the mathematical lexicon. Metaphor plays an important part in conveying mathematical concepts: such as *spherical triangle* to describe a closed three-sided figure drawn on a sphere; the *face* of a solid; or, physical occurrences *obeying laws*. Problems arise when novices, in particular, take the metaphor literally as in *right-angled* triangles and by false analogy, *left-angled* triangles (Pimm, 1987, p.87).

Consequently, initiating students into the world of mathematical ideas requires that they gain linguistic competence and communicative competence (Bechervaise, 1992; Pimm, 1994a) through the acquisition of the vocabulary, syntax, semiotics and semantics of the mathematical register in order that they not be "speechless mathematically" (Arzarello, 1998, p.250). The *metalinguistic understanding* (an awareness of the structures and relationships of a language) by students of their natural language, as well as the standard forms of (in this case) English and of mathematical language, has been shown to be important in the development of students' ability to deal with complex ideas and make generalisations, and in the understanding of algebra in particular (MacGregor, 1993).

Usiskin (1990) united both literal and metaphoric perspectives when he wrote that "mathematics *is* like a language because it *is* a language" (p.232). Mathematical language is a sub-set of the set of languages. Hence it shares characteristics of the natural languages, and acquires some features that are unique.

The ideas of mathematics and how they are understood by individuals are conveyed through language. This is the case whether mathematics is an activity conducted through a language that exists as an extension of a natural language, whether mathematics is considered as a language in a metaphorical sense, or as a foreign, or natural, language for the purposes of pedagogical discussion,. "Language is so central to the whole of the educational process that its role was never even talked about, since no-one could conceive of education without it." (Halliday, 1985, p.96)

## LANGUAGE AND MATHEMATICS: COGNITIVE CONNECTIONS

Although low-order concepts may be formed and used without language, language is essential for the formulation and communication of higher-order concepts (Austin & Howson, 1979, p.167). So much so that speech is more than a "window on the mind". It is inseparable from thinking, with neither speech nor the thoughts generating that speech being "prior to the other" (Sfard, 2001, p.29). Pirie, however, considered speech, and language in general, to act as a mediator for thought, rather than thought itself and warned that "the only knowledge we can have of the pupils' understanding is gained from *our* interpretation of *their* communication to us through symbol or word" [Pirie's italics] (Pirie, 1998, p.10). This said, language is the only way we have of gaining insight into students' thinking, and possibly the only way students have of gaining insight into mathematical (and other) ideas (Austin & Howson, 1979, p.167).

The link between language development and cognitive development was made by Piaget (in Austin & Howson, 1979) who suggested that the growth of a child's language succeeded the development of "concrete operational" thought. Freudenthal (1973) disagreed with this interpretation of the connection between language and understanding criticising the fact that Piaget "almost never tests whether they [his subjects] understood the language in which the task was formulated" (p.120). Development of understanding relies on the acquisition of adequate and appropriate terminology is necessary for understanding to develop (Bruner and Vygotsky in Austin & Howson, 1979). The Vygotskian perspective adopted by Laborde (1990), later also by Sfard, is that both mathematical content and the associated linguistic aspects need to be learnt together, that "cognitive and linguistic aspects intervene simultaneously in the comprehension and in the use of different means of expression" (Sfard, 2001, p.65).

This section discusses research into the connections between language and mathematics from developmental perspectives of cognitive growth. The role of the teacher's language in developing students' mathematical understanding is discussed in the first sub-section. A second sub-section considers how students' ability to generalise is reflected in their use of language. Third and fourth sub-sections discuss recent research into students' spoken and written language, and what it can reveal of their mathematical development.

## The role of the teacher

The acquisition of mathematical ideas and their associated language occurs principally through classroom discourse. The linguistic features of the teacher's language can influence the students' understanding (Sfard, 2001, p.65), as can the structure of the student's natural language (Galligan, 1995 in Ellerton & Clements, 1996; MacGregor, 2002). Discourse encourages students to verbalise their ideas. Such verbalising is "developmental and continuing" and provides situations where a teacher can "discover the level and development of student understanding" (Mousley & Marks, 1991). This implies that teachers have the necessary language skills to model the appropriate forms of language and to become aware of them in the discourse of their students.

MacGregor, in a study of the language used by students in the context of comparing quantities, found that pre-service teachers needed to "acquire proficiency in the language necessary for explaining mathematical ideas" (2002, p.4). This is true for all aspects of teaching mathematics (or any other discipline). In other words, teachers need to have developed an appropriate mathematical, and classroom, register.

## Language use and generalisation

The continuous development of students' verbalising of mathematical understanding need not be as consistent as might be inferred from the hierarchical organisation of curriculum documents. Students who can verbalise mathematical understanding of one particular aspect may not be able to do so for another. Closest to recognising the to-and-fro nature of students' development of understanding are the ideas articulated in the *Principles and Standards for School Mathematics* (NCTM, 2000). It was acknowledged here that whenever students encounter a new idea they will first articulate their understandings, either verbally or in writing in informal ways. Only once the concepts have been understood could students express them in more formal ways that have been traditionally recognised as mathematical (p.63).

Generalisation and deductive reasoning are at the heart of mathematics (Mason, 1996, p.74). Deductive reasoning is characterised by the use of logical connectives such as *so*, *because*, *if … then*. Somewhere between the ages of five and ten years, children develop the ability to make logical causal inferences from the temporal order of events, instead of reasoning that because one event follows another in time it must be caused by that preceding event. Children develop the ability to deal with causal constructions (for example, the use of the word *so*) or respond appropriately to and use linguistic clues that signal that an “explanation is in the deductive mode” (Donaldson, 1986, p.146). As a consequence, Donaldson suggested that the single-word response *'Cos* (because) to questions of *Why?* reflected a lack of knowledge rather than an inability to provide an explanation of concepts intuitively apprehended.

Deductive reasoning proceeds from an apprehension of properties of a class to that which is perceived to be a particular instance of that class. It is the ability to make generalisations from particular instances, to see an overall pattern in repeated similar instances that characterises mathematical behaviour. This type of cognitive development occurs together with linguistic development (Laborde, 1990). Generalisation, abstraction and metaphor are vital to the understanding of complex thought in any field (Dawe, 1995) and are mastered at different times in a child’s language development.

There appear to be few linguistic or cognitive problems for children moving from specific instances to generalising because words imply classes of things or ideas. Children can make abstractions from concrete experiences before the age of five years. However, the use of metaphor begins at eight or nine years of age (Halliday, 1985). This has implications for the learning of mathematics. Pimm (1987) pointed out that many mathematical concepts are couched in metaphoric terms, and consequently, very young students might still interpret the metaphor literally. The student and the teacher might be using the same words but have different concepts attached to those words.

Being able to think of a particular example as “representative of a more general justification of a class of examples” is the outcome of a “cognitive act” (Rowland, 2000, p.40) expressed through language. The quality of the language used is a pointer to the quality of thought.

## Spoken language

Generalisation and abstraction mark a certain detachment of the speaker from the subject. This detachment is characterised by certain linguistic structures such as the person, tense and modality of the utterances. It is these features that underpin the “pragmatic” judgements teachers make of a student’s understanding or competence.

*Pragmatics* is the area of linguistics which considers the way in which language is used. Rowland (2000) examined ways in which students use oral language to indicate a “degree of commitment to a proposition” (p.46). He identified the use of pronouns, hedges and modal forms as indicative of the extent of “cognitive vulnerability” (p.210).

The work of Rowland was drawn on by Bills and Grey (2001) and Bills (2002) in their studies of the explanations of mental computation process given by students in primary school. Students whose explanations used the general *you* rather than *I*, who used particular examples to refer to a class of arithmetic objects, or who quoted a rule, and whose explanations were given in the simple present tense, tended to be more successful than students whose explanations were in the form of a personal recount of action on a particular item.

Spoken language has only recently been recognised as being important in a child’s education, but more as a skill to be acquired than as a vehicle of learning (Halliday, 1985, p.96). However, “the incoherence of discourse is often a pointer to an inability to organise the relevant meanings in relation to each other” (Halliday & Hassan, 1985, p.96). Incoherent writing may also be an indicator of poor cognitive constructs (MacGregor, 2002; Sibley, 1999; Southwell, 1993).

### Written language

Studies outlined in this sub-section suggest that: the coherence of students’ written explanations of their understanding of mathematical ideas can be associated with their success in mathematics; and that students’ awareness of appropriate language structures in their natural language has been found to assist in the development of their mathematical understanding. The ability to make generalisations about mathematical ideas would appear to be associated with the development of the *register*, the use of terminology, syntax and grammar, appropriate to the context, in which to express those understandings.

The quality of students' explanations of their mathematical understanding has been found to be associated with quality of understanding of the mathematics. Boero, Douek, and Ferrari (2002) reported that the quality of the written explanations for an algebraic procedure given by first-year university science students studying mathematics was associated with their algebraic proficiency and subsequent retention in the course. They concluded that poor language skills and an "inability to provide a coherent mathematical explanation" (p.248) were a sufficient condition for failure. Thus, a "register awareness" was also necessary for students to be mathematically successful, as well as a "metalinguistic" awareness (MacGregor & Stacey, 1994).

MacGregor's research focused on the associations between the syntax of language used in and about mathematics, the understanding by students of the structure of their natural language and their success in mathematics, particularly algebra. Mathematics achievement and proficiency in the standard natural language were related (MacGregor & Price, 1999). In a 2002 paper reporting on a study of the "association between articulation and cognition in the context of comparing quantities", MacGregor suggested that "grammatical form reflects conceptual structure and can reveal that a concept is vague, under-developed, unstable or incorrect" (p.2).

The importance of facility with the mathematical register becomes more marked as students are required to use formal *proof* structures. *Proof* is both a process and a product. The process of proving mathematical conjectures requires that one use argumentative processes, whilst the product of proof in all its conventional formality requires "sophisticated metalinguistic awareness and linguistic competence" (Boero et al., 2002, p.255). The formal conventions of mathematical proof become barriers to students struggling to express their understanding in unfamiliar ways. Students may also have difficulty in expressing their understanding in formal ways, not because they do not have the concept, but because they lack the notation (Laborde, 1984, reported in Jacobsen & Dawe, 1984; Rowland, 2000).

Developing students' facility with the notation of mathematics would seem to have been a predominant focus of mathematics teaching (e.g., Board of Studies NSW, 2002, p.19). In mathematical contexts, particularly that of algebra, the symbols "are always uppermost as the focus of attention", and often that is the point where students remain. They manipulate the symbols according to rules, but without understanding (Frid, 1993; Pimm, 1987). Pimm drew an analogy between students learning to write their natural language where they focused on the letters and words only, and so lost the meaning of the whole sentence. Understanding requires that students connect the mathematical symbols with their own mental constructs. The enduring nature of early constructs may be a barrier to further mathematical growth (Laborde, 1990). This lack of “languaging” (Sierpenska & Lerman, 1996) to develop understanding can result in a disjunction between the symbol system of mathematics and the conceptual system (Austin & Howson, 1979).

If generalisation and deductive reasoning are at the heart of mathematics (Mason, 1996, p.74), then an increasing ability to abstract concepts and to make generalised inferences from particular instances indicates levels of cognitive growth in mathematics. The research discussed in this section has identified aspects of the inextricable association of language development with the development of mathematical understanding. It has described the importance of a developing mathematical register, from natural language to the conventionalised terminology of the register. The register, it is noted, uses metaphor to describe complex ideas, and novices encounter difficulties moving from a literal, or natural language, interpretation of terminology. Vocabulary, syntax, and semantics of the ways in which students articulate their understandings, either spoken or written, have been found to indicate levels of competence and understanding.

## LANGUAGE IN MATHEMATICS: DISCOURSE IN THE CLASSROOM

Mathematical understanding is communicated to others through writing, speaking and representations such as tables, graphs, diagrams, symbols or combinations of these. These aspects comprise the formal discourse of a mathematics classroom. The partners in the discourse are the teacher and the students. Discursive interactions are those of teacher and student(s), and students with other students.

The following section outlines how research has found that spoken discourse can contribute to the development of students’ mathematical understanding. Firstly, interactions in the discursive classroom are described. Secondly, ways in which research suggests that mathematical discourse can be developed are examined.

## Interactions in the discursive classroom

Discourse is the act of communicating thought through speech (Shorter Oxford English Dictionary, 1975), but it is more than just the language (Mousley & Marks, 1991, p.9). Meaning is conveyed in speech not only through words, but also through the intonation, volume and rhythm of the sounds, and accompanying gestures and facial expressions, the pauses and the paraphrases (Halliday, 1985, p.30), as well as the context in which the discourse takes place and the "interlocutors' histories" (Sfard, 2001). Speech differs from written communication in important ways, and serves different purposes. Speech is anchored in the here and now, writing is not. It "has different functions, different contexts and consequently it 'means' in different ways" (Halliday, 1985, p.32). Spoken language contains more grammatical items. These serve to indicate and reinforce the subject about which the conversation revolves (particularly the little words).

Written language contains more lexical items – items whose meaning is important. Consequently, the written form of language is meaning dense, while the spoken language of discourse is more sparse (Halliday, 1985, pp.61- 62). Meaning has to be inferred from the total verbal act.

There are different types of discourse, and different types of mathematical discourse – the everyday, the school discourse and the discourse engaged in by professional mathematicians. At times, this last type of discourse is different in nature and focus to the discourse of mathematics educators (Sfard, 2002a).

In a classroom, teacher and students interact, as do students with each other. These interactions help the participants to develop a set of jointly understood meanings for the mathematics being undertaken. To do this, parties have both to listen and to talk. The relationships in a classroom influence, and are influenced by, the roles of speaker and listener ascribed to the members of the classroom by the culture of the class. Two examples of discursive relationships, the role of teacher as listener and the purpose of student talk are discussed in the following.

### ***Teacher as listener***

Not only is the meaning attached to concepts conveyed through talk, but so too are attitudes towards the propositions made (Rowland, 2000) or the subject itself (Bills, 1999). Classroom discourse reflects social interactions, power relationships, and the province of control over knowledge (Mousley & Marks, 1991), as well as the quality of understanding that students have of the mathematics. Studies of classroom interactions between teacher and students, and between students, consider the negotiation of meaning that occurs from a social perspective and from a cognitive perspective.

Cobb, Wood, and Yackel (1995) take a constructivist approach to the teaching of mathematics, realising that approach through problem-solving activities to produce substantive learning. Substantive learning results in “cognitive restructuring” (p.232). Their study of classroom discourse focused on the social norms of the classroom and how they constrained “what is problematic for the students and what might count as an acceptable solution” (p.236). A similar study (though having a different focus) was conducted by Sfard (2001). She studied interactions between two students as they negotiated their way through a problem, and also interactions between a student and her teacher. The focus in this case was on the cognitive relationships between the interlocutors and the effect on their communication of ideas, and the ultimate effect on the substantive learning by the students.

If “communication …is almost tantamount to thinking itself” (Sfard, 2001, p.13), then discursive practices in the classroom must serve to foster and to reveal the mathematical thinking of students. Fostering mathematical thinking involves students acquiring the terminology and the language patterns necessary both to formulate mathematical concepts and to communicate these with others. There is a second reason for talking – to communicate with oneself. In a classroom situation, both these purposes of talk may provide insight for the teacher who listens (Pimm, 1987, p.23).

Notwithstanding the fact that "*spoken words* do not equal *thoughts in the mind"* (Zack & Graves, 2002, p.265), all that teachers can do is interpret the student's utterances within their own frames of reference (Pirie, 1998). Listening to students as they negotiate meaning with each other, with the teacher, and with themselves, "invites [one] to eavesdrop on their thinking" (Zack & Graves, 2002, p.241). By listening to students, teachers gain insight into what students might understand and how well they understand.

Learning mathematics involves more than developing proficiency in manipulating symbols and following standard procedures in order to obtain correct answers. This can be done without an understanding of the meaning of the mathematics (Frid, 1993). Learning mathematics involves students establishing many deep and rich connections between ideas.

### ***Students talking***

Talking about mathematics enables students to verbalise their ideas, and to expose them to a critical audience of their peers and teachers. Verbalising is necessary "for abstracting, generalising and categorising" (Dawe, 1995, p.233), what it means to do mathematics. Speech is immediate and exists in the present, presenting a dynamic view of the world where phenomena happen (Halliday, 1985, p.97). By talking mathematics, students demonstrate their becoming knowledgeable in the discipline, part of which involves "learning to talk in the manner of the practice" (Adler, 1999, p.1).

The relationships that exist between everyday language, school language and mathematical language have been variously modelled by, for example, Quesada (1996) and Ellerton and Clarkson (1996). The models situated mathematics in a social and cultural context the interrelationships of which were negotiated through and by language. By focusing students' attention on the language they use as they verbalise ideas, and by modelling the appropriate register, teachers develop the metalinguistic and register awarenesses that are essential for the development of mathematical understanding (Adler, 1999; Boero et al., 2002; MacGregor & Price, 1999; Reeves, 1990) and for the students' participation in the culture of mathematics (Sierpenska & Lerman, 1996; Steinbring, 1992).

Induction into a culture requires that there be “experts” who are already part of the culture who can introduce students to the ways in which that culture views the world and thinks about that world. That world, in this instance, is the world of mathematics. Classroom discourse provides the social setting for students to develop their mathematical ideas, to think creatively and create an “awareness of thought” (Reeves, 1990, p.97), and to learn to use the words that communicate their ideas about mathematics in precise and clear ways.

## Developing mathematical discourse

Several studies (e.g., Austin & Howson, 1979; Bickmore-Brand, 1993; Pimm, 1987; Reeves, 2000) undertaken have examined how students acquire the language skills needed to communicate mathematical ideas clearly. These examine the strategies used by teachers, the purposes and potential problems.

Cobb, Wood, and Yackel (1995) studied “how the evolving network of obligations and expectations” influenced students’ growth of mathematical understanding. They concluded that in order to have students recognise their mistakes and refine their ideas, a sense of trust has to be nurtured. The talk in a mathematics classroom becomes therefore a conversation, not a test (Reeves, 1990). The effectiveness of classroom talk depends on the purposes of the teacher in encouraging it, and on the management of the talk through various strategies (Austin & Howson, 1979).

Bickmore-Brand (1993), adopting a literal interpretation of mathematics as a language, discussed how adaptation of language arts strategies could contribute to students’ understanding and to their induction into the culture of mathematics. Bickmore-Brand proposed strategies such as *immersion* in a mathematical environment where mathematical activities are carried out in appropriately mathematical ways in contexts which link students’ experiences and thinking to new mathematical contexts. As in language classes, the use of scaffolding techniques such as probing questions to elicit explanations, joint constructions of reports or explanations and explicit teaching of the mathematical register to help students develop mathematical ways of inquiry and communication were suggested.

The particular case of mathematics being taught using language arts techniques to help students whose first language is not English was also discussed by Bell (1993). Such artfully structured discourse guided students into "talking themselves into understanding" (Pimm, 1991). A "socio-psycho-linguistic" framework can map how students move from the use of real world language to the language of the classroom, to mathematical language, to the construction of meaning in mathematics (Gawned, 1990, p.40).

Talking *about* mathematics, or talking *of* mathematics, needs to be "an integral part of the daily program" (Sibley, 1999) and leads into students writing about their mathematics. Although mathematics has been perceived as being conducted primarily in the written form, the words and symbols and syntax of that writing need to be connected to students' own constructed meaning if they are to manipulate the condensed symbolic forms with any meaning (Frid, 1993). This learning may take place more effectively through the spoken form, and requires of the teacher particular roles and expertise.

To encourage discourse that is productive, teachers use strategies such as having students work in small groups, using "gambits" (Pimm, 1987) to structure and control the talk in the classroom, modelling the appropriate language and providing some direction for students to adopt the model of mathematical language (Bickmore-Brand & Gawned, 1990). Group work does allow students to talk about the mathematics and the mathematical task in hand, but not all student talk may be productive of better mathematical insight or understanding (Gooding & Stacey, 1993; Laborde, 1990; Pimm, 1991). The talk has to "create an awareness of thought" (Reeves, 1990, p.98), and be "focused, explicit, disembodied and message oriented" (Pimm, 1991, p.45) rather than be a means of social negotiation, such as the exchanges between two students, Ari and Gur, described in Sfard (2002b). Video studies of students working together revealed to Zack and Graves (2000) the "intricacies and depth of their reasoning"; the video record enabled the researchers "to hold it [the students' reasoning] in place and think about what their talk means" (p.265).

The social structure of a classroom is asymmetrical – the teacher has the expertise and the control. Consequently, the talk in a classroom is asymmetrical. Typically, the teacher does most of the talking and controls the talk of students. The questions teachers ask direct the verbalisation by the students. Love and Mason (1995) discussed the effects on classroom discourse as teachers tell students information or, alternatively, ask students for information. How that information was obtained was either by means of questions that directed (*funnelled)* students to the teacher-required answers or those that focused the students on the subject of the discussion. The latter type of questions left students free to provide answers, which might not have been those that the teacher expected, but which might have been more revealing of the students' understanding.

The nature of the discourse may, according to Reeves (1990), "blinker or expand a child's thinking" (p.94). Teachers therefore have to think about the language they use (Jacobsen & Dawe, 1984, p.267), to be teachers of language as well as teachers of mathematics. They need to be aware of language structures, and model and explicitly teach those structures, particularly the mathematical vocabulary and syntax that typify the mathematics register.

Adler (1999) described a study of how students in a South African multilingual classroom were helped to refine their language skills and hence their understanding of the mathematics they were learning. In order for students to develop an educated discourse, Adler suggested that teachers needed to attend to what students said as well as how they said it. In the instance described in Adler's study, the teacher's attention was on the pronunciation by the students, their use of verbalisation as a tool for thinking (communicating for themselves), for demonstrating understanding (communicating for others), and to help the teacher understand the pupils' thinking. The teacher's attention was divided between noting the form of the speaking as well as the content.

Although, initially, this might have been problematic for the teacher, Rowland (2000) claimed that, with practice, one might listen to the language as well as listen for the mathematical content. For students to become fluent in the language of mathematics, they need to move between everyday language use and technical use of the language, and develop a *robustness* (Austin & Howson, 1979) to cope with the different usages of the language.

The research discussed in this section focuses on the importance of establishing a discursive mathematics classroom, ways in which this might be done, and some of the difficulties that could arise. Often, mathematics is seen as providing little to discuss. Established procedures are to be followed and answers marked as either correct or incorrect. Against this traditional view is the idea that there is much to be spoken about in mathematics, and that discussion is valuable in helping students to learn mathematics.

Such a classroom has students engage in purposive talk, about the mathematics in hand. To achieve this, classroom discourse needs to be carefully orchestrated (but not always controlled) if it is to achieve the desired end of the development of mathematical understanding. Teachers need to be aware of the social and cognitive relationships that exist in the classroom, their own use of language and the use of language by students.

By establishing an environment where students feel free to expose their thinking through conversations with themselves or others, teachers and peers, students develop mathematical understandings that are meaningful, and teachers gain insight into the thinking of their students. Talk allows students to acquire the form and structure of the mathematical register, and the ability to move from their natural language to the privileged use of the language in the discipline.

## LANGUAGE FOR MATHEMATICS: FROM NATURAL LANGUAGE TO MATHEMATICAL LANGUAGE

Reeves (1990) noted that the use of oral language has been widely advocated since, for example, the Cockcroft Report (1982, p.93). Although simple concepts may not require language (and possibly also some higher mathematical concepts as suggested by Schweiger, 1994), more complex ideas are mediated by and manipulated through language (Austin & Howson, 1979, p.243; Boero et al., 2002). Learning mathematics proceeds from the student carrying out an activity in which the mathematical ideas are considered (thought about), to the student telling his or her thoughts. The activity needs student-talk and teacher-talk to “turn the doing into thinking mathematics”. The quality of the talk determines the quality of the learning

(Reeves, 1990, p.94), and this is determined by the learner's mastery of the natural language as well as the distance of the learner's language from that of the teacher and from that of mathematical language (Austin & Howson, 1979; MacGregor, 1993).

This section examines research that focuses on the development of students' mathematical language as it evolves from the informal, natural language to the formal, conventional language of the mathematics register. Firstly, some of the difficulties that research has found students encounter are identified. Secondly, models that describe how students' language assumes a more formal structure are examined, with a focus on spoken language. Thirdly, research that considers the development of students' written language and the move to symbolic language, particularly that of algebra, is discussed.

### Difficulties students encounter

The difficulties learners encounter when moving from the everyday register and conceptualising of natural language to that of a mathematical register reflect their development of understanding of both the mathematical concepts and the ways of expressing that understanding. Austin and Howson (1979), Laborde (1990) and Reeves (1990) have made the point that the two processes are interdependent. This implies that the language of mathematics needs to be taught explicitly with the content, and linked with the concepts.

Difficulties experienced by students in the learning of mathematics have been attributed in part to a formalist teaching of mathematics when the "connection between the natural and intuitive reasoning and formal calculations is not made" (Vergnaud, 1997, p.25). Austin and Howson (1979) warned that the premature introduction of formal language might be "mathematically unproductive" (p.179), particularly when vocabulary is taught without the students having "knowledge of the concept or entity described" (p.180).

Frid (1993), in her study of high school students, also found that when the focus is on use of the correct register, and particularly on the associated mathematical vocabulary, students may have learned "to replicate the cultural conventions of formal mathematical language, but they [did] not appear to have developed any intellectual autonomy in using mathematical language to interpret and give meaning to their mathematical activity" (p.29). The result was that students came to "devalue" their own ways of making sense: "I'm not perfect in interpreting it in the correct mathematical language" (p.39).

The teaching of mathematical registers without reference to their function, and without making explicit the reasons for particular constructions and styles, results in students dealing with mathematical language without taking care of the meanings. In cases such as those examined by Frid, this lack of meaning resulted in students carrying out procedures without seeing the need to attend to the logic or the meaning of the results. The focus became merely one of students listening to, or reading, instructions in order to find out what to do and then providing an answer (MacGregor, 1990, p.103).

## Models of language and cognitive development

Research has identified that language development is closely associated with mathematical understanding. Some models that explain the developmental shifts are described below.

Gawned's (1990) model of development of mathematical language consisted of four areas that contributed in some degree to a mathematical activity and to the building of the associated concept. The basis of the model was the real world language of the child that became the acceptable *classroom language,* which then privileged specific *domains* for the language of mathematics. This specialised domain then aided in the final step: construction of meaning in mathematics. Encouraging learners' use of everyday language to discuss mathematical concepts required them to master complex language skills (MacGregor, 1993) and to have a knowledge of language in general (Laborde, 1990, p.55). It also helped learners to value their own way of making sense (Frid, 1993).

Although students need to be encouraged to use their own language, teachers need to model and encourage the use of appropriate vocabulary as well as the associated linguistic structures needed to record information (Pengelly, 1990; Waywood, 1993), and make mathematical explanations, conjectures and justifications (Malone & Miller, 1993). MacGregor (1990) commented that having students to write explanations was not a familiar mathematical activity. Explanatory writing requires students to engage in a particular style of writing that is different to the recount genre most often used (Mousley & Marks, 1991). This genre does not support the development of abstraction or generalisation in the way that genres of explanation and reportage may do. Where students can use their own language to verbalise their ideas, and where they can "talk themselves into thinking" (Mousley & Marks, p.77), they can build the mathematical language necessary for clear conceptualisation of the mathematical ideas.

Vergnaud (1997, p.14) described how a learner's conceptual progress can be demonstrated through different linguistic aspects of either spoken or written explanations. As students' understanding develops, they moved from "adjectival" use of a mathematical term to a "substantive" use, from using a single proposition to "several-termed related propositions", from the explanation in a singular instance to using the explanation of a particular instance to exemplify the behaviour of a class of mathematical objects. Higher-order thinking was accompanied by an increase in linguistically complex statements, that were coherent and well-formed (Bills, 2002; Bills & Gray, 2001; Boero et al., 2002; MacGregor, 2002).

Teachers need to use the appropriate mathematical register as a model for their students, becoming mediators for students' thinking and understanding, and becoming "cultural mediator[s]" by providing important "voices" as models of mathematical linguistic structures (Boero et al., 2002, p.263). Merely hearing examples of appropriate mathematical language need not guarantee students will adopt the practice (Laborde, 1990; MacGregor, 2002; National Council of Teachers of Mathematics, 2000); there needs to be explicit teaching of the linguistic skills. Those students who do adopt the register of the teacher are likely to be those students who understand the mathematics (Bills, 2002). Adler (1999) and Rowland (2000) found that, for teachers, balancing the focus of what students say with the manner of the saying was difficult, but could be refined with practice.

Parzysz (1984, in Jacobsen & Dawe, 1984, p.267) gave an example of how French students were moved from using everyday language to using mathematical language. The text language was "special" French that

> present[ed] peculiarities in vocabulary and syntax and [varied] with the age of the student, being 'dynamic' in the beginning, but growing progressively 'static'

> under the influence of the teaching received, and tending to become closer to symbolic language structures. The most utilised written or verbal language [was] therefore an intermediate one, with the purpose of bridging the gap between truly natural language and symbolism with various possible levels in between.

Because symbolic language is exclusively written, teachers needed to do much "translating". It is this aspect of mathematics that exemplifies the literal perspective of mathematics as a language, adopted by, for example, Bechervaise (1992).

## Developing the written language of mathematics

Research and discussion about the mathematics-language interface has been directed at ways in which spoken and written language in natural language develops to the written language of mathematics. Halliday (1985) pointed out the particular differences between spoken and written language, the different purposes and, consequently, the different ways that spoken and written language "mean". In turn, the written language of mathematics differs from that of written English in several ways that have been identified by several authors (e.g., Austin & Howson, 1979; Bechervaise, 1992; Ferrari, 2002; Halliday, 1985; MacGregor, 1990, 1993; MacGregor & Stacey, 1994).

Mathematical language and the associated symbol system are not a shorthand form of natural language (MacGregor, 1993, p.55). Writing English requires that there be a variation in vocabulary and structure. Mathematics, on the other hand, has a limited set of linguistic patterns in order that similarities are made apparent (Austin & Howson, 1979).

Mathematical language is low in redundancy, but, because of the restricted register, often becomes highly ambiguous. Written forms of everyday language use redundant statements that help to reinforce meaning and remove ambiguity (Halliday, 1985). The syntax of mathematical statements is important and differs from that of natural English. It is not linear; rather, as MacGregor (1990) found, meaning was made through paying attention to prepositions and order of words and their relation to each other (p.110). Bechervaise (1992) found that mathematics needed to be read analytically and accurately for it to be comprehended. Therefore, competence in one's natural language would seem to be a precondition for success in reading and understanding mathematical language (Boero et al., 2002; Laborde, 1990; MacGregor, 1993; MacGregor & Stacey, 1994; Padula, Lam, & Schmidtke, 2001).

Mathematical writing is context free. Unlike written forms of everyday language, where meaning is constructed within a context, or from association with experience (Ferrari, 2002; MacGregor, 1990), written mathematical text does not depend on the context of a situation for meaning to be made. It also has a limited vocabulary where each word has a precise and an unvarying meaning. Often, it is the strictly defined uses of everyday words that make understanding of mathematical ideas difficult. Words such as similar are used in a particular way in mathematics, that is like everyday use but not synonymously so (Austin & Howson, 1979, p.181).

Pimm (1987) gave the example of the word *diagonal*, an adjective in English; in geometry, it is a noun. Nunes (1997, p.41) cited the change from the use of the *plus* and *minus* signs as verbs in arithmetic expressions indicating operations to be performed, to that of adjectives in the contexts of directed numbers. These signs, and the other arithmetic operators also act as "conjunctions" in expressions of relationship between numbers (e.g., $x + 4$). This is a particular aspect of algebra that students moving from arithmetic contexts find difficult.

The move by students from using the structures of natural language to those of algebraic language has been studied by Padula et al. (2001) and, extensively, by MacGregor and her associates (1993, 1994, 1999). MacGregor and Stacey (1994) found a correlation between the algebraic competence of students and their "metalinguistic" competence in their natural language. Some students, whose metalinguistic skills were sound however, were not so successful in algebra. MacGregor and Stacey suggested that these students might benefit from instruction in the syntactical aspects of algebraic language.

Mathematics, in the end, is carried out in the written form. Students need to move from informal talk and representations to using formal language, whether spoken or written. In the preceding section, the development of mathematical understanding through classroom conversation was considered. Writing is also valuable, particularly if the appropriate use of different genres is encouraged (Mousley & Marks, 1991).

Different genres, like different registers, provide students with different "voices" with which to express their understanding.

Writing exposes misconceptions and misunderstandings, gaps in language skills and non-standard use of the language, because writing requires the use of carefully structured language, and an understanding of the concepts, in ways that spoken language does not (MacGregor, 1990).

The models of written language students most often encounter are those of the textbook and the teacher. These are finished products, models of completeness. Pimm (1987) argued that this might not necessarily be useful as instanced by Frid (1993, p.39) who writes of a student, who confessed to being unable to express ideas in the accepted formal ways. The student was left without a voice with which to articulate his thoughts. Such an example demonstrates the challenges for teachers and textbook writers to set a fine example of a finished product while at the same time encouraging students to develop their own understanding in their own ways.

Perry (2001) advocated "messiness as an assessment tool", describing writing, and other forms of representations, as ways students have of negotiating meaning. This is writing for oneself. The "messy" artefacts are records of students' mathematical process. Writing for others requires a polished record, the product of the mathematical activity. Writing has also been found to help in the development of mathematical understanding of Aboriginal students (Howard, 1999). Writing provided an opportunity for students to express their mathematical thinking on paper, to clarify and re-order their thinking, to synthesise ideas, and to open their ideas to critiques.

In their 1996 review, Ellerton and Clements cited various approaches to the use of writing to learn mathematics. These approaches included the use of journal entries, conference writing, the use of narrative or letter writing, projects and reports. They suggested the need for further research into the value of these activities in improving student performance or understanding or attitudes to mathematics.

The research described above, recognises that the development of a mathematics register from the informality of one's natural language is a matter of negotiating meaning. Firstly, some of the difficulties that arise as students negotiate new meanings for familiar words that are adopted into the mathematical lexicon are considered. Secondly, ways in which developing language facility and mathematical understanding are linked are discussed. Thirdly, research into the conceptual development of meaning in written mathematics, and ways in which the sparse, and often ambiguous mathematical symbol system can be more explicitly linked with meaning through the use of everyday language is described.

As students' mathematical knowledge advances, so must their linguistic competence in the mathematical register. Students talk about their mathematics before they can write it, so they are able to move from using everyday language to describe mathematical events, to using the mathematical register.

## LANGUAGE OF MATHEMATICS: THE MATHEMATICS REGISTER

The language of mathematics may be spoken, or written. In its written form, natural language and conventional language structures may be used, or, as is most common, the particular symbol system of mathematics may be used. The language of mathematics is not "natural", nor is it usually considered as a language itself, except in a metaphoric sense (Pimm, 1987).

This section firstly reviews the literature on the general nature of mathematical language. Secondly, how research sees the influence of language in the transition of thinking from the arithmetic to the algebraic is discussed. Thirdly, ways in which it has been found that natural language influences and indicates the development of mathematical thought are described.

### Some general comments on mathematical language

Mathematical ideas can be expressed in any natural language, notwithstanding the limitations of the languages of some cultures when dealing with concepts in western mathematics. The language of mathematics has been described as a "creole" (Bechervaise, 1992) or as a "register" (Boero et al., 2002; Halliday, 1976; Pimm, 1987). The reasons for these descriptions stem from the fact that the particular vocabulary of mathematics consists of words specially coined, words borrowed directly from natural language or given special meanings (Boero et al., 2002; Dawe, 1995). Of particular importance in mathematical language are prepositions, grammatico-logical operators (e.g., *or*, *if…then*, *and*) and quantifiers (e.g., *less than*, *larger*, *times*).

The language of mathematics uses no sounds other than those of the natural language in which it operates, but it does use new symbols (Schweiger, 1994). Its syntax is different to that of natural languages, and this is the most important constituent of meaning in mathematical statements (Austin & Howson, 1979; Boero et al., 2002; Gawned, 1990; Laborde, 1990; MacGregor, 1990, 2002; MacGregor & Price, 1999; MacGregor & Stacey, 1994; Nunes, 1997; Pengelly, 1990; Pimm, 1987, 1991; Pirie, 1998; Reeves, 1990). The syntax of mathematical language is restricted (Austin & Howson, 1979) and is characterised by use of the passive voice and nominalisation of verb forms, together with the "removal of references to the author of the action" (Laborde, 1990). The language of mathematics has no word classes such as verbs, adverbs, and nouns as does a natural language (Schweiger, 1994).

The written forms of natural languages are more lexically dense than the spoken forms (Halliday, 1985). The written forms of mathematics are even more so, lacking almost all grammatical items and with no redundancy. The lack of redundant items and the lack of context for written mathematical language mean that, unless meaning can be derived from an understanding of the syntax, the written statement is often difficult to read and comprehend.

Research shows that students experience difficulties when translating word problems into mathematical statements when they read according to the conventions of natural language text (Boero et al., 2002). By so doing, students seek meaning through the context of the question rather than through understanding the mathematical ideas. When the mathematical ideas are understood, the information can be restructured into the necessary well-formed formulae (MacGregor & Price, 1999). The problems and difficulties students find in developing the mathematical register and acquiring the language of mathematics become most apparent in their transition from arithmetic to algebra.

## Language and the transition from arithmetic to algebra

Algebra may be introduced to students as a means to express arithmetic generalisations (Cobb et al., 1995). Ordinary language can be used to express these generalisations but the use of an algebraic symbol system broadens the ideas (Boero et al., 2002), by enabling the manipulation of the symbols according to a set of rules to create new patterns and reveal new mathematical insights. The symbols can be treated as mathematical "objects" and moved around according to formal rules (Sfard, 2002b) but often without students attaching any meaning to either the resultant patterns or the moves made to arrive at the new arrangement (Frid, 1993). The connection between the concepts and the expression of those concepts is lost.

When students have to work from a problem written in ordinary language, they have to attend to the semantics of the problem as well as to the symbolism. They may be able to solve the problem informally, but the demand of also using the symbol system may, for novices, be too great (Ellerton & Clarkson, 1996). The structure of the symbolic expressions and the translations of the verbal statements in problems are not "semantically congruent" and lead to errors such as the "reversal error" of the "students and professors" problem (Boero et al., 2002)

It has been suggested that the development of understanding by students of the structures of algebra may follow the historical development of algebra (Kieran, in Boero et al., 2002). From a verbalisation of the general patterns of numbers and functional relations developed a symbol system that used elements not found in the natural language, such as brackets and the equality sign. Also, the devising of processes for solving equations and using lexically dense symbolic expressions to communicate the logical steps of the solution, rather than verbalised procedure, broadened the applications to which algebra could be put.

The necessary attention to syntactical elements of mathematical language has meant that students need to have developed a competence in their own natural language to the extent that they can reflect on the structures of the natural language, either implicitly or explicitly. Hence, they come to appreciate the structures of mathematical language whether in verbal (oral or written) form or expressed symbolically. MacGregor and Stacey (1994) referred to this as "metalinguistic awareness" and drew analogies between word and syntax awareness of one's natural language and symbols and syntax awareness in the mathematical register. MacGregor and Price (1999) discussed the example of students required to find the answer for "6 – 10". Many students noticed the lexical items: a *6*; a *10*; a *subtraction sign*, but failed to take notice of the order in which the items were written. Lacking a real world referent, such as a drop in temperature or the spending of money, the students had no way of deciding on the correct response.

Schmidt (1996) remarked on the influence of students' early experiences when teachers introduced them to a limited, and often rigidly adhered to, range of meanings for arithmetic symbols. Typically, the *equal sign* was introduced as meaning "makes" or "gives the answer"; problems were presented as a series of arithmetic operations on numbers on the left hand side of the equals sign, to be "closed" to a unique answer on the right hand side. Schmidt noted that this was a direct translation of, for example, five plus three "makes" eight ($5 + 3 = 8$). It made little sense, however, to state "eight makes three plus five" ($8 = 3 + 5$), although mathematically this is a well-known formula with the truth-value intact.

The problem also occurred when students met expressions such as "$x + 2$". Often, students' experiences have led them to associate the signs for addition, subtraction, multiplication and division solely with the arithmetic instructions to do something. They do not see the signs as symbolising relationships between numbers and so fail to understand the many meanings of the expression "$x + 2$" (MacGregor & Stacey, 1994) that need to be taken account of when students embark on algebra. Austin and Howson (1979) recommended that teachers needed to be aware of these various meanings, and give students a chance to encounter them in order that they have a "robust" understanding of the mathematical ideas.

Algebraic thinking requires significant steps in mathematical abstraction and generalisation. The representations of algebraic concepts and processes of manipulation exemplify the meaning-dense nature of mathematical language embodied in the symbols of algebra. Although students can manipulate these symbols according to a learnt procedure, they find it difficult to explain the meaning behind the procedure. Various authors conclude that deep understanding of mathematical ideas is associated with the ability of students to express those ideas clearly, using complex sentence structures and arguments. For example, Boero et al. (2002) found that students in their first year of university who could not explain in appropriate, clear, and distinct language, elementary algebraic process were most of those who failed the course.

## Natural language and the development of mathematical concepts

Identification of the development of mathematical understanding is often focused on the content of students' discourse. Cognitive competency may also be displayed through the linguistic forms of a student's mathematical verbalising or writing – how the students convey the content. The development of the linguistic forms peculiar to mathematics, the mathematical register, occurs in concert with the conceptual development of the student.

Douek (2002) found that students' use of the relational and logical connectives *if*, *so*, *but* and *or*, and the quality of the links made through these speech items indicated the level of understanding of mathematical ideas possessed by the student. By the age of eight years, students used such connectives logically to explain causal relations (Donaldson, 1986). This meant that lack of such constructions in students' mathematical verbalising was indicative of cognitive challenge. Laborde (1984, in Jacobsen & Dawe, 1984, p.264) "noticed that if a mathematical notion is not well assimilated by students they might have considerable language problems when they try to speak about it".

Discourse of a mathematics classroom involves the speaker's conceptions of the mathematical objects, knowledge of the language in general and in the specific environment, conceptions of the interlocutors and the aim of the linguistic activity (Laborde, 1990). How the language operates in this social setting reflects attitudes, social boundaries and understanding of the material that is the content of the discourse.

Pragmatics is the branch of linguistics which examines and analyses language in use – "how words can be used to do things, to achieve one's ends" (Rowland, 2000). To study the use of language in the classroom, and what it conveys about students' understanding of mathematics, one needs to be aware of behavioural and attitudinal markers in the conversations. Examples of these have been identified in studies by Rowland (2000), Bills and Gray (2001), and Bills (2002). Rowland remarked in his studies on the students' use of pronouns, the attendant vagueness, and the modality of their conversations as indicators of conceptual security.

Bills and Gray (2001) and Bills (2002) reported studies of interviews of primary school students (six- to ten-years old) about their thinking as they carried out several mental computations, and as they explained non-mathematical procedures. They found that students who were more confident or successful mathematically tended to generalise. Such generalisation was marked by the use of the general *you*, and of procedures described in the simple present tense rather than the recount form of past tense. The more successful students tended also to disregard specific examples in their explanations, or to use a particular item as exemplifying a class of similar computations.

Both teachers and students use "extremely subtle pragmatic interpretive judgements [...] regularly in the course of mathematics teaching and learning..." (Pimm, 1987, p.167). The judgements, Pimm reported, were based partly on the mathematical content of what students said (or wrote) and partly on the way in which the students used the language.

Mathematical language has its own peculiar characteristics; like, yet unlike the natural language in which it is conducted in the classroom. This section has discussed the nature of mathematical language how language changes influence the move from arithmetic to algebra. When students encounter algebra they take their first, formalised steps in making and using generalised statements about mathematical behaviours. They also need the associated language with which to effect the transition from arithmetic to algebra. Recent research has shown that it is not only the mathematical content of what students say, but the way in which they express their thoughts that also provides insight into their understanding.

## CONCLUSION

Much of the research into the connections between mathematics and language has focused on the development of a mathematical register, and the associated difficulties encountered by teachers and their students. Developing a mathematics register for speakers of the native language in which the mathematics is conducted is challenging. Students need to acquire a vocabulary that often echoes the less precise vocabulary of everyday speech. They have to develop an awareness of the subtle differences in mathematics syntax to that of their everyday speech. Students also have to learn to interpret a set of meaning-dense symbols. These challenges are often made more difficult for students learning mathematics in a language that is not their native tongue. Translating mathematics to non-western cultures also might be problematic. However, the research agrees that acquisition of the mathematics register is essential for students to advance their understanding of the discipline.

For students to develop the mathematical register, an immersion model, adapted from that used for teaching a second language, has been found to be effective. Students need to work in an environment where mathematics is discussed. The research reviewed indicates that student-student discussion and student-teacher discussion in the mathematics classroom are important for several educational reasons. Being able to communicate their mathematical ideas allows students to organise and clarify their thinking, experience alternative perspectives and understandings, and enables teachers to identify levels of cognitive development. These levels of development are realised in students' utterances by the mathematical content, principally the mathematical terminology. Much of the research has focused on this aspect.

However, studies cited in the final section of the chapter suggest that while content knowledge might be necessary to establish the extent of a student's mathematical understanding, it is not sufficient. The form of the utterances – the coherence, the structure and modality – characteristics of natural language in use, can also be used as pointers to students' cognitive growth. It is these aspects that are discussed in detail in the following chapter.

# CHAPTER 2: THEORETICAL PERSPECTIVES AND RESEARCH QUESTIONS

The review of the literature suggests that it is only recently that connections between mathematics and language have been investigated in terms of the structural relationships that exist between mathematical and natural languages. Such research has focused on issues such as the role of natural language in mediating understanding of mathematical concepts (Boero et al., 2002), the extent to which students' understanding of the structures of their language may influence their mathematical competence (MacGregor & Stacey, 1996), and particular features apparent in students' natural language that may indicate their level of mathematical understanding (Rowland, 2000; Bills & Gray, 2001).

The research mentioned above examines two aspects of natural language as used by students. MacGregor (2002) and Boero et al. (2002) considered students' written responses to problems. Rowland (2000), Bills and Gray (2001), Bills (2002) and also Boero et al., considered the spoken explanations or discussions of students as they engaged in solving problems or took part in interviews. The research examined responses from students in primary schools (Rowland; Bills & Gray; Bills; Boero et al.), secondary schools (MacGregor & Stacey, 1996), and universities (Boero et al.; MacGregor, 2002). The research is essentially pragmatic, in that it examines the language "in use" by the students. The research provides a possible basis for a framework that articulates an empirical foundation for many of the judgements about the cognitive growth of particular students in the course of a regular class.

This chapter considers, in detail, empirical models of student cognitive growth in mathematics, developed from research documented by Boero et al. (2002), Rowland (2000), Bills and Gray (2001), and Bills (2002), in conjunction with the theoretical models (SOLO) developed by Biggs and Collis (1982) and later elaborated (Biggs & Collis,1991; Collis,1994), and Tall (1991). The algebraic context in which these models might be situated is discussed in a further section. Together, these aspects give rise to research questions, which are articulated in a final section of the chapter.

# EMPIRICAL MODELS RELATING STUDENTS' LANGUAGE USE AND MATHEMATICAL UNDERSTANDING

The research described in this section identifies characteristics of written or spoken language used by students that correlate with their mathematical success. The models based on these characteristics are considered to be "empirical" because they apply in a particular context, using quite specific descriptors of pragmatic language use that have been identified in particular clinical situations. Firstly, research on the written responses of students by Boero et al. (2002) is described. Secondly, the pragmatics of student responses as analysed by Rowland (2000) are discussed. Thirdly, research by Bills and Gray (2001) and Bills (2002) into possible links between the verbal explanations of primary school students and their success in arithmetic is examined.

## Students' written responses

Studies of language used by students when discussing mathematical situations have found that the complexity and organization of their speech is related to their understanding of the subject matter. This sub-section discusses research by Boero et al. (2002) into written explanations of an arithmetic problem by students entering university.

Douek (2002) suggested that a student's use of the relational and logical connectives *if, so, but* and *or*, and the quality of the links made through these speech items indicated the level of understanding of mathematical ideas possessed by the student. Donaldson (1986) found that by the age of eight years, students used such connectives logically to explain causal relations, rather than as associating events that simply occurred one after the other. Donaldson concluded that lack of such constructions in students' mathematical verbalising could indicate that students were struggling with the concepts. Laborde (1984, in Jacobsen & Dawe, 1984) "noticed that if a mathematical notion is not well assimilated by students they might have considerable language problems when they try to speak about it" (p.269). Involved in the discourse of a mathematics classroom are the speaker's conceptions of the mathematical objects, knowledge of the language in general and in the specific environment, conceptions of the interlocutors and the aim of the linguistic activity

(Laborde, 1990). How the language operates in this social setting reflects attitudes, social boundaries and understanding of the material that is the content of the discourse.

In a synthesis of their work on the role of natural language in the learning of mathematics, Boero et al. (2002) cited a study they conducted with 45 first year computer science students at university. These students were asked to solve a problem that was a "simple middle school…arithmetic problem" (p.247). The students were also required to explain their answer although no direction was given as to the form the explanation should take. Almost all of the students provided an explanation in words, sometimes accompanied by diagrams. The papers submitted by the students were classified according to the language used by the students in their explanations. The researchers used three levels of classification:

> Level 0(L0): no verbal comment, rambling words, poorly organised sentences;
>
> Level 1(L1): well-organised, semantically adequate, simple sentences; few compound and no conditional sentences;
>
> Level 2(L2): a good number of well-organised, semantically adequate compound sentences, including conditional ones (p.247).

Boero et al. (2002) further investigated the numbers of the students in the study who left the university and the performance of the students remaining in the algebra course. Eight of thirteen students who left the university had responses graded as Level 0, and five were from Level 1. The explanations of ten students were graded as Level 0. No Level 0 student managed to pass the algebra examination before May [the end of course] (pp.247-248). The researchers concluded that the inability on the part of students to express their ideas clearly, cogently and in some complexity was closely correlated with failure. They did not, however, claim that success was closely correlated with the ability to produce explanations that were at Level 1 or Level 2. Rather, they suggested that students whose explanations were at Level 1 or Level 2 possessed some "insurance against failure" (p.248).

Written responses, such as those provided in the research cited above, are likely to be a more refined product than spoken responses. A connection can exist between mathematical success and organised, informative and clear responses as they occur in regular classroom discourse. In these circumstances, both teachers and students use "extremely subtle pragmatic interpretive judgements [...] regularly in the course of mathematics teaching and learning..." (Pimm, 1987, p.167). The judgements, Pimm reported, were based partly on the mathematical content of what students said (or wrote) and partly on the way in which the students used the language. Judgements were based on language use as well as on behavioural and attitudinal markers that occurred in the conversations. Studies by Rowland (2000), Bills and Gray (2001), and Bills (2002), discussed in the following sections, identified such markers and their connection with students' levels of mathematical achievement. The examination and analyses of "how words can be used to do things, to achieve one's ends" (Pimm, 1994, p.167) is termed pragmatics, a branch of linguistics.

### The pragmatics of student explanations

This sub-section discusses aspects of Rowland's (2000) pragmatic analysis of student discourse. Rowland' analysis focused on mathematics talk recorded in transcripts of interviews and classroom discourse. Although pragmatics takes into account the use of language by participants in discourse to convey knowledge and negotiate social relationships, the focus in the instances cited by Rowland was on how the language used conveyed mathematical understanding. His main thesis was that vagueness in mathematics talk could be interpreted developmentally. Various features that are identified as characteristic of "vague" mathematical talk could be used as pointers to students' understanding of the mathematics under consideration.

Rowland (2000) interviewed primary school students from the ages of five-to-twelve years about various mathematical activities in which they were engaged. He used in-depth, "contingent" interviews and 230 "quick" interviews. The contingent interviews were clinical interviews based on particular questions, but which developed according to the responses from students. The shorter interviews made use of a tighter protocol of a series of set questions to which students' responses were recorded, but which the interviewer did not follow up "in depth". Rowland also drew on transcripts of interviews with his own undergraduate students and on records of classroom exchanges from early primary years to university classes in order to characterise developmental features of vague mathematical talk.

Rowland (2000) distinguished between *vagueness* and *generality*. Generality makes an indeterminate reference to anything in a class of things (mathematical objects) whilst vagueness makes use of an indeterminate class boundary (p.61). General statements are about well-defined classes, although the members of that class may not be differentiated. Vague statements imply an uncertainty about the class itself, and hence the members of that class. Vagueness helps students "say what they don't know how to say" (Rowland, p.67), and implies that the students are operating in a conjectural mode rather than with certain knowledge. Three features of vague mathematical talk, pronouns, hedges and modality, are examined in detail below.

### ***Pronouns***

A *pronoun* is "a word used instead of a noun substantive, to designate an object without naming it, when that which is referred to is known from context or usage, has been already mentioned or indicated…" (Shorter Oxford English Dictionary, 1975).

Pronouns are potentially vague; their meaning is dependent on the context (Rowland, 2000, p.73), and on correct syntax. Pronouns, used unambiguously, must be in close proximity to the singular, principal [noun] for which they are proxy. Nor should they come before the principal (Fowler, 1965, p.481). It is often through violations of these grammatical rules that pronouns achieve their vagueness. For example, "You put it on there and take it away from that" (verbatim quote from a student) is a statement that is ambiguous without the presence of the objects indicated by the pronouns. Of particular interest in the context of mathematical discourse is the use of the pronouns *I, you* and *it*, and their plurals.

Rowland (2000) concluded from his research that the pronoun *it* was used by students to denote concepts and generalisations that they could not name in the discourse, because they were not operating in the required mathematical register. The concept may have been formed in an intuitive way in the mind of the student, but because of its provisional and unclear nature could not be articulated in a conventional mathematical manner. The vagueness of *it* becomes apparent when transcripts are reread in the absence of the conversational context. Whilst *it* may carry clear references in the progress of a conversation, where the participants' use of gestures and/or voice modulations provide clues to aid the establishment of meaning, these are not readily apparent in the taped or transcribed versions of a conversation.

The use of *we* was found to be as a social marker. *We* refers to a school of practice or thought that may or may not include the participants in the discourse. A teacher may use *we* to denote the group of experts – teachers or mathematicians, for example – or to include the class members in an oblique instruction – "This is the way we behave". *We* may also be used by students, or their teacher, to include all members of the class group in describing a habit of practice in which they all participate (Pimm, 1987).

The use of the pronoun *you* has been found by Rowland (2000, p.75) to be an effective pointer to thinking that involves generality. Although teachers may use *you* in a personal sense, indicating a particular student, or group of students, students were found to rarely use *you* when addressing their teacher. This was because of the asymmetrical power relationships that exist in a classroom (p.209). Rowland found that students used *you,* instead of the more formal *one,* as denoting an impersonal, general, detached observation or explanation about a mathematical idea or process. The shift from pupils' use of *I*, *me* or *my* to the use of *you* or *your* marked a shift from explanations about experiences and feeling to explanations that were a more "generalised objective comment" (p.76). Students tended to use first or second person pronouns depending on whether they were talking about actions or knowledge (p.208).

### *Hedges*

For Rowland (2000), the presence of "hedges" and other modal forms gave "linguistic pointers to uncertainty and attendant cognitive vulnerability" (p.210). Identification of hedges is central to his characterisation of "vagueness".

The mode of an utterance conveys a speaker's attitude to that utterance. The mode may be conveyed by voice or gesture, or by certain linguistic formulations. In the context of mathematical development the forms by which students conveyed their commitment to or conviction about the factual content of the statements uttered – "epistemic modality" – became the focus. *Hedges* are words or phrases that are indicators of the attitude of the speaker or the modality of an utterance.

Rowland (2000) provided a taxonomy of types of hedges, which is summarised as follows. A hedge may be a *shield* or an *approximator*. Shields may be further subdivided into *plausibility shields* and *attribution shields*. Plausibility shields imply a position on the matter under discussion, using terms such as *I think, maybe, perhaps*. Attribution shields involve a third party, such as appeals to authority, and imply a neutral stance, or a shifting of responsibility to another, marked by phrases such as *mother said*.

Where shields mark a speaker's commitment to the truth-value of a proposition, approximators act within the proposition itself and serve to modify, or modulate, the truth-value of the proposition (p.65). Approximators may attach vagueness to the nouns or verbs or adjectives used, acting as "adaptors". For example, the use of the phrases *sort of, a little bit, about* makes class boundaries less distinct. Approximators act also as "rounders", particularly in the context of measurement. Rowland also noted that the words *around* or *about* might also be used as shields, as protection against error, thus indicating uncertainty as to the truth of the statement made.

### *Other Modal Forms*

Rowland (2000) has also mentioned other modal forms that indicate uncertainty, such as *prosodic hedges*, the use of *well* and *dispreferred second parts* in a conversation. Prosodic hedges are those where the voice is used to convey attitude, for example the rising intonation at the end of a statement, or students dropping their voice or mumbling. In Australia, the rising intonation at the end of a statement is a common speech habit (Fletcher & Loakes, 2006) and consequently might not be so significant in indicating uncertain knowledge.

The use of *well* by students to introduce a statement indicates that "the hearer should not expect an account that is clear and convincing" (Rowland, 2000 p.137). In a similar vein, participants in a conversation used what Rowland referred to as dispreferred second parts. These are attempts by one member of a conversation to modify a refusal or mark a lack of knowledge in such a way as to maintain the social relationships necessary to the conversation. Dispreferred second parts are marked by delays in responding to a question, the use of prefaces such as *well* or *I'd like to*, or the concoction of overlong and elaborate explanations that serve to mask a refusal, or lack of knowledge (p.92). The mood of uncertainty, or "lack of commitment to a proposition" could be further expressed in speech by long hesitations, use of the word *well* and interruptions by words such as *um* or a rising vocal inflexion at the end of a statement. The use of the word *well* can occur even when writing an explanation of mathematics if the student is at all unsure (Callingham, personal communication, 2005).

The following important developmental aspects noted by Rowland (2000) are: Modal forms and hedges, and hence vague language, developed with age, and were established by the end of primary school (p.150); modality might have been suppressed by students if they did not wish to appear not to know (p.169); and the absence of hedges implicated certain knowledge on the part of the student (p.141). Vagueness in statements might arise because the child lacks the particular vocabulary or because he or she has only a poorly formed concept of the mathematics to be explained. Vagueness, is important in a "conjecturing" environment and Rowland (2000) cited instances of students' speech patterns that indicated their attempts to develop new ideas.

### Primary school students' verbal responses

The following sub-section details work by Bills and Gray that elaborates on some of the speech patterns used by students as they described their mathematical thinking. Bills and Gray (2001), and Bills (2002), reported studies with middle primary school students (ages 6-to-9 years) that found particular linguistic characteristics which provided pointers to mathematical understanding.

The study reported by Bills and Gray (2001) was conducted over a two-year period with seven-to-nine year-old students, who, over the course of six interviews were asked to carry out 45 different mental calculations. Having done so, they were asked to describe "what was in their heads" as they did the calculation. Their responses were categorised as *particular, generic* and *general*.

Particular responses were those in which students used a "concrete" representation as a description such as specific numbers in a procedure or specific objects when explaining a concept. General responses were those which were most abstract, where a rule was given for a procedure and a definition for a concept.

Generic responses were those in which a specific instance was used as an example of a more general set of instances.

Students were also asked questions that were non-numeric and non-mathematical. Their responses to these were also categorised as particular, generic or general.

Bills and Gray (2001) found that: (i) the use of particular expressions was most likely to be accompanied by incorrect answers, while correct answers were more often accompanied by generic or general answers; high-achieving students (as measured by SAT scores) were more likely to use non-particular expressions than low-achieving students; students at all achievement levels used general forms in non-calculation contexts; and, (ii) the more successful students switched to particular expressions when faced with more difficult questions, whilst the least successful students used the same proportion of general and particular expressions. In more global terms these results can be summarised as: accurate answers were most likely to be explained in general or generic terms; and, students whose calculations were more likely to be accurate tended to use more general or generic terms than those students who made more computational errors.

It was also found that the students who made fewer errors tended to reserve "non-particular" expressions for answering questions involving calculation. On the other hand, these students tended to use more particular expressions when answering non-calculation questions.

In a further study, Bills (2002) analysed two thousand responses from twenty-six students aged six-to-nine years for pronouns, tense and causal connectives used in their descriptions of mental computations. He also asked students questions that were non-computational and observed the teaching in the classroom. Drawing on work by Rowland (2000), Bills examined the association of the pronouns *I, you* and *it*. *It* was often used as a general pointer to some mathematical object. *I* was used in a particular sense associated with the description of a singular instance, or in a more general sense when students cited a rule. *You* was invariably used in a general way, as a commonly accepted replacement for the more formal *one*.

Comparisons were also made between the success of students who phrased their explanations in the past tense and those who used the simple present tense, as well as those who elaborated their explanations using causal connectives such as *because*, *so*, *if* and *then*. Bills found that: there was the same spread of linguistic characteristics across the abilities when students responded to non-calculation questions; a higher proportion of causal connectives was associated with correct calculations; teacher language was often adopted by the more successful students; and, correct calculations were most often accompanied by explanations using *you*, simple present tense and using explanatory words of deduction.

The conclusions reached by the two associated studies by Bills and Gray (2001) and Bills (2002) are that differences in the linguistic structure of student's descriptions and explanations of mental computations may be pointers to their different mental constructions and to different qualities of thinking.

The research described in this section has identified certain characteristics of language used by students that might be associated with different levels of mathematical competence or understanding. The models based on these characteristics are "empirical" because they apply in a particular context, using quite specific descriptors of behaviour that have been identified in particular clinical situations.

Linguistic features identified by the researchers as pointing to differences in cognitive development include students' use of pronouns, verb tense, and non-particular statements. Students, whose statements are confidently voiced, clearly organised, and who used logical connectives to elaborate ideas, were more likely to be successful. Modality features such as the use of linguistic hedges, voice inflexions and gestures, were also found to provide clues to levels of mathematical understanding. These levels of mathematical understanding reflect stages of cognitive development in the student.

## THE SOLO MODEL OF COGNITIVE DEVELOPMENT

Researchers (e.g., Austin & Howson, 1979, Pirie, 1998; Sfard, 2001) have noted that language development and cognitive development are closely associated. Language is essential if teachers are to gain insight into students' thinking.

Although, as Pirie pointed out, simple ideas can proceed without language, language is the only way we can communicate more complex ideas. Thus, findings from the empirical studies described in the foregoing section can be situated in a theoretical framework that describes levels of cognitive development. Such a framework is that of the SOLO (**S**tructure of the **O**bserved **L**earning **O**utcome) Model, developed by Biggs and Collis (1982).

In 1982 an Occasional Paper on *Language Development and Intellectual Functioning* was presented by Kevin Collis to a UNESCO forum. It linked language development from words to propositional speech with a post-Piagetian model of "intellectual functioning" (p.5). The language development so identified has been further supported by the findings of Boero et al. (2002) described in the previous section. The model of intellectual functioning was a broadly-based, and broadly-applicable theoretical model developed by Collis and co-author Biggs as the SOLO Taxonomy (Biggs & Collis, 1982), later renamed as the SOLO Model.

This section firstly describes the rationale behind the development of the model. Secondly, the original model is described and this is followed by a description of subsequent changes to the model. A discussion of the connection between the SOLO model and cognitive theory follows. Ways in which the model can be used to explain language development and development in mathematical understanding are also described. Comparison of the SOLO model with one other cognitive model is also made.

## Rationale for the SOLO Model

The SOLO Taxonomy was developed by Biggs and Collis (1982) as a theoretical model to describe levels of cognition as demonstrated by a student's response to some particular item. The SOLO Taxonomy developed out of Biggs' and Collis' concerns with the methods used for the evaluation of learning by school students. Throughout its evolution, from the model described in 1982 (Biggs & Collis, 1982) to an elaborated version (Biggs & Collis, 1991) the SOLO model has sustained its research base in school teaching and learning practices. Its attractiveness is: its "objectivity" (the value-free nature of its assessment of task outcomes); its focus on observable outcomes of learning; its broadening applicability to ways of thinking other than academic learning; and, its emphasis on the quality of learning, the depth of understanding, rather than solely on the quantity of material consumed and recalled by the student.

Along with other workers in the field, Biggs and Collis became increasingly uncomfortable with the Piagetian model of cognitive development, particularly in its identification of individuals' development with their responses to tasks set. The SOLO model resulted from research into students' responses, which were seen as linked to the stage of their cognitive development, but not identical with it. The model attempts to classify specific responses of particular students to particular tasks at a particular time. The labels used in the model are attached to the response to the task rather than to the student, as was the case with the Piagetian stage model of cognitive development (Biggs & Collis, 1982, p.23).

Piaget's model could not satisfactorily account for apparently atypical behaviours of students, what Piaget termed "décalages". Students whose cognitive development had been labelled as "concrete symbolic" (able to use words and symbols to express ideas) would suddenly revert to using pictures. This represented a lower order of thinking according to the Piagetian model, which should have been left behind as the student entered the concrete symbolic stage. Biggs and Collis (1982) found that such behaviours were not exceptional. They did not necessarily signify carelessness, poor teaching or poor learning strategies, or aberrant development; they were, simply, the response of the child at that time to that particular task. This view differed from that of a cognitive development model, based on Piagetian stages, which attempts to identify and measure *ability*, as distinct from *attainment*. Biggs and Collis drew a fine distinction between ability and attainment:

> …Ability, expressed as […] IQ, is relatively stable over time and is effectively independent of instruction. Attainment refers to the result a student obtains in a particular test about circumscribed material that has specifically been the subject of instruction (p.22)

Because education needs to focus on aspects of the student which might be altered through instruction, Biggs and Collis felt that there was a need to identify and describe those aspects rather than focus on aspects that could not be influenced by teaching. It was the stated aim of the SOLO model to provide a criterion-referenced measure of the quality of learning (Biggs & Collis, 1982, p.7).

## Stage Theory and the Early SOLO Model

Stage theory as described by Piaget had assumed that: the logic possessed by a task was the same logic applied by the student in completing the tasks; logic used other than that required according to the developmental stage of the child placed the child in a [usually] lower stage of development; and, performance on one set of tasks typified performance on all tasks. These assumptions led to the conclusion that, as developmental stages of cognition were fixed, they determined how a student would respond to tasks in any context, and could learn new material only when *ready* (Biggs & Collis, 1982, p.21).

The focus was on logical functioning or academic intelligence – how students performed on school tasks in school situations. The work of Biggs and Collis also concentrated, initially, on academic school achievement. They did not entirely reject the structuralist view-point (Biggs & Collis, 1982, p.213) but adopted a modified framework which acknowledged that: developmental stages [*modes* in SOLO terminology] apply insofar as they provide an upper limit of functioning for the student, but they do not dictate that the student must function at that level all the time (p.22); such stages are age-related; biological influences affect development and these are linked to the age-relatedness of stages (p.223); and, stage is a function of working memory capacity which is influenced by the novice-expert progression (p.212).

The departure from the Piagetian model of stage development is that, once available, all modes may be accessed by the student as new knowledge is acquired. This has important implications in discussions of the development of expertise in areas of knowledge that are principally physical or pictorial, and which are addressed in later versions of the SOLO model.

In establishing the SOLO framework, learning responses were categorised according to the quality of the response, involving four dimensions: working memory capacity; operations relating task content; consistency within a response and the need to close on a response; and, general overall structure that results from the interaction between the previous three dimensions (Biggs & Collis, 1982, p.31).

## Organisation the early SOLO model

The SOLO model describes development in terms of *modes* and *levels*. The modes describe the broad thinking behaviour of the student, which is identified by the abstractness of the task and the subsequent response to the task (Pegg, 2003, p.242). The modes increase in complexity and abstraction, moving from the world of immediate experience to that residing in the mind. Levels describe the development within each mode. "A level refers to a pattern of thought revealed in what a student says or does" (Pegg, 2003, p.243). These levels also increase in complexity and abstraction.

### *Modes of the SOLO Model*

The Modes describe typical ways of knowing – how an individual might function cognitively. The early SOLO model followed the Piagetian model of four main stages, identified as: *sensorimotor, intuitive, concrete*, and *formal* modes. This last was elaborated as first, second, and third orders, suggesting that the authors considered the possible existence of further, unresearched developmental modes (Biggs & Collis, 1982, p.216). The modes are described in the following.

#### *Sensorimotor mode*

The sensory-motor mode is available from infancy, although its nature changes in adult cognition (Biggs & Collis, 1991, p.190). In this mode, the "learner is concerned with learning motor responses and sensory discrimination"(Biggs & Collis, 1983, p.153).

#### *Intuitive mode (later termed Ikonic, 1989)*

The ikonic mode develops in pre-schoolers, wherein actions are internalised, or visualised as images. In this mode, much of the focus is also on "the mastery of oral language" (Biggs & Collis, 1982, p.211). This mode, in adults, is different in kind, but as discussed later, is far more important in adult thinking than Biggs and Collis realised in the initial formulation of the SOLO model (Collis, 1992, p.28).

*Concrete mode* (later termed *Concrete symbolic,* 1989*)*

The Concrete mode is typically used during the years of formal schooling, and consequently the focus of early SOLO models. It is concerned with the basic concepts and operations which are "directly tied to the empirical world" (Collis, 1992, p.26) and with their representation in a consistent symbolic form.

*Formal mode*

The formal mode arises during adolescence, during which individuals begin to hypothesise about matters beyond their immediate experience.

*Post-Formal modes*

Post-formal modes are those in which the structures of existing theories and disciplines are considered. The existence of these modes was anticipated in the early model by Biggs and Collis proposing second, and third orders in the formal mode (1982, p.218).

### *Levels of the SOLO Model*

Within each mode of development, Biggs and Collis identified five *levels* of learning response. The two extreme levels may be described in terms of their relationship to the succeeding or preceding modes. Thus, learning responses manifest as learning cycles within each mode (Biggs & Collis, 1982, p.214). Just as developmental modes increase in complexity and abstraction so too, within each level do the responses gain in complexity of the detail that can be handled and in the organisation of that detail. It is the ability to reorganise information that determines how much detail the student can handle. Similarly, it is the need to organise detail that prompts the student to move to another level, or another mode. The characteristics of the SOLO model are schematised in Figure 2.1.

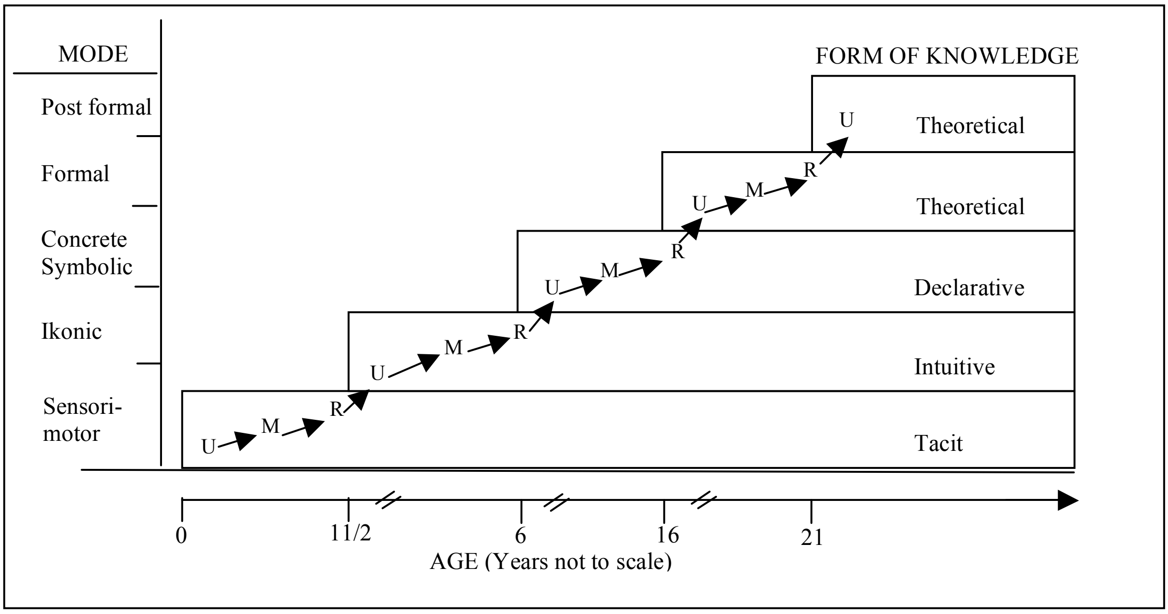


*Figure 2.1: Modes, learning cycles, and forms of knowledge of the SOLO Taxonomy (Collis, 1992, p.370)*

Students, no matter the mode in which they operate, progress through five levels that share common cognitive developmental features. Students progress from:

*Prestructural* level where they demonstrate a disinclination to engage with a task or concept (e.g., Biggs & Collis, 1982, p.79); to

*Unistructural* level where the task or concept can be dealt with if one relevant aspect only is considered; to

*Multistructural* level where several aspects of the task, or numerous data, are accumulated but addressed in isolation; to

*Relational* level where all relevant aspects are assembled into a coherent and consistent whole with respect to the data given; to

*Extended Abstract* level, where students attempt to go beyond the data given, taking a more global perspective. This level of reasoning becomes the unistructural level of the next mode.

## Changes to the SOLO Model

By 1991 Biggs and Collis (and other researchers, e.g., Leder, 1992, p.26) realised that other forms of intellectual functioning coexisted with, or even overrode, academic thinking in a large proportion of the population, most of the time. Because their focus was on *evaluating the quality of learning outcomes*, Biggs and Collis tied SOLO to a learning model that took account of the interaction of teacher expectations and instructional style with the student's own prior knowledge, intentions, motivation and expectations (Biggs & Collis, 1989, p.10). Such factors are not necessarily attached to any one developmental stage, but influence the extent to which a person can effectively function within a developmental stage. Thus, a developmental stage provides an upper limit of functioning, but does not dictate that one must function at that upper limit (Biggs & Collis, 1991, p.22).

More importantly, Biggs and Collis recognised that once developmental modes appeared, they remain, not subsumed by later modes, but available for use at all stages of learning. This was the concept of *multi-modal* thinking, foreshadowed in the initial, 1982 model (p.227). Depending on the skill to be learnt, or the knowledge to be obtained, a student operated in a *target* or predominant mode (Biggs & Collis, 1982, p.187), supported by other modes as needed. Thus, a gymnast would operate at the adult sensorimotor level (which is significantly different from that in a child) but would employ the theoretical perspective of the concrete symbolic mode and the visualisation skills of the ikonic mode to reflect on and enhance a performance.

It has become apparent that when dealing with unfamiliar concepts adults, even experts, may use ikonic thought to deal initially with a task, and then move on to concrete symbolic and formal modes. The proposition that adult experts do not always function in the formal modes was highlighted by Hadamard (1954, in Collis, 1992, p.28), whose work has prompted cognitive theorists to re-examine their definitions of "intelligent thought". When expert mathematicians or scientists rely on flashes of insight to come up with profound theories, it points to a danger in considering developmental modes of thinking as strictly hierarchical.

By accommodating the fact that acquisition and exercise of different types of knowledge might require different modes of cognition, the SOLO model ceased to be perceived as strictly hierarchical in nature. However, "other" modes, used to support a target mode, were thought to be relegated to a peripheral role once expertise in the target mode was attained. Research on problem-solving and adult learning strategies revealed that students switch between modes of thinking in order to complete a task (Collis, Watson, & Campbell, 1993).

The SOLO framework has developed from a model designed to describe intellectual functioning in a school-based, academic context to a model which can be applied to learning in a variety of contexts and modes. The following, elaborated model describes the modes and levels in detail that convey the broader applicability of the SOLO model. Some descriptions contained in that of the earlier model are repeated for the purpose of clarity.

### ***SOLO modes in the revised model***

In the revised model, the modes of learning can be described as follows.

The *sensorimotor mode*, begins to operate soon after birth. It is the mode in which *tacit* learning, in response to the physical environment takes place. This is the mode in which the very young acquire motor skills and in which, in older age groups, the skills associated with various sports, or the learning of a musical instrument (for example) are predominantly situated and developed. Individuals operating in this mode may be able to carry out an action, or series of actions, but be unable to describe or justify that behaviour.

The *ikonic mode*, begins to operate around the age of two years. Actions are internalised as images and the young child develops a lexicon of words and images that represent objects and events. Later on, this is the mode that functions in the development of appreciation of art and music. Cognitive understanding in this mode may be considered *intuitive*.

The *concrete symbolic mode*, becomes available from about the age of five-to-six years. "Concrete" experiences, rooted in the real world are represented and understood through the mental manipulation of systems of symbols (words, numbers, or music notation) (Pegg & Tall, 2005). This is the mode in which students in the later years of primary school and secondary school are expected to function, albeit with the support of the two earlier modes. Cognitive understanding in this mode may be considered as *declarative*.

These last two modes are often used in support of learning in the sensorimotor mode, particularly at expert levels (e.g., Sport coaching in Biggs & Collis, 1991, p.71)

The *formal mode* begins to operate at about the age of 15 or 16 years when the individual considers abstract concepts without their necessarily being linked to some concrete referent.

*Post formal modes* develop at around the age of 22 years and become apparent when an individual can consider the underlying structure of theories or disciplines (Biggs & Collis, 1991, p.66; Figure 2.1)

### ***SOLO Levels in the concrete symbolic mode***

The five levels, which operate within each mode, have been described earlier. Here, they are described as they appear in the concrete symbolic mode, the mode in which much of formal schooling takes place (*target mode*).

*Prestructural level (P)* responses are characterised by their being non-existent (students cannot begin to answer the question), irrelevant or inconsistent. Students operating at this level provide answers very quickly without any thought as to their appropriateness (the "shot-in-the-dark" response). Responses at this level are essentially responses in a mode acquired earlier, usually the ikonic mode, though sometimes the sensorimotor mode. Thus, knowledge may be described as being either intuitive or tacit.

*Unistructural level (U)* responses demonstrate understanding or recognition of only one relevant aspect. Conclusions reached by a student may be based on only one piece of information. They may be produced quickly (rapid closure) and thus there are often inconsistencies between the response and the question or information.

*Multistructural level (M)* responses are typified by many aspects of the task being listed as singular, isolated concepts. The student may be aware of conflicting information or inconsistent conclusions, but does not appear to be able to resolve these (the "shopping list" response) in an attempt to close on an answer or solution.

*Relational level (R)* responses demonstrate that a student can relate all relevant facts identified at the previous multistructural level into a coherent whole within the limited framework set by the context. When the student moves beyond the given information or beyond his or her experience, then the responses can be inconsistent. Students can delay closing on a response until all data have been considered and accounted for.

*Extended abstract (EA)* level responses are characterised by successful attempts to deduce results from general principles or to try to establish general principles from the relationships understood between specific facts. Students at this level can resolve inconsistencies and can consider alternative solutions or the effects of different constraints on a given situation. Thinking at this level means that the student has moved from the concrete symbolic mode into the formal mode.

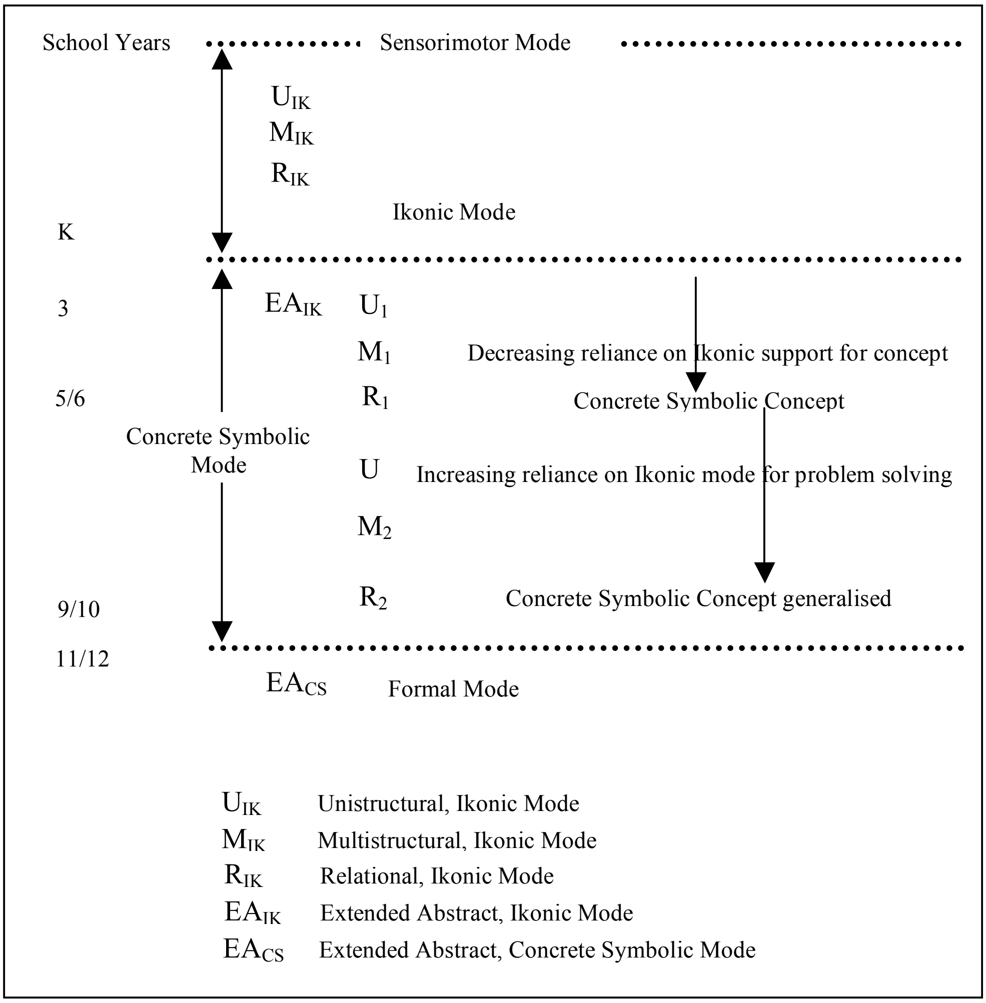


*Figure 2.2: Learning Cycles (The structure from* ***K-12****) (Collis, 1994, p.346)*

Later studies (e.g., Collis, 1994; Collis, Watson, & Campbell, 1993; Pegg, 1992) established that in the concrete symbolic mode there are at least two cycles involving the unistructural (U), multistructural (M) and relational (R) levels. The first cycle (represented as U1, M1, and R1) represents "an interface between responses in the ikonic mode and [those in the] concrete symbolic" mode (Pegg, 2003, p.244). The student moves from using ikonic support for concept formation to the concept established in the concrete symbolic mode. The student comes to understand a concept as it is represented by symbols – either as words, words in sentences, or mathematical symbols such as numbers, or numbers in arithmetic expressions. During the second cycle (represented as $U_2$, $M_2$, and $R_2$), the student uses the concept developed during the first cycle to solve problems, firstly with particular examples, but at the $R_2$ level, the concrete symbolic concept has become generalised. This movement is summarised in Figure 2.2.

The diagram illustrates the UMR cycles as they appear in the concrete symbolic mode. Similar cycles have been identified in earlier modes, and in the formal mode (Pegg, 2003, p.245). Also illustrated is the recognition that learning in the concrete symbolic mode may be supported by thinking in the earlier modes. The first UMR cycle uses understanding established (usually) in the ikonic mode to build concrete symbolic concepts. In the second UMR cycle, the ikonic mode is thought to provide important support for problem-solving (use of the concepts previously established) in the concrete symbolic mode. Students move through the two UMR cycles before reaching the extended abstract level in the concrete symbolic mode, the unistructural level in the formal mode. How, or why, movement from one level to another, or from one mode to another, is stimulated or effected is discussed in the following sub-section.

### The SOLO model and cognitive theory

The SOLO model was not intended to provide a theory of cognitive development. Biggs and Collis suggested that one may infer cognitive development from behaviours exhibited by students, a point also made by Pirie (1998). They were wary of going beyond the concrete evidence of learning responses into postulating the existence of a "generalised cognitive structure of the individual", that was not directly measurable (Biggs & Collis, 1989, p.22). However, their taxonomy of response levels is related to student's thinking. Mechanisms whereby a student moves through levels and modes needed to be postulated in order to tie the responses being analysed to the cognitive attributes they signify.

SOLO levels depend upon information-processing abilities, explained in terms of working memory. Working memory (or short-term memory) is “hypothesized as that part of the mind in which immediate conscious thought processing takes place” (Collis, 1982, p.6, footnote). Students may need to hold a quantity of unrelated data in working memory in order to complete a task (successive synthesis); or they may keep related data in working memory in order to form interrelationships between the various items (simultaneous synthesis). It is postulated that working memory becomes larger as children develop, although only to a certain limit (some learning theorists suggest seven disparate items). Therefore, an increase in working memory beyond this hypothetical limit becomes a matter of organising isolated data into categories or familiar patterns, which can then be connected.

If data can be fitted into pre-existing organisational frameworks in one’s mind and thus related to prior knowledge possessed by the individual, then that data can be dealt with more efficiently by considering patterns of interrelationships. Case (cited by Collis, 1982, p.6) regarded working memory as consisting of a component into which the data to be worked with is held, and another where the data are operated on. As working memory has a limited capacity, the more attention paid to the discrete data items, the less space remains available to process or connect the data. As ideas become more familiar, they become more automatic. Automatisation of data moves it from working memory into long-term memory. Data become *encoded* – the more interconnections formed among data, the more deeply embedded the data become and the more readily recall can be triggered by a cue in working memory.

Encoding of data fits information into a pattern or *schema,* and this schema provides an organisational basis for identifying potentially similar ideas or information. By having in working memory a set of schemata (more or less elaborate, depending on the intellectual level of the task and the complexity of the associated detail) working memory can accommodate more data. Relationships can then be established, not between isolated data items but between patterns, or schemata. When schemata develop a global perspective they can become generalised principles that enable individuals to deal with data not yet within their immediate sphere of experience. The student has moved beyond one mode into the next – from the extended abstract level of the earlier mode to the unistructural level of the next, later mode.

The move from the sensorimotor mode to the ikonic mode seems to be a "natural" progression (Collis, 1992, p.23). Societies with no formal schooling systems have adults functioning well in that society who can visualise circumstances and informally communicate ideas. However, when cognition is confined to the sensorimotor or ikonic modes it is difficult to communicate complex concepts in an unambiguous way. Without a consistent symbol system, it is difficult to abstract and form general principles, which in turn hinders clear, efficient thought and communication.

Our society places great value on the development of abstraction from the real world of immediate events to generalised hypotheses about those events. Communication of our ideas about the world demands a system of symbols that relate logically to the world, as well as sustaining an internal logic. Mastery and refinement of the symbol system by which we gain control over our environment is the purpose of formal schooling. The data in this context remain attached to real situations, but become increasingly complex so that handling the data effectively requires an increasing degree of organisation and a developing perception of interrelationships.

Because the situations dealt with are still founded in the everyday world, schooling is largely carried out in the concrete symbolic mode. However, it is apparent that as students develop ever-broadening schemata, ever more complex patterns to deal with data, they need to develop more general principles to encompass all the details of specific contexts. The consequence is that they need to take a longer view – to step back and see the forest, rather than individual trees. The symbolic system at this point dissociates itself from the thing-it-represents to assume a life of its own that characterises an academic discipline. Students move from the concrete symbolic mode into the formal mode. Their thinking has progressed through the unistructural, multistructural and relational levels in the concrete symbolic mode to the extended abstract level in that mode. This may also be thought of as the unistructural level in the formal mode.

It is the transition from the relational level in the concrete symbolic mode to the extended abstract in that mode that requires a huge cognitive leap (Biggs & Collis, 1982, p.224). Students have to be confronted by a task for which new strategies, not available at the relational level, have to be developed. How students deal with this situation depends on endogenous factors such as motivation to solve a problem or complete a task, as well as being able to assimilate information in ways not provided by considering just context-specific data. The student is forced either to move from inductive reasoning from the data to propose a theoretical principle which transcends the given, and so form deductions from that hypothesis which is tested against the given data, or ignore the problem. The first response is that of students moving to the unistructural level of the formal mode (i.e., operating at the extended abstract level in the concrete symbolic mode). The second response is that of students operating at the relational level in the concrete symbolic mode, but which is considered to be Prestructural with respect to the formal mode (i.e., Pre-formal) (p.225).

The first response indicates that individuals can organise data in such a way that a general, more global pattern can be found. Working memory can be cleared and information can be more easily dealt with simultaneously because there is an efficient and effective framework within which data can be encoded. The second response indicates unwillingness or inability (in terms of insufficient experience or knowledge) on the part of the student to be engaged in the task of further abstraction.

## The SOLO Model and Language Development

In the 1982 paper presented at the UNESCO forum Collis linked the SOLO framework to the development of language (p.5). His schema is summarised as follows.

The *Sensorimotor* mode requires and uses no language.

The *Ikonic* mode (in Collis' paper it is termed Intuitive/Pre-Operational) is the mode in which ideas are verbalised, words being attached to images (real or mental).

In the *Concrete Symbolic* mode (*Concrete Operational*, 1982) words are arranged into written sentences.

In the *Formal* mode, the sentences become propositional statements.

In each of these modes, Collis proposes that levels such as those described above operate in a sequence as:

> …Single-word sentences [are] the achievement of the sensori-motor stage; sentences at a mature level of complexity for communication purposes the achievement of the next stage; and finally fully propositional logic being the end-result of the concrete operational stage (Collis, 1982, p.6).

The language model can be further elaborated.

The *Sensorimotor* mode requires and uses no language. This is immediately apparent when one observes infants, but is also apparent when older children (or adults) can copy an act, or demonstrate an action, but are unable to tell how something can be done.

The *Ikonic* mode is when words are spoken, either singly in the early stages, or in sentences at later stages. Communication at this level may also be in the form of drawings or in the use of phrases such as *like* or *looks like*, as noted by Davey (1988). It may be that when operating in this mode, older children, or adults, use vague terms such as *it* (as discussed by Rowland, 2000), to refer to an image, either mental or real, but for which they have no associated word.

In the *Concrete Symbolic* mode, words begin to be arranged into written sentences. As this is the mode in which, as noted earlier, formal schooling takes place, the development and use of language and the associated symbol systems is closely linked with the development of understanding in a number of academic contexts. The complexity, detail, relevance and coherence of the sentence or sentences, reflect the level of understanding. Biggs and Collis (1982) and Collis (1982) provided several examples representing how the levels of understanding are expressed in this mode.

In the *Formal* mode, the sentences become propositional statements, reflecting the student's ability to abstract ideas from the immediate context and make statements about generalised principles.

Mathematics, or mathematical understanding, develops through these modes. In the *sensorimotor* or *ikonic* modes, young students use concrete materials or pictures to illustrate mathematical (arithmetic) concepts such as counting or measuring, or to carry out arithmetic operations. In these modes, teachers may use appropriate mathematical terminology, and encourage the students to use it, gradually incorporating the mathematical words into simple sentences that describe the mathematical ideas.

Mathematics, although beginning in the early modes operates predominantly in the *concrete symbolic* mode and *formal* modes at school. Support from earlier modes continues and, in the case of some particular problems, remains essential to the establishment of understanding as students operate in different modes according to the task and the stage of solution of the problem (Biggs & Collis, 1991). The move from the *concrete symbolic* mode to the *formal* mode requires, as Biggs and Collis (1982) acknowledge, a huge cognitive leap that is not unlike the *proceptual divide* described by Gray and Tall (1991, 1992a, 1992b). The cognitive model proposed by Gray and Tall, has much in common with the SOLO model, and is of particular relevance to this study.

### The proceptual model of Gray and Tall

The SOLO model bears family resemblances to other cognitive models, both in the global stages and in the UMR sequences which operate as *local cycles* in each of the modes (Pegg & Tall, 2005).

> …The UMR cycle … operates in the construction of new concepts as the individual observes what is initially a new context with disparate aspects that are noted individually, then linked together, then seen as a new mental concept that can be used in more sophisticated thinking (Pegg & Tall, 2005, p.470).

One model that is closely related to the SOLO model (and which is of particular interest to this study because of its focus on the development of algebraic understanding) is the framework of Gray and Tall (1994). It describes the development of understanding of a mathematical concept from the application of a particular procedure, through the acquisition of a collection of procedures that achieve the same mathematical end point, defined as a *process* to an understanding of the connections between the procedures. The framework attempts to describe how students develop a versatile understanding of mathematical concepts (from early counting in the ikonic mode) and the associated symbols (in the concrete symbolic mode).

This understanding is an essential precursor to algebraic thinking. An expression, such as "$x + 2$" needs to be considered either as a "potential operation" to be carried out if $x$ is known, or as a statement of the relationship between $x$ and 2. If the latter, "$x + 2$" has to be treated as a mathematical object to be manipulated. This versatile understanding, seeing such statements as either symbolising a process or standing for a mathematical concept, Gray and Tall termed a *procept*, an amalgam of the notions of process and concept (1992, p.2). The development of mathematical understanding thus may proceed from the use of a particular procedure on an already conceptualised mathematical object to the use of various procedures, but without conceptual understanding. Once procedures can be used flexibly, students conceive of the mathematical symbols as signifying a process, or set of procedures, one of which they may choose to use. The final stage in development is where students see the mathematical symbols as a procept. The development may therefore be seen as procedure-multiprocedure-process-procept. This aligns with the SOLO local cycle of unistructuctural-multistructural-relational-unistructural, which may occur in a new cycle in the same mode, or in a higher mode (Pegg & Tall, 2005).

The SOLO model was developed to provide criteria for evaluating the quality of (school) student understanding through the products of their thinking. It was based on the Piagetian model of stages of cognitive growth, but different to that model in that each stage remained accessible to the student at all times. The stages were described as modes of thinking – sensorimotor, ikonic, concrete symbolic and formal. Within each stage, levels of development of thinking were proposed – unistructural, multistructural, relational and extended abstract (unistructural in a higher mode). In later formulations of the model, these levels of thinking were conceived to exist as cycles in each mode, each cycle building cognitive sophistication until conceptual understanding became qualitatively different, thus moving to another, higher mode.

One context for the application of the SOLO model is that of students moving from arithmetic to algebra. This movement typically occurs at the beginning of secondary schooling, although the foundations have been established to some extent in primary school.

## ALGEBRA

> Mathematics is a reasoning and creative activity employing abstraction and generalisation to identify, describe and apply patterns and relationships…The symbolic nature of mathematics provides a powerful, precise and concise means of communication. (Board of Studies, 2002a, p.7)

The abstract, general and symbolic nature of mathematics is exemplified, at least in school mathematics, by Algebra. The New South Wales (NSW) Year 7-10 Syllabus (2002a) emphasises the communicative power of mathematics, and the processes of doing mathematics such as making and testing conjectures, applying logical reasoning and providing proof of conclusions reached. This pedagogical stance that underlies the content of the mathematics syllabus means that students need to learn not only how to do mathematics, but to learn to think mathematically. Hence, the development of algebraic understanding involves more than the development of proficiency in manipulating symbols according to sets of rules. Students need to perceive the general relationships that exist between numbers and systems of numbers and which are encapsulated in mathematical symbols.

Students need to express their understanding in ways that make meanings clear – to them as well as to others. The upshot of this is that the development of algebraic thinking, like the development of all mathematical thinking, involves students in verbalising their thoughts – either informally, or, increasingly using the conventions of mathematical discourse, developing the mathematical register.

Thus, how students use language to describe their algebraic thinking is as much a part of their algebraic, and more general mathematical, development as is their use of certain procedures to effect mathematical solutions. This section considers some aspects of Algebra, and the association between Algebra and the SOLO model.

## Aspects of Algebra

Foundation algebra concepts are built in the primary years of schooling through pattern recognition and completion, developing descriptions of patterns from everyday language to mathematical (arithmetic) language, developing understanding of number relationships, and representing numeric and geometric patterns that involve one operation as tables and word sentences (Board of Studies NSW, 2002b). In the first years of secondary school, introductory topics in algebra include: the use of letters to represent numbers; the translation of word sentences to algebraic symbols; the generalisation of number patterns; the manipulation of letters according to rules first encountered in arithmetic contexts; and the solution of linear equations.

Students' success in algebra depends, therefore, on their making the conceptual shift from the use of particular numbers to the generalised letter, and in perceiving relationships between numbers in terms of general principles rather than as particular instances. This means that students need to understand the arithmetic symbols (or words) for addition, subtraction, multiplication and division not only as instructions to operate on a set of numbers, but also as describing relationships between numbers (Warren, 2003). The intent of the arithmetic symbol must be elicited from the context and so students are required to move between two different, but related, concepts. Tall and Thomas (1991) referred to this conceptual flexibility as a *procept*. Students with a proceptual understanding conceive of expressions such as "$x + 3$" as either, indicating that three is to be added to $x$ or that there exists a number that is three more than some other number. Even the arithmetic expression "$4 + 5$" can be conceptualised as the result of adding five to four, giving a result of nine, or simply as representing a number that is five more than four (or *vice versa*).

Much of the literature into the nexus between mathematics and language has also focused on algebra (e.g., MacGregor & Price, 1999; MacGregor & Stacey, 1994; Sfard, 2001; Sierpenska & Lerman, 1996). The focus of the research has been on understanding how students give meaning to the symbols and syntax of algebra, and the consequent results of their manipulations of algebraic expressions.

## Algebra and SOLO

Algebraic thinking involves the ability to perceive relationships between numbers and to generalise those relationships. This generalisation results in the use of letters rather than particular numbers, and the manipulation of relational expressions as mathematical objects (Bills & Gray, 2000; Sfard, 2002b; Tall & Thomas, 1991). Therefore, in terms of the SOLO model as outlined in this chapter, the thinking required to carry out algebraic processes successfully occurs in the formal mode.

However, many introductory algebra examples can be dealt with arithmetically, particularly the solution of simple one- or two-step linear equations which have small positive-integer solutions (Tall & Thomas, 1991). Arithmetic thinking, dealing as it does with symbols which represent abstractions from the real world, occurs in the concrete symbolic mode of the SOLO model. It is not until students meet examples that demand their attention to number relationships (and they conceive of the idea that the obtaining of a unique, closed answer is not always the object of mathematical manipulation) that they develop an orientation that might be described as *algebraic thinking*.

Therefore there is a transitional stage between purely arithmetical thinking (that deals with particular operations on particular numbers to arrive at a unique solution) and algebraic thinking (that deals with general relationships between undefined numbers). This conceptual development marks the transition between the concrete symbolic and the formal modes of the SOLO model. Thinking in the formal mode allows students to perceive, and hence manipulate, symbolised relationships as mathematical objects according to relationship rules, not to arrive at an unique solution, but to reorganise concepts and develop new patterns.

Because algebra involves the manipulation of objects and the perception of patterns, thinking in the ikonic mode supports algebraic thinking and processes in the formal mode. Although necessary to perceive patterns of symbols, this mode of thinking is not sufficient. The mathematical relationships represented by patterns of algebraic symbols have to be comprehended for thinking to be considered as truly algebraic. To be able to express these relationships, to make them explicit, students need to develop the appropriate language. As students' algebraic thinking evolves, their language for communicating these ideas needs also to evolve, the more effectively to capture and convey their mathematical understanding. This is not merely a matter of acquiring the terminology, but the meanings behind the words. Research indicates that the quality of students' understanding of these meanings may be determined by a study of their use of language.

## RESEARCH QUESTIONS

Research findings and theoretical perspectives from the disparate and independent contexts of linguistics, cognitive theory and the learning of algebra, when taken together, suggest the possibility of a model that explicitly links linguistic features of students' utterances with their mathematical success and conceptual understanding. The broad sweep of the question as to whether the linguistic characteristics of students' verbal explanations and their success in executing mathematical procedures can be compared and also mapped on to the SOLO framework requires that a specific mathematical context be chosen as the vehicle by which the investigation can proceed.

The context chosen is that of certain aspects of introductory algebra. This was chosen for two reasons: introductory algebra (and in particular, elementary algebraic procedures[2]) could be seen to parallel the number understanding of 6- to 10-year olds who were studied by Bills and Gray; and, at a deeper level, algebraic development reflects students' cognitive growth as they first connect aspects of arithmetic knowledge and then make representations and articulate generalisations about arithmetic systems. In SOLO terms, development of algebraic understanding would seem to bridge the concrete symbolic and formal modes. How these aspects might be connected, together with the associated linguistic features, becomes the focus of this research.

---

[2] The term *algebraic procedure* is used by Tall & Thomas (1991) to describe a particular, step-by-step solution regime, whilst the term *algebraic process* is used as a general term for a number of different procedures that lead to the same solution outcome.

The overarching aim of the study is to devise a developmental model that aligns linguistic features of students' explanations with the quality of their algebraic understanding.

Data from exploring the following six research questions will inform the model:

1. What are the levels of student ability (attainment) in algebra?
2. What algebra items present difficulties to students of different abilities?
3. What algebra understandings can be associated with students of different abilities?
4. What are the linguistic features of explanations of students of different abilities?
5. How do linguistic features of student explanations can be associated with the development of algebraic understanding?
6. How might linguistic features and associated algebra understanding be mapped onto a conceptual framework such as the SOLO model?

Research suggests that across a range of mathematical contexts and ages, the ways in which students use language to express their thinking appear to reflect the quality of their understanding. The studies cited have focused on students in primary years, or university entrance. The study reported in the following chapters examines responses from students in the intervening years, the early years of secondary education. These years are often pivotal in establishing students’ attitudes to, and involvement in mathematics.

Establishing deep conceptual understanding is facilitated through encouraging the development of the appropriate register. Conversely, the quality of the language used by students might also indicate the quality of understanding. This study aims to demonstrate how this might appear in the context of introductory algebra. The following chapter discusses the methodology used. Subsequent chapters (4, 5, 6, and 7) describe and analyse the results.

# CHAPTER 3: METHODOLOGY

Assessment of students' mathematical understanding is an ongoing concern for teachers. The judgements made by teachers in their conduct of classes serve to inform their teaching in subtle, and not so subtle, ways. The judgement is often based, in part, on what students say, and how they say it. Whilst it is relatively easy for teachers to focus on the content of students' utterances, students' use of non-mathematical language as well as speech patterns and modality contribute to the information on which teachers make judgements. These aspects might only be implicit. Consequently, justifying "professional judgement" on matters of pedagogy becomes difficult. It is the aim of this research to make explicit aspects of student utterances that might contribute to implicit teacher judgements, and to align those utterances with a model of cognitive development.

Numerical data such as that obtained through testing provide only part of the picture of students' mathematical development. Research, such as that conducted by Stacey and Steinle (2006), reveals that students might obtain correct answers in a test situation but as a result of inappropriate mathematical thinking. Data obtained through interviews, video recordings and observations of students as they go about their work in classrooms provides insight into their thinking.

In this study, the structure of students' explanations and its relationship to their understanding was the primary focus. Therefore, the data collected needed to provide information on the status of the algebraic knowledge of the students involved in the study as well as information that allowed analysis of their use of language, and subsequent insight into their thinking.

The study, consisting of two aspects, was developed to investigate the question as to whether there exist linguistic features of students' explanations that might act as pointers to their conceptual growth, in the context of introductory algebra. These two aspects were, firstly, a test of students' knowledge of introductory symbolic algebra and, secondly, audio-taped interviews with selected students. In this chapter, the context of the study is described in the first section. A second section describes the overall design of the study; the third describes the data collection instruments; the fourth describes the quantitative analysis of the data; and, the fifth section the qualitative analysis. The sixth section describes the tests of significance used. The final section discusses methodological issues.

## CONTEXT OF THE STUDY

Students and teachers from three schools agreed to participate in the study. The following section describes the geographical context; the participants' background and the rationale for their selection; and, the mathematical context in which the study was conducted.

### Geographical

The participating schools were located in a regional city in NSW, population approximately 22 000. The town is an educational centre. It has a university, a TAFE college, eleven primary schools and six high schools within the town limits. Included in these numbers are three independent schools, one Catholic secondary school, and one Catholic primary school. The schools draw students from within the town and surrounding rural towns and villages. The independent schools also draw students from across the state, interstate and overseas.

Many in the town's adult population are employed in the three education sectors: Schools – state, independent and Catholic; TAFE; and, the university. Others are employed in the various services, trades and professions necessary in a large centre, and as rural workers.

### The participants

The participating students were drawn from Years 8 and 9, the second and third years of secondary schooling in NSW. One of the schools was a coeducational secondary school in the Catholic system. The other two were independent, K-12 schools, one an all-girls school, and the other an all-boys school. The schools were chosen for several reasons. They were "convenient" in that they were local to the researcher's base, they were non-selective schools whose student population was drawn from similar demographics, and the school administrators were in a position to make independent decisions regarding the researcher's access to students and teachers. All three schools had participated in previous research projects and were familiar with the requirements of their involvement.

The total number of students answering the survey was 222. Their ages ranged from 13-to-16 years and represented the ability range of the Year 8 and 9 cohorts. In all three schools, mathematics classes in Years 8, 9 and 10 were graded on student performance in class tests and assignments throughout the school year. Two of the schools permitted all students in the Year 8 and 9 groups to participate. The third school decided that only a group of "able" students in Year 9 would be involved. Able students were those students selected by the school as being "most able", and therefore in the class group identified as the "top" class, and taking the Advanced mathematics course.

Table 3.1 details the numbers of male and female students, and the total number of students in each year group for all participants. Year 8 (and Year 7) students are taught the same Stage 4 syllabus content. Year 9 (and Year 10) students nominate whether to be taught one of three different courses at Stage 5 (Board of Studies NSW, 2002a).

*Table 3.1: Survey respondents by gender and course*

| Year Level* | Mathematics Course | Female | Male | Totals |
|---|---|---|---|---|
| 8 | N/A | 71 | 34 | 105 |
| 9 | Advanced | 34 | 35 | 69 |
|  | Intermediate | 21 | 8 | 29 |
|  | Standard | 14 | 5 | 19 |
| Totals |  | 140 | 82 | 222 |

* at time of completing survey

Table 3.2 provides a further break-down of the student numbers according to school year level, mathematics course, sex, and school. The differences in numbers of participants from each school were a function of the size of the school year cohorts. School 1 was by far the largest school, and male and female students were represented in approximately equal numbers.

School 2 was a small girls' school. The 37 students identified made up the entire Year 8 cohort, in two graded classes. The Year 9 cohort of the school was also small. There were two Year 9 classes: one class of 22 was taught the Advanced mathematics course, whilst 17 students were taught the Intermediate or Standard mathematics courses in a combined class. School 3 was a boys-only school, and only one class of students participated.

*Table 3.2: Survey respondents by school and course*

| Year Level* | Mathematics Course | School 1 | | | School 2 | School 3 | Totals |
|---|---|---|---|---|---|---|---|
| | | F | M | Tot | Female only | Male only | |
| 8 | N/A | 34 | 34 | 68 | 37 | 0 | 105 |
| 9 | Advanced | 12 | 11 | 23 | 22 | 24 | 69 |
| | Intermediate | 12 | 8 | 20 | 9 | 0 | 29 |
| | Standard | 6 | 5 | 11 | 8 | 0 | 19 |
| Totals | | 64 | 58 | 122 | 76 | 24 | 222 |

*at time of completing the survey

The consequence of the different numbers of participants from each of the schools, and the differences in numbers from the grades was that no analysis based on school performance, school years, or sex differences was made. Analyses used the aggregated data from all participants ignoring any background differences. The differences between the average raw scores for School 3 and the other two schools were significant ($p<0.05$). However, the participants from these other two schools were from the entire Year cohorts (Years 8 and 9) whereas those from School 3 were from the most able Year 9 class only.

When raw scores for participants from all three schools in the ability range represented by students from School 3 were compared, there was no significant difference in the average raw scores. Combining the results from all three schools was also possible because the schools used common textbooks, the teachers met frequently to share concerns and ideas, and each of the schools had similar teaching programs (some of which had been developed cooperatively).

## Mathematical Context

The particular context in which the algebra used as the basis for the study was situated is that outlined by the *NSW Mathematics Syllabus 7-10* (Board of Studies NSW, 2002a) (the first four years in secondary school) that was followed by all schools in the study. At the time of data collection, the syllabus was undergoing a period of transition. The new Years 7-to-10 Syllabus was rewritten as the second part of a continuum of expected mathematical development from Kindergarten – the first year of compulsory schooling – to Year 10, the final year of compulsory schooling (Board of Studies NSW, 2002a, 2002b). Development of mathematical understanding is described for five content strands of mathematics and one process strand by statements of *outcomes* organised into *stages*.

With the exception of the first stage, which describes mathematical expectations for Kindergarten, each stage covers two years of school grades. The content strands are Number, Measurement, Space and Geometry, Data, and Patterns and Algebra. The process strand is termed Working Mathematically and includes aspects of student questioning, communicating, applying strategies and reasoning. The staged outcomes are further elaborated through *key ideas,* and *indicators*, which describe specific content in terms of expected behaviour by which students might demonstrate a degree of attainment of a particular outcome.

Staged outcomes are written from the perspective of the logical development of mathematical understanding and frame the more specific content indicators. They describe that which typically students can do if they are operating at that stage. Stages are not necessarily aligned with ages, or school grades. In practice, however, this is often the case.

Stage 4 covers the first two years of secondary school in NSW (Years 7 and 8). The mathematics content is the same for all students in these grades. Schools may choose to organise students into classes based on performance at some time during this stage. Stage 5 (Years 9 and 10) however has three levels of outcomes. Historically, those students who attained high levels of achievement in Years 7 and 8 took the “Advanced” mathematics course (content equivalent to that in Stages 5.1, 5.2 and 5.3, Board of Studies NSW, 2002a). Students whose mathematical achievement was poor took the “Standard” course, the content of which is outlined in Stage 5.1. Other students studied the “Intermediate” course, the content of which is that included in Stages 5.1 and 5.2. In each course some of the content was common, some was modified to suit the level, and other content was unique to each course. Included in the study, were students from all levels of attainment (Table 3.2).

The Syllabus from which the students in the study were taught was the earlier version (Board of Secondary Education, 1990; Board of Studies NSW, 1996a, 1996b, 1996c), in which broad outcomes were not made explicit, but in which content was listed, along with suggested teaching activities. These have, for the most part, become the indicators of the new syllabus. In 1999, a separate document detailing outcomes associated with the syllabus content was published for teachers to use in conjunction with the syllabus (Board of Studies NSW, 1999). In 2002, a new syllabus that presented the mathematics in terms of staged outcomes on a continuum from Kindergarten to Year 10 was introduced (Board of Studies NSW, 2002a, 2002b).

Two important changes made to the syllabus (K-10) by the 2002 document are: the incorporation of process outcomes in the Working Mathematically strand into each of the content areas, and the explicit introduction of the topic areas of Data, and Patterns and Algebra in the primary years (K-6).

In the earlier syllabus, the process outcomes for Years 7 and 8 had been included in the 1999 document as a set of outcomes for a "Problem Solving" strand that teachers were expected to incorporate into the teaching of each content strand (Board of Studies NSW, 1999). In the rewritten syllabus for Years 9 and 10 (Board of Studies NSW, 1996), the process outcomes were gathered under the heading of "Working Mathematically", but remained as a set of overarching outcomes rather than integrated with content descriptions. The new syllabus (Board of Studies NSW, 2002a) makes explicit within each of the content outcomes, indicators that address the process outcomes. The present, and new, requirement that teachers of primary school students explicitly address pre-Algebra skills and knowledge may well influence findings from subsequent, similar research.

## DESIGN OF THE STUDY

The study was designed to provide both quantitative and qualitative data on student performance and understanding of algebra. Linked to this was a qualitative analysis of the language used by students as they explained their thinking about items in the survey. The data collection consisted of two parts: a survey (test) of basic algebra followed by interviews with students. A pilot study ($N$=40) was conducted to test the items in the survey together with an interview protocol. Modifications to both the survey and the interview protocol were made following the pilot study (Appendix A).

The main study retained the two-part format for data collection. The first part consisted of a 40-item survey (test) of introductory algebra items for which answers only were required. The survey was marked and students from each of the participating schools were selected for interview, on the basis of raw score results. The interviews were audio-taped and then transcribed for analysis. The interviews constituted the second part of the data collection.

The analysis phase involved Rasch-modelling of the survey responses and qualitative analysis of data from both the survey and the interview. The interviews provided substantive evidence for conclusions drawn from the Rasch model. The results of analysis of linguistic features of the interviews were used, together with the data from the survey to develop a linguistic-based model of conceptual development in early algebra. The design is summarised in Figure 3.1

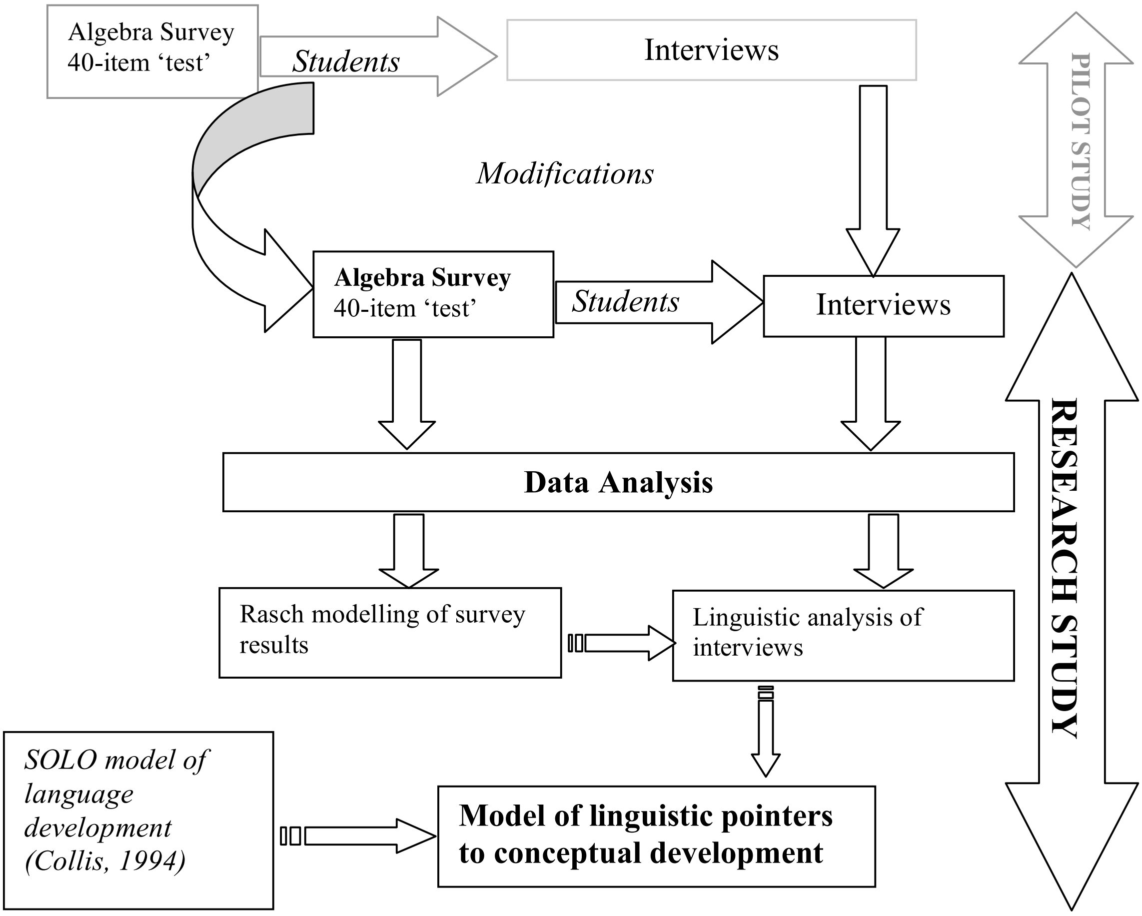


*Figure 3.1: Schematic Overview of the Research Design*

The pilot study provided information about the relevance of the survey items to the students – whether or not students in the target years were, on the whole, familiar with the algebra content. Interviews with students in the pilot were loosely structured around each item in the survey, and responses from the students were used to establish the structure of the final interview protocol

The main research study consisted of data collection from the survey of forty algebra items, which were then supplemented by interview data collected at a time following the administration of the survey. Unlike the pilot study, these interviews were structured around a protocol of sets of items from the survey. The data from the survey were Rasch-modelled; interview data were subjected to an empirical analysis of linguistic features. Both aspects of the data were combined, using the SOLO model as an organising framework, to produce a model describing linguistic pointers to conceptual development.

## DATA COLLECTION

The instruments, the rationale and the background to the development of the data collection instruments are described in this section. Firstly, the survey used is described and discussed. Secondly, the interview protocol is considered.

### The survey

The survey was a written test consisting of 40 algebra items based on examples in the *NSW Mathematics Syllabus, 7 – 10,* Stage 4 (Board of Studies NSW, 2002a), from textbooks used by the participant schools, as well as items from Küchemann's (1981) study. These last sets of items were chosen because they were also typical of the Stage 4 items chosen from the NSW Syllabus, and because they had well-documented research findings accompanying them (Appendix B).

Stage 4 Outcomes from *NSW Mathematics Syllabus, 7 – 10* state that at the end of the stage a student:

- "uses letters to represent numbers and translates between words and algebraic symbols" (PAS4.1);
- "creates, records, analyses and generalises number patterns using words and

algebraic symbols in a variety of ways" (PAS4.2);

- "uses the algebraic symbol system to simplify, expand and factorise simple algebraic expressions" (PAS4.3); and,
- "uses algebraic techniques to solve linear equations…" (PAS4.4) (Board of Studies NSW, 2002a, pp 82- 86).

Items involving manipulation of expressions comprised the first 25 questions, the remaining 15 questions required students to solve simple linear equations. The items were not necessarily ordered according to difficulty.

The responses by students in the pilot study to questions in the survey resulted in some minor changes, but confirmed that the items included in the survey were appropriate to the stage and experience of the target groups in the main study. The minor changes are noted in the summary of the pilot study (see Appendix A: Pilot Study).

### *Rationale for the survey*

The survey was developed to provide data on the algebraic knowledge of the students as it related to their ability to manipulate elementary algebraic expressions and equations. There existed no external, system-wide tests (such as SAT) that could be used to determine the achievement levels of individual students across a set of common measures.

Although each of the schools involved in the study graded their students on achievement, this was measured in various ways, and represented a holistic assessment of student performance over a period of time. Individual schools devised their own assessment regimes to measure student achievement. The assessments might include assignments, projects and tests. The tests, however, need not target algebra specifically, and would not be common across the participating schools.

### *Participants*

The total number of students who took the survey was 222. In two of the schools, the survey was conducted with all classes in Years 8 and 9. In one school, only those students in the top Year 9 class participated. (For a summary of the participants by school, grade, mathematics course and sex, see Table 3.1.)

The students' usual classroom teachers administered the survey so that it was situated in a setting familiar to the participants. Such tests were not uncommon in any of the three schools. The survey was to be given at the teachers' convenience, in the time of a normal lesson, approximately 50 to 60 minutes. There was no time set for the completion of the survey, and calculators were not to be used.

### *Constraints*

Factors that might have contributed to data corruption include the asynchronous administration of the survey in the three schools, and the absence of some of the Year cohort. To avoid corruption of the data, the surveys ideally needed to be administered at the same time and under the same conditions. As three separate schools were involved, this was not possible. However, it is unlikely that any communication between teachers, or students, would significantly compromise the results, particularly as it was emphasised that the survey was exploratory, and not for purposes of comparison.

It was not possible to insist that the surveys be administered on the same day, or even the same week, in each of the schools. In two of the three schools, the surveys were administered near the end of the fourth term. In the third school, the survey was conducted the following year, in the last weeks of the first term. The time taken to obtain the various permissions needed (see the section on ethics) meant that the survey was in the schools in the last term of the year. This is a particularly busy time, and a disrupted time. Before school examinations, teachers were loth to take up valuable class time. After the examination period, many students were involved in activities that took them out of the classroom. Hence, there were some small numbers of participants from classes that were already small in number. In one of the participating schools, the survey was postponed until the following year, and then only one class could be involved.

The results from each school were aggregated. No analysis of the performances by individual schools was undertaken. Although the school teaching programs were similar, based as they were on the Stages of the NSW Mathematics Syllabus, the timing of the teaching of the different topics is a matter for the schools. Consequently, there might have been some influence from recent teaching on some results. (One student noted this fact during an interview.) By treating the data set as one sample, individual differences between schools, and between students would not unduly influence the patterns of behaviour sought through analysis of the data. To analyse individual schools' performance, or student responses, there also needed to be detailed background data collected. This was not possible, given time constraints and the extent to which schools were willing to be involved in the study. The survey data were used principally to establish ability (achievement) rankings that were used to select students for interviews. The ability groupings were also used as the bases for comparison of the linguistic features of the students' explanations.

## The interviews

Interviews were conducted after the administration and marking of the survey. Because the survey was to be administered by schools in Term 4, the interviews could not be conducted until the following year. This was not considered a problem with the integrity of the data as the survey was designed solely to identify ability ranges of students for selection for interview. A detailed description of the students, their mathematical background and sex, and details of the interview protocol are described in the following.

### *Participants*

The participants ($n$ = 31) were purposively selected on the basis of their results (in raw scores) on the survey items. The selected students represented a small subset of the total student population, across a range of results. Table 3.3 provides an overview of the student population by school year level, mathematics course, sex and total numbers of students in each year group.

The interviews were conducted several months after the survey was taken, and in the case of two schools, this meant that the students had moved into the next Year. (i.e., those students who did the survey in Year 8 were, when interviewed, in Year 9, and those in Year 9 when they did the survey were in Year 10 when interviewed.) Students in the third school were interviewed in the school term after their having sat the survey.

*Table 3.3: Participants interviewed*

| Year Level | Mathematics Course | Female | Male | Totals |
|---|---|---|---|---|
| 9 | Advanced | 4 | 9 | 13 |
|  | Intermediate | 6 | 1 | 7 |
|  | Standard | 1 | 1 | 2 |
| 10 | Advanced | 4 | 4 | 8 |
|  | Intermediate | 1 | 0 | 1 |
| Totals |  | 16 | 15 | 31 |

Students were selected for interview to represent the range of raw scores on the survey. The students' results on the survey were ordered from those who correctly answered all questions to those who scored zero correct responses. Only those students who attempted all or most of the questions were considered for interview. It had been found from the pilot study (Appendix A) that students who did not make an attempt to answer all, or most, questions were unable or unwilling to contribute substantive data in an interview. Table 3.4 provides details of the participants at the time of their being interviewed.

*Table 3.4: Participants interviewed by school and course*

| Year Level | Mathematics Course | School 1 |  |  | School 2 Female only | School 3 Male only | Totals |
|---|---|---|---|---|---|---|---|
|  |  | F | M | Tot |  |  |  |
| 9 | Advanced | 1 | 1 | 2 | 3 | 8 | 13 |
|  | Intermediate | 3 | 1 | 4 | 3 | 0 | 7 |
|  | Standard | 1 | 1 | 2 | 0 | 0 | 2 |
| 10 | Advanced | 4 | 4 | 8 | 0 | 0 | 8 |
|  | Intermediate |  | 0 |  | 1 | 0 | 1 |
| Totals |  |  | 16 |  | 7 | 8 | 31 |

Given the purpose of the interviews it was important that students be able to communicate effectively (Ericsson & Simon, 1980, p.12). Based on the results of the survey, 50 students were identified for interview. Of these, 31 students agreed to be interviewed. The students who were interviewed were those who were willing to be interviewed, were available at times convenient to the school and the researcher, and who had appropriate caregiver permission.

Approximately equal numbers of male and female students were interviewed, but there was no selection for samples that were representative of the proportions of students in ability groups in the school, or of gender across the sample. However, in each school, a range of abilities was represented in the interview sample. The whole sample of those interviewed was also representative of a range of abilities (with the exception of those with very low scores described above).

### *The Interview Protocol*

The interview consisted of asking each student to respond to nine sets of questions (Appendix B). The first eight sets consisted of items from the survey used to elicit students' explanations of their algebraic thinking. The ninth set required students to respond to questions of a more general nature. Some of these ninth set questions retained a mathematical focus; others were non-mathematical. The questions in this set were adapted from those used by Bills (2001) and sought to establish some background linguistic characteristics of each student.

Interviews were *contingent* (Rowland, 2000). The students were directed to discuss one set of items at a time, by the question, "Tell me what goes on in your head when you see/deal with items such as these [indicating the items in one of the sets]". The interviewer subsequently could ask other questions of the student being interviewed in order to elicit further information. The interviewer used information-neutral probes such as, "Can you tell me more?" or more directed prompts such as, "Can you see anything common in this set [of expressions or equations]?"

Interviews were conducted with individual students, audio-taped and transcribed in full. Each interview lasted from 30-to-60 minutes, depending on the extent of the student responses.

Although the interviews included set questions, they were conducted as conversations, rather than as quizzes to test students' knowledge. The researcher tried to make the students feel comfortable, pointing out that they could leave at any time, that the research was to discover ways that might help them understand algebra better. The interviews were introduced by the researcher telling the students something about themselves and the research, and asking some personal (but irrelevant to the study) details about the students.

Items on the survey were used to frame the interview questions in the eight sets that focused on student's algebra understanding. The sets of items are listed in Figure 3.2.

The items were arranged in the eight sets on the basis of syllabus topics, as well as, where appropriate, superficial or procedural similarity. For example, Set 8 consisted of items from the survey that all had brackets, and elements in Set 1 were all items requiring simplification by the addition or subtraction of like terms.

| | | |
|---|---|---|
| Set 1. Simplify:<br>$3m + 8 + 2m - 5$<br>$5p - p + 1$<br>$2ab + 3b + ab$<br>$5a - 2b + 3a + 3b$ | Set 2. Simplify:<br>$4 \times 5b$<br>$2ab \times a$<br>$4r \times 5t \times 3$ | Set 3. Simplify:<br>$2(x + 5)$<br>$2(x + 4) + 3(x - 1)$<br>$2\ (x + 5) + 8$ |
| Set 4a. Simplify:<br>$\frac{a}{5} + \frac{a}{10}$<br>$\frac{3p}{4} - \frac{p}{8}$ | Set 4b. Simplify:<br>$\frac{4ab}{4b}$<br>$\frac{2}{a} \times \frac{3}{b}$<br>$\frac{2}{a^2} \times \frac{5a}{4}$ | Set 5. Solve:<br>$x + 5 = 7$<br>$2t - 23 = 49$<br>$5a - 4 = 2a + 8$<br>$x + (x + 2) = (x - 1) + 8$<br>$4(p + 3) = 32$<br>$4y = 20$<br>$10y = 5$<br>$ax = 5$ |
| Set 6. Solve:<br>$\frac{x}{4} = 12$<br>$\frac{x+3}{2} = 7$<br>$\frac{63}{x} = 180$ | Set 7. Write without brackets:<br>$(6xy)^2$<br>$(x + y)^2$<br>$(a - b) + b$<br>$8p - 2(p + 5)$ | Set 8. Read these statements out loud, and then tell how you could rewrite them:<br>Add 4 onto $n + 5$<br>Add 3 on to $4n$<br>Take $n$ away from $3n + 1$<br>If $p + q = 5$, then $p + q + r = ?$<br>If $e + f = 8$, then $e + f - g = ?$ |

*Figure 3.2: Questions presented in the first 8 subsets of the interviews*

Ericsson and Simon (1980) discussed in detail the constraints and threats to validity in the use of verbal data to establish cognitive frameworks possessed by individuals about particular concepts/tasks. Although this was not the intent of the interviews in this situation, Ericsson and Simon made the point that when subjects are asked to "think about [… a] perceptual-motor process" (the authors' terminology) rather than execute the process, (they use the example of subjects completing the Tower of Hanoi, p.231), they verbalise their thoughts more readily than if they both carry out the task and verbalise their thinking. By requiring students to think about the algebra in broadly general terms, rather than instruct them to carry out the procedures, their verbalising of their mental processes could, therefore, be more complete, and demonstrate an understanding of algebraic principles. A conversational approach was hoped to foster dialogue that was reflective in nature, and which would provide sufficient data for an empirical approach to qualitative analysis to be used.

## QUANTITATIVE ANALYSIS OF RESEARCH DATA

Quantitative data were used to establish the overall algebra achievement by the participants, on a set of common criteria. Quantitative analysis consisted of the marking of the survey responses, Rasch modelling of the responses, and an analysis of errors. These aspects are described in the following.

### Marking for correct/incorrect responses

The surveys were collected and the items marked as *correct* (1), *incorrect* or *no attempt*. For the Rasch model, both incorrect responses and non-attempts were coded as being incorrect (0). By coding non-attempts as incorrect, the assumption was made that absence of an answer indicated the student did not know how to approach the item, or did not want to commit to an answer. Some responses were crossed out, with no substitute response given. These were also counted as incorrect (non-attempts). Thus, incorrect answers, and no answer were both considered as indicative of conceptual ignorance. Incorrect responses and non-attempts were distinguished for the purposes of the error analysis. Preliminary marking also enabled the identification of students who were likely to be confident to make an attempt at a question, even if incorrect, and hence able to contribute in an interview.

Correct responses (1) were those which were complete (e.g., fractions expressed in lowest terms, all like terms collected, equations completely solved). Some items in particular attracted a variety of partially correct responses, but as students were not required to show working, and many did not, coding these on a spectrum of "correctness" would have required some "second guessing" about the intention of the students.

## Rasch modelling of the test scores

The Rasch model has as its basis Item-Response Theory (IRT), which proposes "the probability of a person's expected response to an item is the joint function of that person's ability, or location on the latent trait, and one or more parameters characterising the item" (Bond & Fox, 2007, p.311). A latent trait is "a characteristic or attribute of a person that can be inferred from the observation of the person's behaviour" (p.311). The probability of an individual's success on an item is calculated from the difference in the ability of the individual and the degree of difficulty of an item.

The Rasch model represents ability levels and item difficulty as a linear progression of development of a single construct *(unidimensional*). Where different aspects of a single construct are used to represent that construct (as with items requiring various algebraic understandings) those aspects need to operate in a connected manner for the construct to be considered as unidimensional.

Thus, relationships between two independent variables (item difficulty and student ability) can be represented on an equal interval scale in such a way that the distances between what are initially ordinal values can be made equal and meaningful (Bond & Fox, 2001). The model is sensitive to the developmental order of the skills or abilities under investigation, and provides numerical estimates of the developmental differences between the ordered skills or persons in the population being studied. This allows the verification of general developmental patterns.

The Rasch model enables the validity of the test instrument to be established by demonstrating the construct reliability and the fit of items to the model. These aspects are discussed in the following two sub-sections.

### *Construct reliability*

In this study, the construct to be analysed was that of basic (i.e., early secondary school) symbolic algebra, the data coded as correct or incorrect responses to the questions. The participant ability rating and the item difficulty estimates are calculated using the natural logarithm of the odds of a successful response to the item (a *logit*). The number is arrived at through an iterative process using matrices of frequency of correct /incorrect responses to individual items against frequencies of correct/incorrect responses from each participant. The calculation takes the ordinal value of the percentage number of correct responses and maps this onto a log-linear scale where equal intervals have equal values. In this way differences in abilities of participants and differences in item difficulties are represented by interval differences. The sizes of the intervals are proportionally significant.

As with all statistical inferences, the greater the amount of available data, the less the error. Thus, the precision of the estimates of individual ability and item difficulty depends on the data available at a particular point. If insufficient items are found to be at the ability level of a participant, then that ability estimate is subject to greater error than one where there are a number of items of a difficulty equal to that ability level. This means that the extremes of the ability levels or item difficulty ranges may be subject to a greater error if the test does not include items that will provide the information. In effect, this means that any test should have no items that allow all participants to give a correct response, and all items should allow at least one correct response.

Where the Rasch model differs from others is the requirement that the data fit the model, rather than the model fit the data. Consequently, construct reliability (validity) of the test instrument is demonstrated by the extent to which both items and persons (cases) fit the model. Construct reliability is dependent on the reliability of the estimates of item difficulty and the ability levels. The mean for these estimates is set, at a default of 0.00. Reliability is measured on a 0 to 1 scale. Thus, the nearer the reliability is to 1, the more confidence can be placed in the replicability of the item placement or ability estimates if the test were given to other suitable samples.

The Rasch model is based on three assumptions: the construct to be measured is unidimensional; the increase in item difficulty represents an increase in construct “quantity” (Wright & Masters, 1982); and, that responses to later items were independent of responses to earlier items (Wright & Stone, 1979).

### *Measures of fit*

Fit statistics are used to detect discrepancies between the model and the data collected. The difference between the expected and observed values results in a “score residual”. By standardising the score residuals, a comparative measure of how each item deviates from the expected value can be arrived at.

There are two types of fit described. *Infit* statistics are calculated as weighted mean square values of the standardised residuals. These statistics are sensitive to patterns of targeted responses, and so provide a good measure of item discrimination. When standardised, the Infit Mean Square, approximates a *t*-distribution, and provides a comparative measure that, when items fit the model, has a mean near to zero and a standard deviation close to one (Wright & Masters, 1982). Infit Mean Square values and *t*-values obtained from the Quest program (Adams & Khoo, 1996) are used as indicators of fit. Where the Infit Mean Square values lie between 0.77 and 1.3, the fit values are taken to indicate reasonable fit of the data to the model.

Unweighted mean square values of the residuals provide *outfit* values. These statistics are sensitive to outliers, “responses to items with difficulty far from a person, and vice versa” (Linacre, 2002, p.878), and provide less information about item discrimination. Quest (Adams & Khoo, 1996) software reports these measures as Outfit Mean Squares and Outfit *t*.

Infit Mean Square values greater than 1.3 indicate that there is more variation than expected between the model and the data. This suggests that fewer then expected students with an ability estimate greater than the difficulty estimate for the item answered the item correctly, whilst more than expected students with an ability estimate less than the difficulty estimate for the item were able to answer the item correctly. In these cases, it might be that constructs other than those of algebraic understandings (in this case) are operating.

Fit values less than 0.7 indicate *overfit*, where patterns of person responses to an item most nearly match Guttmann patterns. This suggests a higher-than-probable degree of discrimination.

Where clusters of item difficulty and student abilities were identified from the Rasch model, differences between the means of the clusters were tested for significance using *t*-tests.

## Further analysis of the survey data

Although the "observable behaviours [from the Rasch model] display more or less the characteristic, […] none of the observation covers all of the [latent] trait" (Bond & Fox, 2007, p.311). Rasch modelling allows the identification of clusters of items and clusters of persons on the same variable (construct). Questions remain, however, as to the nature of these differences, and the nature of the similarities that exist between items (and persons) in the clusters. To answer these questions, data needed to be analysed by considering the items in categories based on mathematical commonalities, and also by examination of errors made.

Mathematical commonalities shared by items were identified from syllabus topics taken from the *NSW Mathematic Syllabus 7 -10* (Stages 4 and 5, Board of Studies NSW, 2002a). These were: addition and subtraction of like terms; multiplication of terms; expressions or equations containing brackets; expressions or equations containing fractions; and, items that couched algebraic concepts in words, or which were equations with no unique solution. These last included items from Küchemann's study (1981), and equations such as "$ax = 5$, find $x$". The categories included both expressions and equations. Several items could have been allocated to alternative groups. Syllabus outcomes were not fully represented by items in the survey. The more advanced topics such as the manipulation of quadratic expressions and the solution of quadratic equations were not included as representative of "early" algebraic concepts. Table 3.5 lists the syllabus outcomes, the categories of items and syllabus outcomes references.

Item categories and syllabus outcomes (or key ideas) are not, by the nature of the items, mutually exclusive. The categories were decided on the basis of the intent behind the inclusion of the items (e.g., no category was given exclusively to the syllabus outcomes of simplifying expressions with indices, although students' ability to use index notation was needed for successful responses to several items.)

*Table 3.5: Algebraic concepts, syllabus references and categories of items*

| Algebraic concepts | Syllabus Outcomes/Key ideas | Syllabus reference | Item Category |
|---|---|---|---|
| Algebraic notation and conventions | Using letters | | |
| | Identifying equivalent expressions | PAS4.1 | Semi-literal+ all others |
| | Index notation | PAS5.1.1 | |
| Manipulation of Expressions | Simplifying expressions by: | | |
| | - Adding & subtracting | PAS4.3 | Addition/Subtraction |
| | - Multiplying & Dividing | PAS5.3.1 | Multiplication |
| | Simplifying expressions by: | | |
| | - Adding & subtracting fractions | PAS4.3 | Fractions |
| | - Multiplying & Dividing fraction | PAS5.2.1<br>PAS4.3 | |
| | Simplifying expressions with brackets | PAS5.2.1<br>PAS5.3.1 | Brackets & Semi-literal |
| | Simplifying expressions with indices | PAS5.2.1 | Multiplication, brackets and fractions |
| Solution of Equations | Solving by adding, subtracting like terms in several steps (up to 3) | PAS4.4<br>PAS5.3.2 | Addition/Subtraction and/or Multiplication |
| | Solving equations with fractions | PAS5.2.2<br>PAS5.3.2 | Fractions and/or Multiplication |
| | Solving equations with brackets | PAS5.2.2<br>PAS5.3.2 | Brackets |

The means of each of the categories were calculated and the differences in means tested for significance using *t*-tests. Errors made by students in their responses to items in the survey were also analysed. The numbers of non-attempts, incorrect responses and correct responses on the survey items were recorded for all participants. Different incorrect responses were noted and counted.

The quantitative analysis of the survey data provided an organisational basis for the qualitative aspects in the study. Because the study aimed to identify an association between student ability and item difficulty and the linguistic features of their explanations, groups of students, and groups of items needed to be established by objective, rather than arbitrary, means. The groups of student ability and item difficulty were resolved by the Rasch model, and subsequently used when examining data from the interviews.

## QUALITATIVE ANALYSIS OF DATA

Interviews provided two sources of data: that of the mathematical content knowledge of students, and that of linguistic structures indicative of mathematical understanding. Interview data were collected primarily for analysis of the linguistic features of the students' explanations about their thinking. However, if such explanations were to be linked to a developmental model, the mathematical content of those utterances also had to be considered. In this section, interview data are described, firstly, as it was used to explain similarities and differences in item categories, and student ability groups, in terms of mathematical content, and, secondly, as linguistic features were identified and examined.

### Mathematical content analysis

Quantitative data about the difficulties of categories of items, and error patterns were substantiated by data from the interviews. These data were used to identify characteristics of clusters of items, by using the explanations of students about the items to which they had responded successfully. Taken together, the data provided a set of descriptors for each cluster of items that posited conceptual changes necessary for students most likely to be able to respond successfully to all items in a cluster; and hence, define characteristics of ability groups.

A *least strategy* was used as the defining characteristic for items in a cluster. (e.g., the least strategy used successfully to solve the equation $x + 5 = 7$ is a "count on by ones" strategy. To solve more complex equations, a least strategy would be a trial-and-error substitution.) Such strategies need not be mathematically correct or useful, even if resulting in a correct response. Clusters of items, although containing items of different types (categories) were then broadly described in terms of the concepts and procedures used by students to answer the items successfully.

Adherence to unsophisticated strategies, which might have been successfully applied to the simpler items, usually meant that students were unable to deal with items of greater difficulty. More able students, although perhaps using unsophisticated strategies to respond to simpler items, would need to access other, more mathematically useful strategies to succeed on more difficult items. Comparing students' explanations to the errors made helped define the limits of such strategies, and therefore indicated points of conceptual development that described the ability groups. Because of the nature of the Rasch model, which aligns ability with item difficulty, the descriptions of the mathematical features and strategies for dealing with item clusters become also descriptions of the conceptual development of the ability group aligned with that cluster.

## Linguistic analysis

The possible relationship between the ways in which students expressed their understanding of the mathematical ideas and the extent of that understanding (success on the survey) was investigated by examining their interview responses. The responses were examined for the following linguistic features:

- Verbosity, the number of words contributed by students during their explanations;
- Use of pronouns *I*, *you*, and *it* (plural forms are included with the singular);
- Type of response, whether *particular*, where the respondent focused on one item, or several discrete items, or *general* when the student provided a rule or a procedure described in non-specific terms;
- Modality as determined by the use of *hedges* and *shields*[3], by pauses, or by false starts;
- Tense: either past, indicating a personal recounting of what the student did or thought; or, simple present, indicating a general statement of a rule or principle;

To avoid individual idiosyncrasies, such as a tendency to prolix or laconic responses, that might mask emergent, more general patterns, linguistic features were analysed in the ability groups identified from the Rasch model. The particular features described above did not always constitute the greater part of individual utterances, but patterns could emerge when the features were considered collectively from all students.

Although the students interviewed represented a range of ability, there were different numbers of students from each ability group[4]. So that comparisons between the groups could be made, two methods were used: raw counts were expressed as an average frequency of a particular linguistic feature per response per group; or, raw counts were expressed as an average frequency of a particular linguistic feature per student per group;

Frequency of occurrence of targeted linguistic features was examined with respect to the different ability groups, and with respect to the difficulty of item sets.

---

[3] The terms *particular, general, hedges* and *shields* are defined in Chapter 2.

[4] The term *ability group* used to designate student attainment on the algebra survey. Retains the terminology used in the literature on Rasch modelling.

Not all sets of items were used for every aspect of analysis. Only those sets for which there was sufficient dialogue to be analysed were used. For example, Set 8, which consisted of items from Küchemann's study (1981), asked students to read the worded problem aloud, and then suggest how it might be rewritten. As a result, many of the features of a freer response were not evident in responses to this set.

One further example is the in-depth analysis of the responses to the sets of equations (Set 5 and Set 6) for the occurrence of the use of the pronouns *I*, *you* or *it* as items became more difficult. Most students, regardless of ability, responded to all, or most items, in these two sets. The sets contained items that challenged students across the range of ability. Therefore data were available across a range of item difficulty, and a range of student ability. This was not always the case for other sets of items.

It could be assumed that fewer students in an ability group would mean that fewer words would be spoken, and *vice versa*. Consequently, any linguistic feature (such as counts of pronouns, or modal indicators) would be proportionately fewer, or more.

To investigate whether changes in the frequencies of linguistic features studied were influenced by the difficulties of types of items, regardless of ability of students, ability groupings were disregarded.

Significance of the differences was tested by using Chi-squared tests. Where differences were not shown to be statistically significant, but patterns of linguistic behaviour were discernable, these were also noted. Results have been illustrated graphically to show the patterns of observed variation in linguistic features.

Verbatim evidence was used to substantiate the findings. This was particularly important when the raw counts of linguistic features were too small, or statistically insignificant, but where, nevertheless, patterns of behaviour could be identified. The verbatim evidence was also used to complement the Rasch model of the survey results and to develop descriptors used in the final model linking linguistic features with conceptual development.

## STATISTICAL ANALYSIS

This section outlines tests of statistical significance used in the study. Tests of significance used were students' *t*- test and chi-squared tests.

### Students' t-test

Students' *t*-tests were used to calculate the significance of the differences between means of ability groups, item clusters (and sub-clusters), and item categories. *t*-values were calculated in an Excel spreadsheet using the formula:

$$t = (m_1 - m_2)/\text{sd means pop}$$

where

"$m_1$" and "$m_2$" represent the means for each of the ability or item clusters;

"sd means pop" represents the standard deviation of the means of each of the clusters and is calculated as

$\sqrt{}$ [(sum of squared deviations$_1$ + sum of squared deviations$_2$)/($n_1 + n_2$ – 2)×($1/n_1+1/n_2$)];

"$n_1$" and "$n_2$" represent the numbers in the populations of which the means are compared;

"deviation$_1$", and "deviation$_2$" are the deviations of the populations being compared.

*p*-values were calculated using Excel software for Mac OSX, and differences were considered to be significant at $p$=0.05 or less.

### Chi-squared tests of significance

Chi-squared tests were used to determine the significance of different frequencies of occurrence of linguistic features. The occurrence of a targeted linguistic feature (e.g., use of the pronoun *I*) was counted for each of the three ability groups, and also for each of the sets of items presented to students during the interview. Observed frequencies were compared with the expected frequencies. The null hypothesis used was that ability, or set difficulty, should make no difference to the frequency of occurrence of the targeted linguistic feature.

*p*-values for Chi-squared tests were calculated using the Excel software for Mac OSX. Differences between actual and expected measures were considered to be significant at $p$=0.05 or less.

$\chi^2$ -values were not calculated using this program.

## EVALUATION OF THE METHODOLOGY

The purpose of the research described in this study is to discover if there exists a relationship between features of students' utterances and the level of their understanding of mathematics – in this case, that of elementary algebra. Consequently, there are both qualitative and quantitative aspects to the research design. With all research, there is the question of whether data would be obtained in other situations by other researchers applying the same methodology and whether the data so obtained suits the aims of the investigation. These are the issues of reliability and validity. This section discusses the methodology from the perspectives of the qualitative and quantitative paradigms, and with regard to the issues of reliability and validity.

### Qualitative and Quantitative Research Design and Analysis

Quantitative research makes measures or counts of "things". It is essentially deductive, emphasising facts, relationships and causes. Quantitative research often takes place in contrived (not natural) settings. There is a focus on individual variables, analyses are statistical, and results are usually numerical. Qualitative data, on the other hand, are holistic in character, drawing general inferences from specific instances and collecting data in natural settings. Qualitative data are essentially descriptive of phenomena that does not lend itself to enumeration (Wiersma, 2002, p.12).

Leedy (1993) stated, somewhat categorically, that "…if the data is verbal, the methodology is qualitative, if it is numerical, the methodology is quantitative" (p.139). He then contrasted the quantitative and qualitative methodologies as respectively: (i) having an outsider versus an insider perspective; (ii) describing a stable rather than a dynamic reality; (iii) focusing on particular factors versus having a holistic focus; (iv) having an orientation towards verification compared to having an orientation towards discovery; (v) the collection and use of objective data versus the collection and use of subjective data, under controlled rather than naturalistic conditions. Quantitative research focuses on the reliability (replicability) of the data whilst qualitative methods are more concerned with the validity (representative truth) of the data (Leedy, 1993, p.144).

However neither methodology is exclusive of the other, and statistical (i.e., numerical) methods can be used to verify qualitative data. The approach of combining both methods is termed "triangulation", of which Leedy (1993, p.143) identifies four types: theoretical triangulation; data triangulation; investigator triangulation; and methodological triangulation. In this study, data triangulation and methodological triangulation are combined.

*Data triangulation* can be achieved by collecting data through two sampling strategies. In this study, a survey (test) of algebra items provided quantitative data on which to base the selection of students to interview. The interviews focused on items in the survey, providing data about the process used by students in responding to the survey items. The data from the interviews was used to explain the item difficulty patterns that resulted from the Rasch modelling of the survey responses.

*Methodological triangulation* was achieved by Rasch-modelling the survey responses, and then using the results of the model to organise and analyse interview data. Because student ability and item difficulty are organised on the same equal-interval scale in a Rasch model, predictions could be made about how students of a particular ability were likely to respond to items of a particular difficulty. The mathematical content, and the linguistic features, of students' interview responses can therefore be linked with objectively measured ability groups, and items.

The study of linguistic features and mathematical understanding adopted an empirical paradigm. Targeted aspects were counted, possible significant differences in counts across the groups under consideration were validated statistically, and verbatim reports from interviews were used as substantiating evidence.

## Validity in Research Design and Analysis

Valid research is "capable of being justified" (Wiersma, 2000, p.4). There are two types of validity: internal validity and external validity. Internal validity refers to the extent to which the results can be accurately interpreted. External validity is the extent to which results may be generalised to populations or conditions, and presupposes internal validity. However, there is a tension operating between these two aspects. Internal validity is context specific, and in educational settings it is difficult to replicate particular contexts exactly in other environments.

Internal validity may be threatened by counter changes occurring over the lifetime of the project, effects of selection of subjects, effects of the presence of an unfamiliar observer, effects of attrition of possible subjects and the problem of drawing incorrect conclusions from the data. These threats to validity and their effect on the research are considered below.

### *Validity and the Rasch model*

It is a requirement of the Rasch model that the data fit the model. The model demands that a single construct be measured. Where items, or cases (persons) do not fit the model, these need to be investigated. Measures of fit, described in a previous section, are used to test the fit of both items and cases. How reliably data fit the model is a measure of the unidimensionality of the construct being investigated, and of the validity of the measurement tool (the survey).

The objective of the linguistic analysis of the interviews was to establish that students of different ability used the language in discernibly different ways when explaining their mathematical thinking. Targeted linguistic features were analysed in ability groups identified from the Rasch model. The average ability estimates for each group were shown to be significantly different. Thus, significantly different patterns of linguistic changes might be validly associated with different ability groups. Similarly, where average difficulties of sets of items, categories of items, or clusters of items could be shown to be significantly different, changes in linguistic patterns associated with those sets of items can be associated with conceptual differences.

Mathematical content data from the interviews (*what* the students said, rather than *how* they said it) was used to explain the clustering of items in the Rasch model. By describing the characteristics of the items in each of the clusters, professional judgement and familiarity with the subject matter and the intent of the inclusion of particular items was also necessary.

### *Changes occurring over the lifetime of the project*

The survey was administered several months before the interviews took place, and in two of the schools at approximately the same time. However, in these two schools the survey was not administered at the exact same time. There was opportunity for students to discuss the questions; this could not be avoided. Given the nature of the questions, and the fact that the teachers conducting the class emphasised that the survey was not a scored test students were, as a large group, unlikely to discuss the survey questions in detail.

More of a problem was the possibility of students copying answers in the class situation. This was not considered a significant threat to the validity of the study because the test was analysed using the Rasch model, used principally to frame the interviews, and finally to select students for interview. The situation of two or three students giving the same response to an item would not be significant in 222 responses.

Responses to the survey could have been influenced by the proximity of the topic in the learning sequence devised by individual schools. Some students might have recently completed the topic; in other cases, a topic might have been addressed in a period well before the survey. However, given that the main focus was data from the interviews, this fact often emerged from the interviews and could then be accounted for in the discussion of results.

Attrition of subjects was not a concern as the survey was administered to several students ($N$=222) and only a small sample ($n$=31), representative of the ability range, from that population was interviewed. The interviews were conducted in each school over two or three consecutive days. No follow-up interviews were conducted.

### *Selection of subjects for participation*

Students were purposively selected for the interviews based on the survey results, rather than as representing a sample of the graded classes. School grading systems tend to be based on a range of considerations, some of which have to do with organisation constraints within each school, rather than solely the educational achievement of the students in each year level. Hence the need for an objective measure of student ability when dealing with algebra.

A range of abilities as measured from the survey results was represented. Although more males than females responded to the survey, nearly equal numbers of males and females were interviewed. This was, as it eventuated, beyond the researcher's control. The students who were finally interviewed were those who consented to participate and who had the permission of their parents, and the indulgence of their teachers. The concern was to obtain a range of abilities, regardless of the sex of the participants.

### *Effects of the presence of an unfamiliar observer*

The effect of a "non-natural" setting – one-on-one taped interviews with an unfamiliar observer – was reduced by setting the students at ease, explaining the purpose of the interview and asking if they minded being taped. As White and Gunstone (1992) found, most participants are cooperative in these situations. The students who were interviewed were volunteers from those selected.

## Reliability in research design and analysis

Reliability refers to the "consistency of research and the extent to which studies can be replicated". There are two types of reliability – *internal reliability* and *external reliability*.

Internal reliability is "the extent to which data collection, analysis and interpretations are consistent, given the same conditions" and external reliability considers "whether or not independent researchers can replicate studies in the same or similar settings" (Wiersma, 2002, p.8). This is of particular importance when interview data are being collected and analysed. The Rasch model provides its own evidence of reliability, and this has been discussed in a previous section. How the research design addresses these two aspects of reliability with respect to the interviews and the analysis of the interview data are described in the following.

### *Internal reliability*

Internal reliability considers whether or not the study is consistent. To ensure consistency in the methods used to collect and analyse data, students' survey responses were written down and the interviews audio-taped and transcribed. The one interviewer (the researcher) conducted all interviews according to the protocol (Appendix B), although the classroom teachers conducted the survey under usual classroom test conditions.

Taping of the interviews and later transcribing the tapes preserved the raw data. The presence of the tape recorder might have threatened the validity of a "naturalistic setting", but was essential for the assurance of reliability.

### *External reliability*

External reliability considers whether or not the study is replicable (Weirsma, 2002, p.8). Educational contexts vary, due to many factors that cannot be controlled by researchers. This study was conducted in one particular, regional city, involving three schools of similar demographics. In order to ensure that the study could be replicated the participants and the context have been described as fully as confidentiality permits. The data-collecting instruments have been described and included in the appendices (Appendix B) and the data-analysis methods described.

The responses on the survey were a snapshot of the student responses at a particular time and might have been influenced by a number of factors – the motivation of the students, the freshness of the topic in the survey, for example. The interviews served to clarify some of the issues, but the students interviewed were a small subset of the sample of students. However, as different instruments were used to collect and analyse the data, some of these issues have been resolved.

## Ethical Considerations

In conducting educational research, care must be taken to protect the interests of the participants involved and the institution represented by the researcher. Ethical issues that need to be considered are: informed consent; provision of sufficient information to the participants; the use of volunteers; and the right to privacy.

Informed consent requires that the participants give approval to be involved before being included in the research. This was obtained from all participants before they took part in the interviews. Because the survey was administered ‘at arm’s length’ by the classroom teacher and the identity of the students kept confidential by the school, only the agreement of the school and participating teachers was obtained. Once students had been selected from the survey results, they were identified and permission sought for their interviews from the school, their teachers, from the student, and, because they were minors, their parents or those with the appropriate legal duty of care.

In requesting agreement from all parties involved – the schools, the teachers, the students and the caregivers – information outlining in “plain English” the nature and purpose of the research was provided, in accordance with the Human Ethics Committee requirements.

The use of volunteers in educational research is essential, particularly because of the power relationships that exist between adults and children, and teachers and students. Only those students who agreed to take part were interviewed. In accordance with child protection requirements, the researcher underwent the appropriate Working With Children checks.

The survey papers were numbered to avoid immediate identification of students. The researcher only had access to the identities of students interviewed. As it was the purpose of the study to investigate possible relationships between students’ algebraic knowledge and their explanations, it was necessary to match the interviewed student with their survey responses, thus complete anonymity was unable to be effected.

Provision for these ethical issues had to be considered in the planning stages of the study, and submitted to the university Human Ethics Committee for approval before research could commence. Matters of confidentiality remain important, however. In the report following, participants mentioned are either unidentified, or identified as “S1”, “S2”, etc. where several, different student responses are reported in close succession. Responses and analyses are not discussed in such a way that any particular school could be identified.

The study uses both quantitative and qualitative approaches, the latter taking an empirical perspective, using verbatim reports as supporting evidence. The quantitative data provide an objective basis for the selection and organisation of students into ability groups for the purposes of analysing their interview responses. The qualitative data are of a type that reflect classroom discourse. Information heard, or listened to, in a classroom is most often concerned with knowledge content. These data provide an opportunity to consider speech patterns of students, whist allowing the mathematical content to play a secondary role.

The following two chapters describe results from the survey, and the modelling and analyses of those results. The data so obtained answer the first three research questions, namely those that explore the nature and quality of students’ understanding of introductory algebra.

# CHAPTER 4: RESULTS FROM THE SURVEY

The first phase of the study, involved students from the three participating schools completing a forty-item survey, presented as a test. The survey items were drawn from the content of the *Mathematics 7-10 Syllabus* (Board of Studies NSW, 2002a), from textbooks commonly used in the participating schools and from Küchemann's research (1981) representing some aspects of elementary algebra (Appendix B).

In this chapter, responses to the survey are described, together with results of the Rasch modelling of the responses. The first section provides an overview of the general features of the survey results, and of the subsequent Rasch modelling. The second section consists of a detailed examination of the survey responses in terms of patterns arising from the item difficulty distribution in the Rasch model. The third section describes patterns of student ability estimates.

## OVERVIEW OF THE RESULTS

Inspection of raw data from the survey shows particular items attracted fewer correct responses than others. Rasch-modelling of the response data elaborated those patterns.

In this section, an overview of the response patterns is provided. The survey results are described firstly in terms of the raw data provided. A second sub-section, describes general results from the Rasch model (see Methodology Chapter 3 for details).

### Responses from the survey

The survey items were marked as "correct" or "incorrect". Correct items were those where only the complete response was given. No part-response was considered as being correct. For example, Item 19 [$2(x + 5) + 8$] was marked as correct only if the response "$2x + 18$" was given, but not for the part-response of "$2x + 10 + 8$". Fractions were to be written in their lowest terms, where possible. Incorrect responses included partly correct responses, erroneous responses, and non-attempts (blanks).

The numbers of correct and incorrect responses to each item are summarised in Table 4.1. Items are arranged in order of decreasing numbers of correct responses. The order reflects the order of the Rasch model.

*Table 4.1: Numbers of correct and incorrect responses to the survey items in decreasing order of numbers of correct responses*

| | Items in descending order of number of correct responses | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Item number | **4** | **27** | **24** | **2** | **1** | **28** | **5** | **23** | **9** | **30** |
| Number correct | 188 | 186 | 183 | 172 | 170 | 164 | 154 | 126 | 123 | 118 |
| Number Incorrect | 34 | 36 | 39 | 50 | 52 | 58 | 68 | 96 | 99 | 104 |
| Item number | **12** | **8** | **21** | **33** | **38** | **29** | **13** | **22** | **32** | **19** |
| Number correct | 108 | 107 | 106 | 106 | 106 | 104 | 98 | 98 | 98 | 96 |
| Number Incorrect | 114 | 115 | 116 | 116 | 116 | 118 | 124 | 124 | 124 | 126 |
| Item number | **31** | **26** | **6** | **18** | **25** | **40** | **17** | **20** | **15** | **37** |
| Number correct | 91 | 85 | 83 | 83 | 81 | 81 | 73 | 73 | 53 | 52 |
| Number Incorrect | 131 | 137 | 139 | 139 | 141 | 141 | 149 | 149 | 169 | 170 |
| Item number | **7** | **14** | **34** | **3** | **35** | **11** | **36** | **39** | **16** | **10** |
| Number correct | 48 | 48 | 45 | 43 | 31 | 29 | 23 | 17 | 16 | 12 |
| Number Incorrect | 174 | 174 | 177 | 179 | 191 | 193 | 199 | 205 | 206 | 210 |

Items in the survey were organised such that the first 26 items were expressions to be manipulated. The remaining items were equations to be solved. The items that attracted the greatest number of correct responses were: three items at the beginning of the survey consisting of expressions (Item 1, $3m + 8 + 2m - 5$; Item 2, $5p - p + 1$; and Item 4, $4 \times 5b$, and three items consisting of simple equations to be solved (Item 24, If $a + b = 43$, then $a + b + 2 = ?$; Item 27, If $x + 5 = 7$, then $x = ?$; and Item 28, If $4y = 20$, then $y = ?$).

The items for which there were the least number of correct responses were Item 10 $[(x + y)^2]$, Item 16 $[2/a^2 \times 5a/4]$ and Item 39 [If $ax = 5$, then $x = ?$]. For the purposes of Rasch-modelling both incorrect responses, and non-attempts were coded as incorrect. When a distinction was made between the types of incorrect responses (see Table 5.1 in the following chapter), it was found that most students had attempted these three items. Students tended to give wrong (but complete) responses to Item 10 and Item 39. Item 16 was most often answered with partially correct responses.

The pattern of correct and incorrect responses is illustrated in Figure 4.1. The graph was obtained by calculating as a fraction the number of correct responses to the number of total responses. Plotting these, clusters of items can be seen.

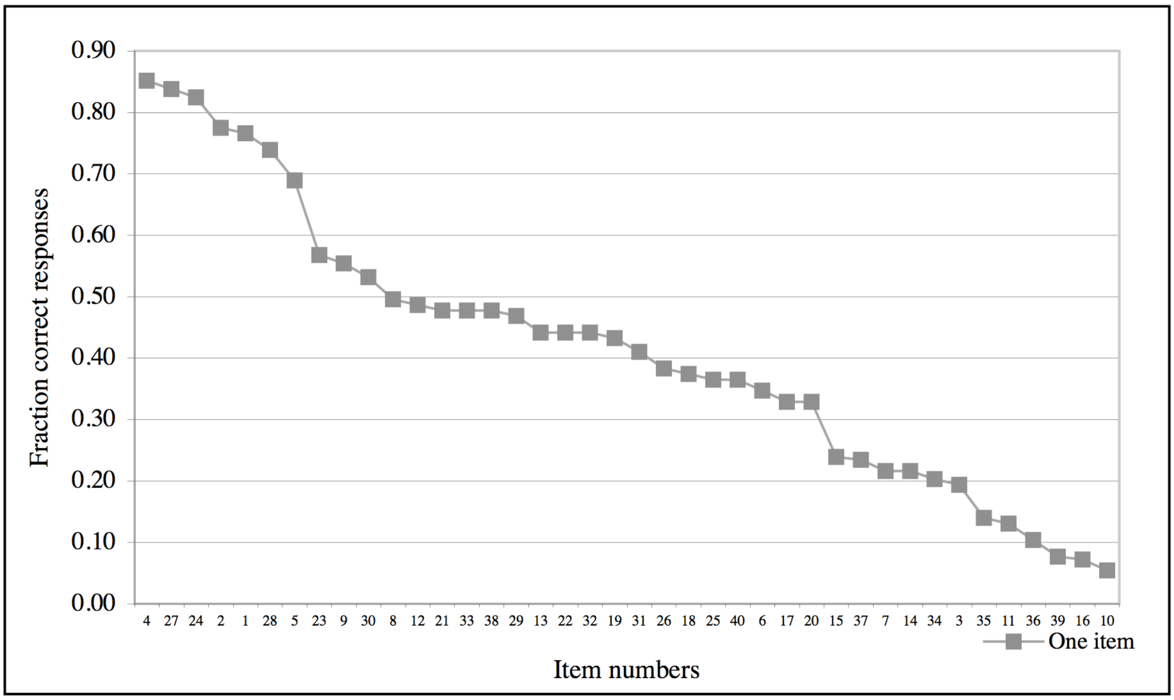


*Figure 4.1: Fraction of correct responses, items in descending order*

The seven easiest items appear as one distinct cluster (items 4 to 5, reading along the horizontal axis), separate from a second cluster (items 23 to 20), consisting of the majority of items. These items were comprised of a variety of different algebra concepts. Two other distinct clusters (items 15 to 3, and items 35 to 10) consisted of items that attracted fewer correct responses. The range of the fraction of items correct within each cluster can be estimated from the vertical axis: in the first cluster the fraction of items correct ranged from 0.85 to 0.69; in the second cluster the range was 0.57 to 0.33; in the third cluster the range was 0.24 to 0.19; and, the fourth cluster, 0.14 to 0.05 correct. These patterns reflected the patterns of the Rasch model.

## Overview of the Rasch-modelling of the survey responses

The Rasch model measures the distribution of persons and items on an equal-interval scale for a single construct. In this study, the construct to be measured is that of basic symbolic algebra understanding. Where data do not fit the model, it can be assumed that the test instrument does not measure the construct it intends. This sub-section describes overall patterns emerging from the model, and statistics that indicate the reliability of the data.

The extent to which the data fit the model is indicative of the validity of the data collection instrument in terms of the overall construct assumed to be measure. The validity of the instrument is measured by the reliability of the item difficulty estimates and the student ability estimates. Item difficulty estimates, with a mean of 1.00 logits have a reliability of 0.99, case estimates, with a mean of -0.64 logits, have a reliability of 0.93 ($p$=0.05). These figures are well within the acceptable limit of 0.7.

Item fit statistics detect "discrepancies between the Rasch model prescriptions and the data…collected" (Bond & Fox, 2007, p.235). Data for all items and all cases are included in Appendix C. The data for item estimates and case estimates, together with the fit statistics for items and cases are summarised in Table 4.2.

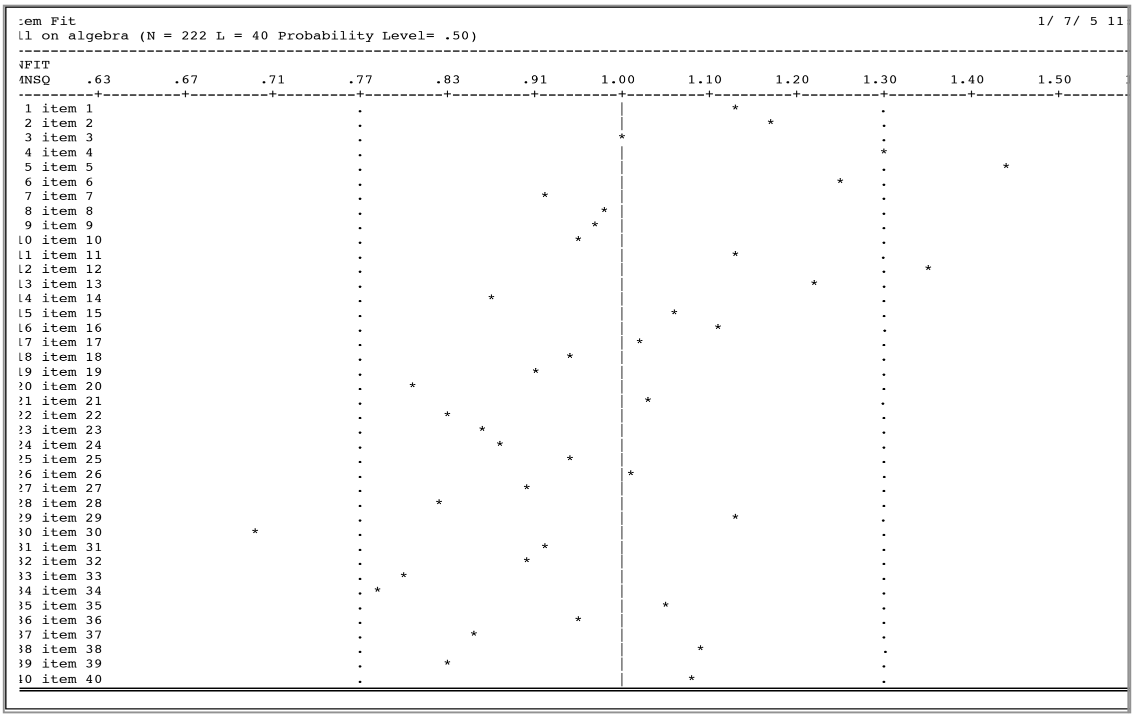
*Table 4.2: Summary Statistics for Items and Cases*

| | | | | | |
|---|---|---|---|---|---|
| Item Difficulty Estimates and Fit Statistics | | | | | |
| Item Difficulty Estimates | | Item Fit Statistics | | | |
| Mean | 0.00 | Infit Mean Square | | Outfit Mean Square | |
| SD | 1.75 | Mean | 1.00 | Mean | 0.99 |
| SD (adjusted) | 1.74 | SD | 0.16 | SD | 0.47 |
| Reliability of estimate | 0 .99 | Infit *t* | | Outfit *t* | |
| | | Mean | -0.05 | Mean | -0.05 |
| | | SD | 1.68 | SD | 1.26 |
| | 0 items with zero scores | 0 items with perfect scores | | | |
| Case Ability Estimates and Fit Statistics | | | | | |
| Case Ability Estimates | | Case Fit Statistics | | | |
| Mean | -0.64 | Infit Mean Square | | Outfit Mean Square | |
| SD | 1.89 | Mean | 0.99 | Mean | 0.99 |
| SD (adjusted) | 1.83 | SD | 0.22 | SD | 0.77 |
| Reliability of estimate | 0.93 | Infit *t* | | Outfit *t* | |
| | | Mean | 0.02 | Mean | 0.17 |
| | | SD | 0.94 | SD | 0.73 |
| | 0 cases with zero scores | 0 cases with perfect scores | | | |

The average difficulty of all items was constrained at 0.0 logits with an adjusted standard deviation of 1.74 logits ($N$ = 40). Most items fitted the model as described by the infit and outfit statistics. No items were answered incorrectly by all students, and no items were answered correctly by all students. Hence, the average ability estimate for all students was -0.64 ± 1.89 logits ($N$ = 222). No student gained a total score of zero, and no student obtained a perfect score.

Figure 4.2 illustrates the infit mean square statistics for each item. To fit the model, the infit statistic for each item must lie between 0.77 and 1.30. These limits are represented by the vertical dotted lines in Figure 4.2. Each item is represented by an asterisk (*). Items lying outside the limits are said to display *overfit* or *underfit*. Three items did not fit the model. These lie outside the vertical dotted lines in Figure 4.2. One item to the left of the vertical line at 0.77 shows *overfit*. Two items to the right of the vertical line at 1.30 show *underfit*.

Overfit items show less than expected variation in patterns of responses. Linacre (2002) suggested that, “other dimensions may be constraining the response patterns” (p.878). The pattern of student responses follows the expected pattern more closely than normal variation would expect: almost all of the students not expected to be able to respond successfully to the item, do not do so; almost all of the students expected to be able to respond to the item, do so. Item 30 [If *x*/4 = 12, what is *x*?], with an infit mean square of 0.7 (infit *t* =-3.72; outfit *t* = -2.2) fits the model better than predicted. Overfit does not detract from the reliability of the test instrument to the extent that items with underfit might do.

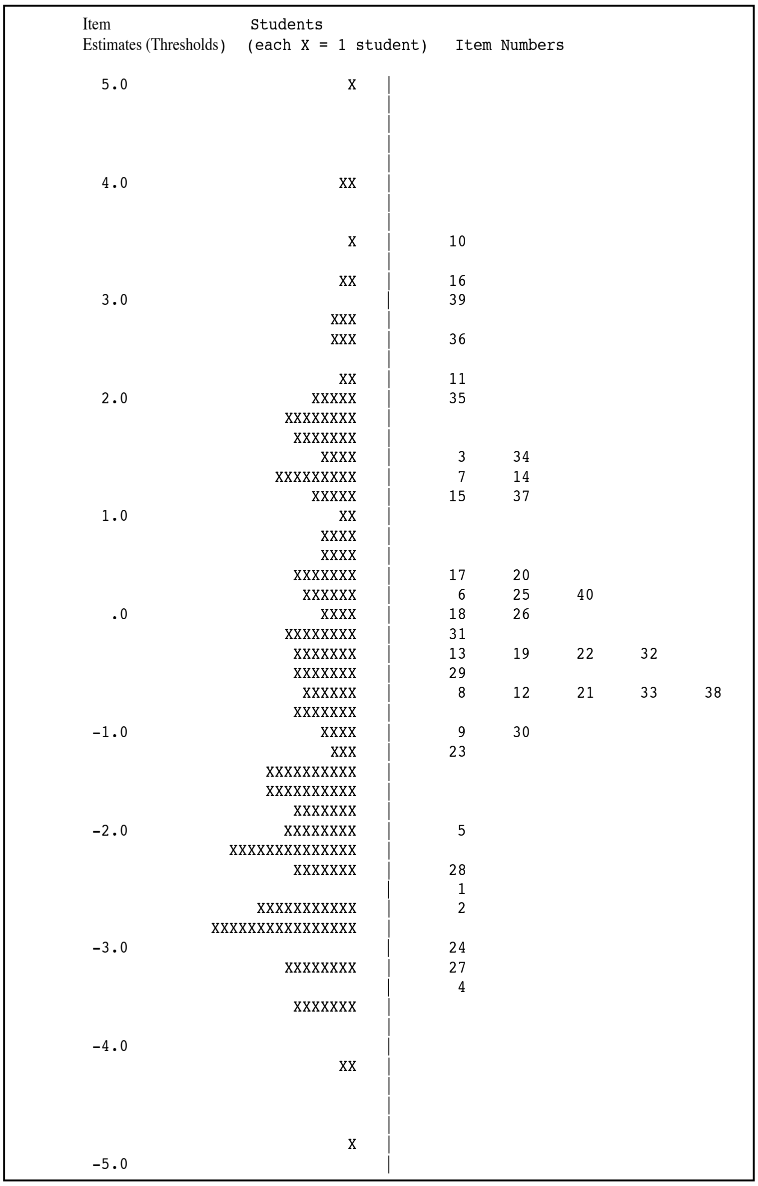


```
 em Fit                                                                                                           1/ 7/ 5 11
 1 on algebra (N = 222 L = 40 Probability Level= .50)
-----------------------------------------------------------------------------------------------------------------------------
 FIT
 NSQ    .63       .67       .71       .77       .83       .91      1.00      1.10      1.20      1.30      1.40      1.50
---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
 1 item 1                              .                             |            *                .
 2 item 2                              .                             |                *            .
 3 item 3                              .                             *                             .
 4 item 4                              .                             |                             *
 5 item 5                              .                             |                             .             *
 6 item 6                              .                             |                        *    .
 7 item 7                              .                    *        |                             .
 8 item 8                              .                           * |                             .
 9 item 9                              .                          *  |                             .
 0 item 10                             .                        *    |                             .
 1 item 11                             .                             |            *                .
 2 item 12                             .                             |                             .    *
 3 item 13                             .                             |                     *       .
 4 item 14                             .              *              |                             .
 5 item 15                             .                             |     *                       .
 6 item 16                             .                             |          *                  .
 7 item 17                             .                             | *                           .
 8 item 18                             .                       *     |                             .
 9 item 19                             .                   *         |                             .
 0 item 20                             .     *                       |                             .
 1 item 21                             .                             |  *                          .
 2 item 22                             .         *                   |                             .
 3 item 23                             .             *               |                             .
 4 item 24                             .               *             |                             .
 5 item 25                             .                       *     |                             .
 6 item 26                             .                             |*                            .
 7 item 27                             .                  *          |                             .
 8 item 28                             .        *                    |                             .
 9 item 29                             .                             |            *                .
 0 item 30                 *           .                             |                             .
 1 item 31                             .                    *        |                             .
 2 item 32                             .                  *          |                             .
 3 item 33                             .    *                        |                             .
 4 item 34                             . *                           |                             .
 5 item 35                             .                             |    *                        .
 6 item 36                             .                        *    |                             .
 7 item 37                             .            *                |                             .
 8 item 38                             .                             |        *                    .
 9 item 39                             .         *                   |                             .
 0 item 40                             .                             |       *                     .
```

*Figure 4.2: Map of Item Fit. Items fit the Rasch model if they map between 0.77 and 1.30 logits*

Underfit items show a greater-than-expected variation in response patterns. The numbers of students expected to respond correctly to the item are significantly fewer than expected. On the other hand, the numbers of students not expected to answer the item correctly are more than expected. The two items which appear as having greater variation than expected are Item 5 [2*ab* + 3*b* + *ab*] with an infit mean square of 1.44 (infit *t* = 4.62; outfit *t* = 3.58) and Item 12 [2/*a* × 3/*b*], with an infit mean square of 1.35 (infit *t* = 3.52; outfit *t* = 2.23). These items lie outside the vertical lines at 1.30, to the right of Figure 4.2. Item 5 had a 44% greater variation than expected, and Item 12, 35% more variation than expected. Possible reasons for the misfit of items are explored through error analyses and interviews with students (Chapter 5).

Quest software (Adams & Khoo, 1996) also prints a map of item difficulty together with case estimates (student ability estimates). This map is reproduced in Figure 4.3.



```
Item                        Students
Estimates (Thresholds)  (each X = 1 student)   Item Numbers

   5.0                     X  |
                              |
                              |
                              |
                              |
   4.0                    XX  |
                              |
                              |
                           X  |   10
                              |
                          XX  |   16
   3.0                        |   39
                         XXX  |
                         XXX  |   36
                              |
                          XX  |   11
   2.0                 XXXXX  |   35
                    XXXXXXXX  |
                     XXXXXXX  |
                        XXXX  |    3     34
                   XXXXXXXXX  |    7     14
                       XXXXX  |   15     37
   1.0                    XX  |
                        XXXX  |
                        XXXX  |
                     XXXXXXX  |   17     20
                      XXXXXX  |    6     25     40
    .0                  XXXX  |   18     26
                    XXXXXXXX  |   31
                     XXXXXXX  |   13     19     22     32
                     XXXXXXX  |   29
                      XXXXXX  |    8     12     21     33     38
                     XXXXXXX  |
  -1.0                  XXXX  |    9     30
                         XXX  |   23
                  XXXXXXXXXX  |
                  XXXXXXXXXX  |
                     XXXXXXX  |
  -2.0              XXXXXXXX  |    5
              XXXXXXXXXXXXXX  |
                     XXXXXXX  |   28
                              |    1
                 XXXXXXXXXXX  |    2
            XXXXXXXXXXXXXXXX  |
  -3.0                        |   24
                    XXXXXXXX  |   27
                              |    4
                     XXXXXXX  |
                              |
  -4.0                        |
                          XX  |
                              |
                              |
                              |
                           X  |
  -5.0                        |
```

*Figure 4.3: Map of student performance (by ability estimates) and item difficulty.*

Logit values are on the left; student ability estimates, for each student, are indicated by "x". On the right of the vertical line, item numbers are arranged according to the level of difficulty. Column labels have been added to the print-out from Quest software for greater clarity of reading. Mean item difficulty was constrained at 0.00 logits. On the same scale, mean student ability is marked at -0.64 logits.

|  | Item Difficulty | Item Numbers | Items |
|---|---|---|---|
| 3.5 — | 3.56 | 10 | 10. $(x + y)^2$ |
|  | 3.17 | 16 | 16. $\frac{2}{a^2} \times \frac{5a}{4}$ |
| 3.0 — | 3.09 | 39 | 39. If $ax = 5$, then $x = \ldots$? |
|  | 2.64 | 36 | 36. If $\frac{63}{x} = 180$ What is x equal to? |
| 2.5 — |  |  |  |
|  | 2.28 | 11 | 11. $8p - 2(p + 5)$ |
|  | 2.17 | 35 | 35. $x + \frac{x}{3} = 4$ What is $x$ equal to? |
| 2.0 — |  |  |  |
|  | 1.59 | 3 | 3. $\frac{a}{5} + \frac{a}{10}$ |
| 1.5 — | 1.51 | 34 | 34. $x + (x + 2) = (x - 1) + 8$, What is $x$ equal to? |
|  | 1.38 | 7, 14 | 7. $(a - b) + b$;<br>14. $\frac{x}{3} \div \frac{y}{4}$ |
|  | 1.22 | 37 | 37. Solve: $5a - 4 = 2a + 8$ |
|  | 1.18 | 15 | 15. $\frac{3p}{4} - \frac{p}{8}$ |
| 1.0 — |  |  |  |
| 0.5 — | 0.46 | 17, 20 | 17. $(6xy)^2$;<br>20. Multiply $x + 5$ by 4 |
|  | 0.33 | 6 | 6. $5a - 2b + 3a + 3b$ |
|  | 0.2 | 25, 40 | 25. Take n away from $3n + 1$<br>40. If $e + f = 8$, then $e + f - g = \ldots$? |
|  | 0.13 | 18 | 18. $2(x + 4) + 3(x - 1)$ |
| 0.0— | 0.07 | 26 | 26. If $p + q = 5$, then $p + q + r = \ldots$? |
|  | -0.12 | 31 | 31. Solve $4(p + 3) = 32$ |
|  | -0.27 | 19 | 19. $2(x + 5) + 8$ |
|  | -0.34 | 13, 22, 32 | 13. $4r \times 5t \times 3$;<br>22. Add 3 onto $4n$;<br>32. If $10y = 5$, $y = \ldots$? |
| -0.5 — | -0.52 | 29 | 29. What is $t$ if $2t - 23 = 49$? |
|  | -0.58 | 21, 33, 38 | 21. Add 4 on to $n + 5$;<br>33. $\frac{x + 3}{2} = 7$, What is $x$ equal to?<br>38. If $r - 82 = 7$, then $r - 83 = \ldots..$?; |
|  | -0.64 | 8 | 8. $\frac{4ab}{4b}$ ; |
|  | -0.7 | 12 | 12. $\frac{2}{a} \times \frac{3}{b}$ |
| -1.0 — | -0.93 | 30 | 30. If $\frac{x}{4} = 12$, what is $x$? |
|  | -1.08 | 9 | 9. $2ab \times a$ |
|  | -1.17 | 23 | 23. What can you say about $m$ if $m = 3n + 1$ and $n = 4$ ? |
| -1.5 — |  |  |  |
| -2.0 — | -1.98 | 5 | 5. $2ab + 3b + ab$ |
|  | -2.33 | 28 | 28. If $4y = 20$ , then $y = \ldots.$? |
| -2.5 — | -2.53 | 1 | 1. $3m + 8 + 2m - 5$ |
|  | -2.6 | 2 | 2. $5p - p + 1$ |
| -3.0 — | -3.02 | 24 | 24. If $a + b = 43$, $a + b + 2 = \ldots$? |
|  | -3.14 | 27 | 27. If $x + 5 = 7$, then $x = \ldots$? |
|  | -3.27 | 4 | 4. $4 \times 5b$ |
| -3.5 — |  |  |  |

*Figure 4.4: Survey items in descending order of difficulty estimates, adapted from the map in Figure 4.3*

Figure 4.4 (above) elaborates Figure 4.3 for ease of identification of the items that fall at each of the difficulty estimates. Each item is listed against the item difficulties in descending order. Most items cluster between the difficulty estimates of +0.5 to -1.17 logits. Gaps between clusters are also discernable from the graph of raw score data (Figure 3.1).

### Summary

Numbers of correct and incorrect responses were used to organise survey items in descending order of numbers of successful responses. This was reflected in the order resulting from the Rasch-modelling of the data. Summary statistics of item difficulty and ability (case) estimates show the test instrument to be valid, testing for a single construct. All but three items fitted the model: one item showed overfit of the model, and two items showed underfit of the model.

Maps from the modelling revealed clusters of item difficulty, similar to clusters seen when fractions of correct to incorrect responses were calculated and mapped.  Different types of items constituted the clusters. Student ability estimates clustered around three significantly different means that appear to align with the mean difficulty of item clusters.

## PATTERNS OF ITEM DIFFICULTY

The Rasch model maps items so that relative difficulties of items are indicated by the size of intervals between the items when arranged in order of difficulty. The difficulty estimates of items imply that particular cognitive demands are placed on students if they are to answer these items correctly. It may be inferred from the clustering that items within each cluster require similar cognitive demands. Patterns of item difficulty estimates are considered firstly, as clusters that appear on the maps in Figure 4.2 and Figure 4.3, and secondly, in categories of related algebraic ideas. A third sub-section discusses items that did not fit the model.

### Clusters of item difficulty

The maps in Figure 4.2 and Figure 4.3 indicate that the survey items have clustered in three statistically significant groups. These clusters are labelled as Cluster 1(items of least difficulty), Cluster 2 (items of average difficulty), and Cluster 3 (items of greatest difficulty).

Cluster 3 consists of three less significantly different groups. These have been labelled as Cluster 3a, 3b and 3c. The ranges of the cluster extremes, the number of items in each cluster and the cluster means, are summarised in Table 4.3.

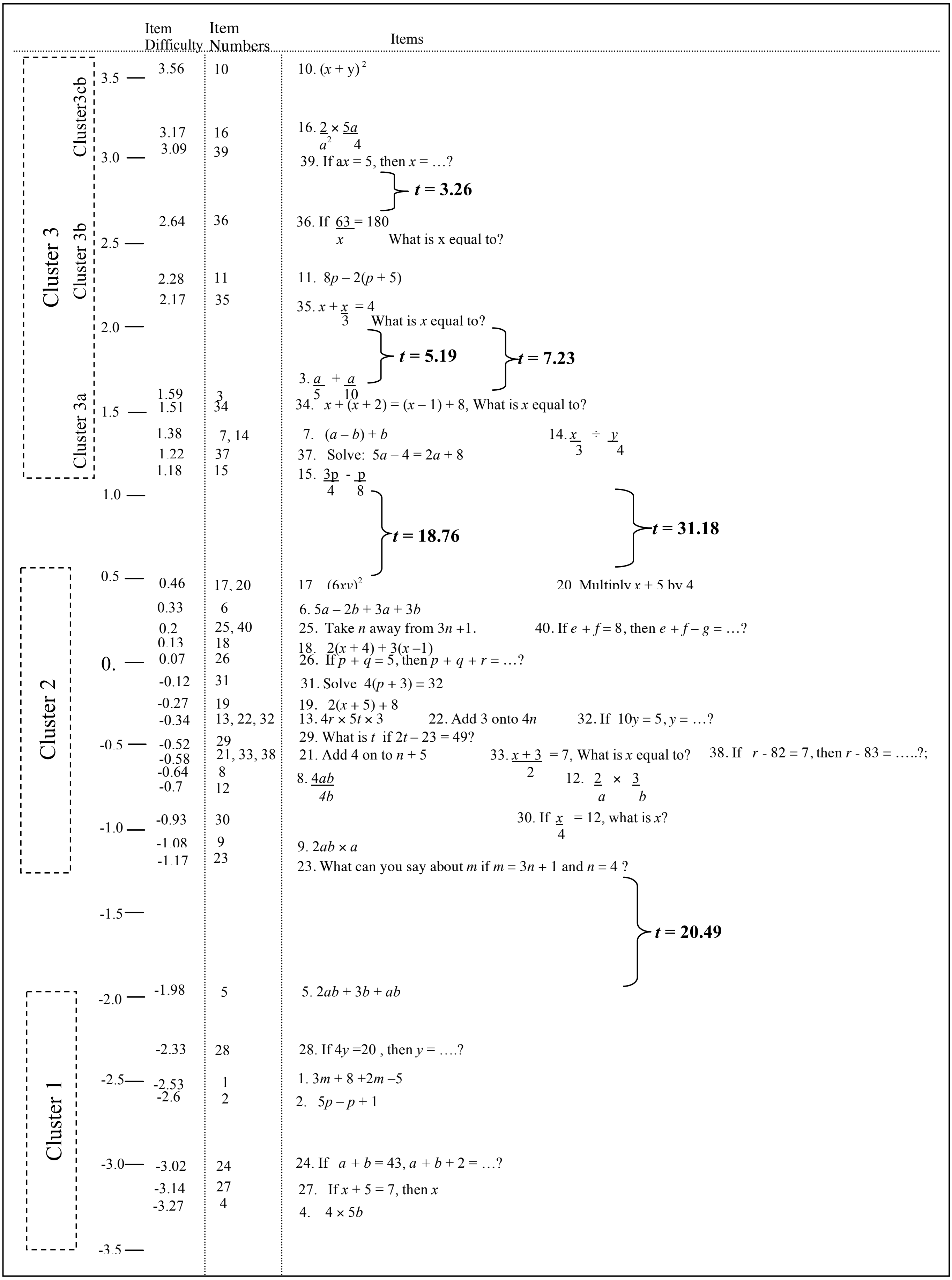
*T*-test results for significance between the means of each cluster, and sub-cluster, are given on Figure 4.5. Each item is listed against the appropriate difficulty estimate (see the original Quest print-out in Figure 4.3), and arranged in descending order of difficulty.

*Table 4.3: Item clusters, with number of items in each cluster, range of difficulty estimates and cluster difficulty means*

| Item Cluster | Number of Items | Range in logits | Cluster mean |
|---|---|---|---|
| 3c: | 3 | 3.09 to 3.56 | 3.273 |
| 3b: | 3 | 2.17 to 2.64 | 2.363 |
| 3a | 6 | 1.18 to 1.59 | 1.377 |
| All 3: | 12 | 1.18 to 3.56 | 2.095 |
| 2: | 21 | -1.17 to 0.46 | -0.302 |
| 1 | 7 | -3.27 to –1.98 | -2.70 |

Items occurring in each cluster are psychometrically similar. This need not imply that some of these items are redundant. Different items embody different algebraic concepts, even though, as occurs in Cluster 2 (–1.17 to 0.46 logits), several items have the same difficulty estimate. This does mean that although the algebraic concepts are different, cognitive demands on students are similar. Also, differences between means of each cluster imply that significant changes in students' thinking are needed for successful engagement with items in the next most difficult cluster. The difference in the means of Cluster 1 and Cluster 2 is highly significant ($t$=20.49, $p$<0.001), as is the difference in means between Cluster 2 and all of Cluster 3 ($t$=18.76, p<0.001). Also significant, is the difference in means between Cluster 2 and Cluster 3a ($t$=31.18,$p$<0.001). The differences in means between the sub-clusters within Cluster 3 are also significant: Clusters 3a and 3b ($t$=7.23, $p$<0.001); and, for Clusters 3b and 3c the difference in means is less significant ($t$=3.26, $p$<0.01).

These data suggest that facility in answering correctly the items in Cluster 2 requires a considerable change to the algebraic conceptualisation used to answer successfully the items



| Item Difficulty | Item Numbers | Items |
|---|---|---|
| 3.56 | 10 | 10. $(x + y)^2$ |
| 3.17 | 16 | 16. $\frac{2}{a^2} \times \frac{5a}{4}$ |
| 3.09 | 39 | 39. If $ax = 5$, then $x = \ldots$? |
| 2.64 | 36 | 36. If $\frac{63}{x} = 180$ What is x equal to? |
| 2.28 | 11 | 11. $8p - 2(p + 5)$ |
| 2.17 | 35 | 35. $x + \frac{x}{3} = 4$ What is $x$ equal to? |
| 1.59 | 3 | 3. $\frac{a}{5} + \frac{a}{10}$ |
| 1.51 | 34 | 34. $x + (x + 2) = (x - 1) + 8$, What is $x$ equal to? |
| 1.38 | 7, 14 | 7. $(a - b) + b$<br>14. $\frac{x}{3} \div \frac{y}{4}$ |
| 1.22 | 37 | 37. Solve: $5a - 4 = 2a + 8$ |
| 1.18 | 15 | 15. $\frac{3p}{4} - \frac{p}{8}$ |
| 0.46 | 17, 20 | 17. $(6xy)^2$<br>20. Multiply $x + 5$ by 4 |
| 0.33 | 6 | 6. $5a - 2b + 3a + 3b$ |
| 0.2 | 25, 40 | 25. Take $n$ away from $3n + 1$.<br>40. If $e + f = 8$, then $e + f - g = \ldots$? |
| 0.13 | 18 | 18. $2(x + 4) + 3(x - 1)$ |
| 0.07 | 26 | 26. If $p + q = 5$, then $p + q + r = \ldots$? |
| -0.12 | 31 | 31. Solve $4(p + 3) = 32$ |
| -0.27 | 19 | 19. $2(x + 5) + 8$ |
| -0.34 | 13, 22, 32 | 13. $4r \times 5t \times 3$<br>22. Add 3 onto $4n$<br>32. If $10y = 5, y = \ldots$? |
| -0.52 | 29 | 29. What is $t$ if $2t - 23 = 49$? |
| -0.58 | 21, 33, 38 | 21. Add 4 on to $n + 5$<br>33. $\frac{x + 3}{2} = 7$, What is $x$ equal to?<br>38. If $r - 82 = 7$, then $r - 83 =$ …..?; |
| -0.64 | 8 | 8. $\frac{4ab}{4b}$ |
| -0.7 | 12 | 12. $\frac{2}{a} \times \frac{3}{b}$ |
| -0.93 | 30 | 30. If $\frac{x}{4} = 12$, what is $x$? |
| -1.08 | 9 | 9. $2ab \times a$ |
| -1.17 | 23 | 23. What can you say about $m$ if $m = 3n + 1$ and $n = 4$ ? |
| -1.98 | 5 | 5. $2ab + 3b + ab$ |
| -2.33 | 28 | 28. If $4y = 20$ , then $y = \ldots$.? |
| -2.53 | 1 | 1. $3m + 8 + 2m - 5$ |
| -2.6 | 2 | 2. $5p - p + 1$ |
| -3.02 | 24 | 24. If $a + b = 43, a + b + 2 = \ldots$? |
| -3.14 | 27 | 27. If $x + 5 = 7$, then $x$ |
| -3.27 | 4 | 4. $4 \times 5b$ |

*Figure 4.5: Items in descending order of difficulty estimates, with t-values for intervals between clusters*

in Cluster 1. This is also the case as students move successfully from Cluster 2 to Cluster 3. The conceptual development required to move from success with items in each of the smaller clusters (3a, 3b, and 3c) might not be as great.

## Categories of items and item difficulty

Possible reasons for the patterns of item difficulty reside in the types of items treated in the survey. In order to explore this, items were organised into categories according to particular algebraic features, based on syllabus topics (Board of Studies NSW, 2002a). These features were: the presence of brackets (Category 1: Brackets); fraction notation (Category 2: Fractions); the need to interpret descriptive relationships between variables (Category 3: Semi-literal); and, the use of one (or more) of the four basic arithmetic operations (Category 4: Addition/Subtraction, and Category 5: Multiplication).

Many of the items could be categorised in other ways. Item 39 [$ax = 5$] might also be included with other items in the Multiplication category because of its superficial resemblance to other such equations. Item 24 [If $a + b = 43$, then $a + b + 2 =$...?] could have been included with items in the Addition/Subtraction category, but was included with Item 21 [Add 4 onto $n +5$], Item 22 [Add 3 onto $4n$] and Item 25 [Take $n$ away from $3n + 1$], which have been termed “semi-literal” (see below for a detailed explanation).

Within each category, the items were arranged according to their Rasch difficulty estimate. The resulting patterns are illustrated in Figure 4.6, which includes the item/case estimate scale (in logits), in the left-hand-most column, and ability estimates for each student (represented by “x”). The items are identified by their survey numbers only, arranged in descending order of difficulty. In each column of items, identified by category, expressions are placed to the left of the column, equations to the right. The mean for each category is indicated in approximate positions on the difficulty scale. The range of difficulty and the average difficulty for each category can be seen. Items involving brackets or fractions appear to be more difficult than items involving addition or subtraction, or multiplication of terms. The category of Semi-literal items has the greatest range, but the average difficulty for this category approximates the mean ability estimate for all students.

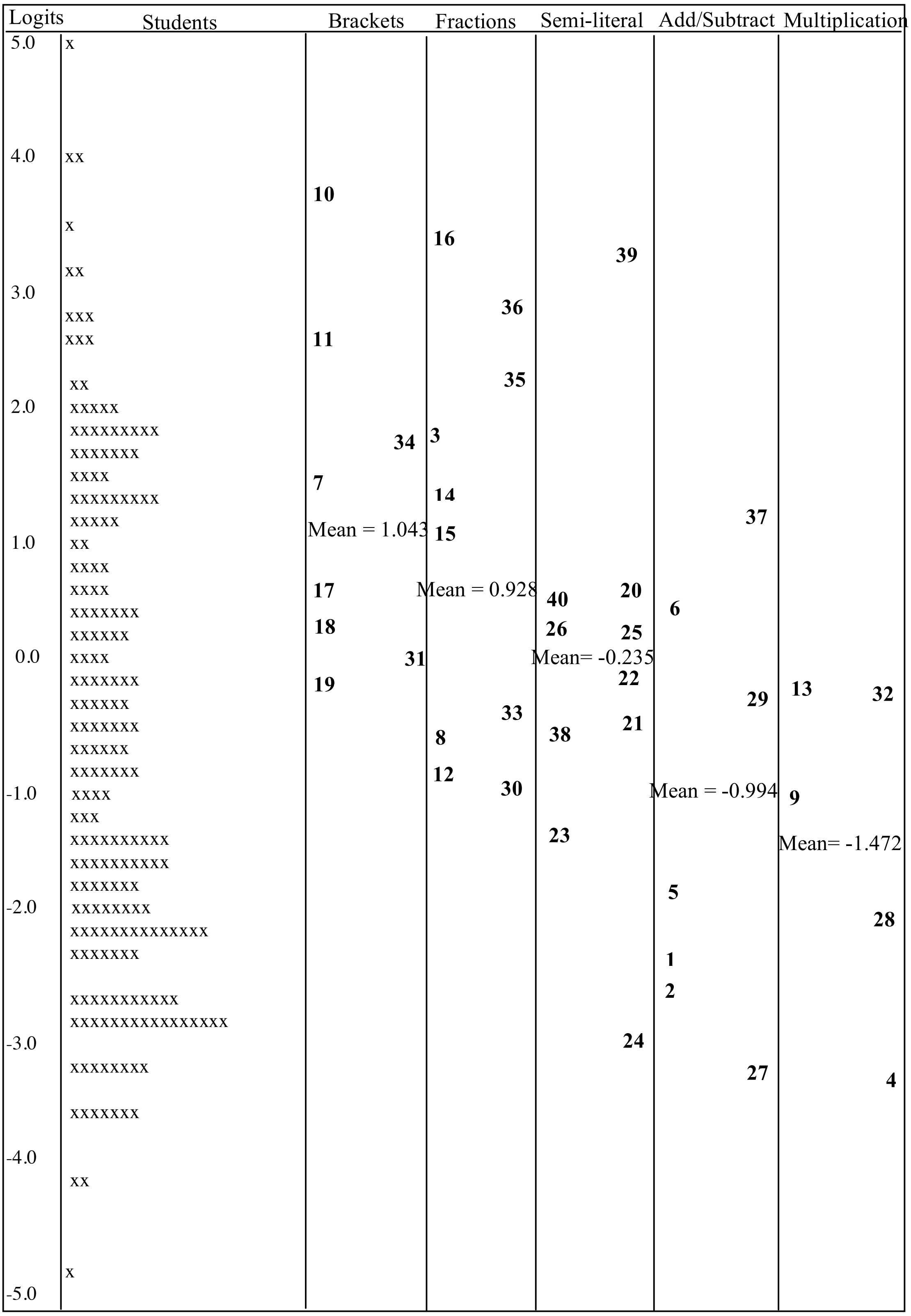


*Figure 4.6: Items (numbers in bold) arranged in categories of syllabus topics. (Expressions are to the left of each column, equations to the right. Individual students are designated with "x", adapted from the Quest map )*

A description of each of the categories together with a list of the items in each category, their Rasch difficulty estimates, the mean difficulty estimate for each group, follows. The significance of the differences between the mean difficulty for each category is also described in the final part of this sub-section.

### *Category 1: Brackets*

Items in this category are described and listed in Table 4.4 together with the survey item number and Rasch difficulty estimate. From the data, it appears that any items with brackets are of a difficulty level such that no more than 25% of students (55 out of 222) could be expected to respond correctly. Successful responses required students to have a flexible and robust understanding of the role of brackets.

*Table 4.4: Items categorised as Brackets in descending order of difficulty*

| Description | Expressions | | | Equations | | |
|---|---|---|---|---|---|---|
| | Item Number | Item | Rasch difficulty estimate | Item Number | Item | Rasch difficulty estimate |
| Expressions or equations requiring the application of distributive law, collection of terms, and the perception of brackets as grouping symbols | 10 | $(x+y)^2$ | 3.56 | 34 | $x+(x+2)=(x-1)+8$ What is $x$ equal to? | 1.51 |
| | 11 | $8p - 2(p+5)$ | 2.26 | | | |
| | 7 | $(a - b)+b$ | 1.38 | 31 | Solve $4(p+3)=32$ | -0.12 |
| | 17 | $(6xy)^2$ | 0.46 | | | |
| | 18 | $2(x+4)+3(x-1)$ | 0.13 | | | |
| | 19 | $2(x+5)+8$ | -0.27 | | | |
| *Mean difficulty: 1.12* | | | | | | |

The average ability for all students on the entire survey was -0.64 logits (Table 4.2), and the least difficult item where brackets were involved was Item 19 [2($x$ + 5) + 8], with a difficulty estimate of -0.27, 0.34 logits above the overall mean. The presence of negatives (Items 7, 11, and 18), the lack of numbers (Item 7), the necessity to interpret notation correctly (Item 17, Item 10), and the syntax of expressions (Item 11) – all appeared to contribute to the increase of item difficulty above -0.27 logits. These understandings were also important for students responding to the items that were equations. The increase in the difficulty estimate for Item 34 may also be because of the unknown being found on both sides of the equation.

### *Category 2: Fractions*

Items in this category are described and listed in Table 4.5, together with the survey item number and Rasch difficulty estimate. These items, which included six expressions to be simplified and four equations to be solved, were categorised on the basis of their inclusion under the same syllabus topics (Board of Studies NSW, 2002a). In the case of the expressions (Items 12, 8, 15, 14, 3, and 16), fractions were to be added, subtracted, multiplied or divided. In the case of the equations (Items 30, 35, 33, and 36), the notation of a fraction was used to indicate division. Four items were scaled with a difficulty estimate near the mean ability estimate (-0.64 logits) for the whole group ($N = 222$). Two of these were expressions to be simplified, and two equations were to be solved.

*Table 4.5: Items categorised as Fractions in descending order of difficulty*

| Description | Expressions | | | Equations | | |
|---|---|---|---|---|---|---|
| | Item Number | Item | Rasch difficulty estimate | Item Number | Item | Rasch difficulty estimate |
| Expressions or equations where there is fraction notation used. Simplification of expressions involves one or more of four arithmetic operations. Solution of equations requires at least one step, using multiplication as the inverse of division. | 16 | $\frac{2}{a^2} \times \frac{5a}{4}$ | 3.17 | 36 | If $63/x = 180$<br>What is $x$ equal to? | 2.64 |
| | 3 | $\frac{a}{5} + \frac{a}{10}$ | 1.59 | | | |
| | 14 | $\frac{x}{3} \div \frac{y}{4}$ | 1.38 | 35 | $x + x/3 = 4$,<br>What is $x$? | 2.17 |
| | 15 | $\frac{3p}{4} - \frac{p}{8}$ | 1.18 | | | |
| | 8 | $\frac{4ab}{4b}$ | -0.64 | 33 | $\frac{x+3}{2} = 7$<br>What is $x$ equal to? | -0.58 |
| *Mean difficulty: 0.928* | 12 | $\frac{2}{a} \times \frac{3}{b}$ | -0.7 | 30 | If $x/4 = 12$,<br>what is $x$? | -0.93 |

Structurally, Item 8 [$4ab/4b$] and Item 16 [$2/a^2 \times 5a/4$] are similar, and require similar manipulations in order to simplify the expression completely. The difference in difficulty estimates (Item 8, difficulty estimate: -0.64 logits and Item 16, difficulty estimate: 3.17 logits) might be explained through examination of the errors together with interview data.

Answers to Items 30 [$x/4 = 12$] and 33 [$(x + 3)/2 = 7$] could be found by students recalling number facts, whereas this was not so easily done for Item 36 [$63/x = 180$]. Item 36 needed students to realise that division could result in a quotient greater than the dividend, and hence an answer that was not a whole number. Item 35 [$x+x/3 = 4$], was much more structurally complex item than Item 33. Hence the increased difficulty estimates for these Items 35 and 36.

### *Category 3: Semi-literal Items*

Items in this category are described and listed in Table 4.6, together with the survey item number and Rasch difficulty estimate. Items in this group state relationships between numbers as letters, and require students to make inferences from those relationships. Item 24 [If $a + b = 43$, then $a + b + 2 = ...$?] can be answered by the use of addition in much the same way as Item 27 [$x + 5 = 7$] (from Category 4). Item 20 [Multiply $x + 5$ by 4], could have been included in the category of Brackets, but was better placed in this category as the need for brackets is only implied by the question, rather than being made explicit by the visual structure of the question. The ambiguity in the statement appears to have made Item 20 more difficult than others in the Semi-literal group, with the exception of Item 39 [$ax = 5$]. Other Items (23, 38, 26, 40, and 24) required an understanding of additive relationships that, in each of these cases, could be established by the substitution of simple positive integers.

*Table 4.6: Items categorised as Semi-literal in descending order of difficulty*

| Description | Expressions | | | Equations | | |
|---|---|---|---|---|---|---|
| | Item Number | Item | Rasch difficulty estimate | Item Number | Item | Rasch difficulty estimate |
| Expressions or equations requiring the reading of an algebraic statement describing the relationship between numbers, and using that relationship to rewrite the expression solve an equation. | 20 | Multiply $x + 5$ by 4 | 0.46 | 39 | If $ax = 5$, then $x = \ldots$? | 3.09 |
| | 25 | Take $n$ away from $3n + 1$ | 0.2 | 40 | If $e + f = 8$, then $e + f - g = \ldots$? | 0.2 |
| | 22 | Add 3 onto $4n$ | -0.34 | 26 | If $p + q = 5$, then $p + q + r$=? | 0.07 |
| | 21 | Add 4 on to $n + 5$ | -0.58 | 38 | If $r - 82 = 7$, then $r - 83 = \ldots$? | -0.58 |
| *Mean difficulty:*<br>*- 0.167* | 23 | What can you say about $m$ if $m = 3n + 1$ and $n = 4$? | -1.17 | 24 | If $a + b = 43$; $a + b + 2 =$? | -3.02 |

Item 39 appears, on the surface, to involve simple multiplication. It has, however, further conceptual complexity because there is no unique solution. Instead the item is a general statement about the relationships existing between terms. For this reason, the item was included in the Semi-literal category. The conceptual complexity was confirmed by the Rasch difficulty estimate (3.09 logits), as well as by the types of responses given by the students.

### *Category 4: Addition/Subtraction*

Items in this category are described and listed in Table 4.7, together with the survey item number and Rasch difficulty estimate. The difficulty estimates of the items in the Semi-literal category closely aligned with those for items in the category of Addition/Subtraction. This may be attributed to their conceptual similarities. The one exception was Item 37 [$5a - 4 = 2a + 8$], a linear equation, with the unknown on both sides. It was structurally similar to Item 34 [$x + (x + 2) = (x - 1) + 8$], but lacking the superficial, but obvious, feature of brackets. However, it was also similar to Item 29 [$2t - 23 = 49$], in that the manipulation of the terms in the equation in order to arrive at a solution consisted predominantly of a sequence of addition/subtraction steps.

*Table 4.7: Items categorised as Addition/Subtraction in descending order of difficulty*

| Description | Expressions | | | Equations | | |
|---|---|---|---|---|---|---|
| | Item Number | Item | Rasch difficulty estimate | Item Number | Item | Rasch difficulty estimate |
| Expressions or equations that require the collection of terms by adding and subtracting, an understanding of syntax in verbal descriptions, application of number facts.<br>*Mean difficulty: -1.317* | 6 | $5a - 2b + 3a + 3b$ | 0.33 | 37 | Solve:<br>$5a - 4 = 2a + 8$ | 1.22 |
| | 5 | $2ab + 3b + ab$ | -1.98 | 29 | What is $t$ if<br>$2t$-23 = 49? | -0.52 |
| | 1 | $3m + 8 +2m -5$ | -2.53 | 26 | If $x + 5 = 7$,<br>then $x$ = ? | -3.14 |
| | 2 | $5p - p + 1$ | -2.6 | | | |

The high difficulty estimate for Item 37 could be attributed to the large number of non-attempts (119), coded as incorrect, a number was almost equal to that of Item 34 [$x + (x + 2) = (x - 1) + 8$] for which there were 120 non-attempts, also coded as incorrect (see Chapter 5, Table 5.1). It should be noted that Item 5 [$2ab + 3b + ab$] did not fit the model as expected; some other factors contributed to the correct responses of students whose ability estimates were lower than the difficulty estimate for this item. Interview data, together with that from the error analysis, suggests possible reasons for the misfit of Item 5 and of Item 12, (Category 2: Fractions, Table 4.5) noted earlier in the overview of the results of Rasch modelling of the survey data.

### *Category 5: Multiplication*

Items in this category are described and listed in Table 4.8, together with the survey item number and Rasch difficulty estimate. All items with brackets (implied multiplication) had a greater difficulty estimate than items that involved simple, one-step multiplication procedures. In the cases of equations (Item 32 [$10y = 5$] and Item 28 [$4y = 20$]), the multiplication involved was in the appearance of the equation. *Algebraically*, the solution required the process of division. Students were able to answer Item 32 and Item 28 using multiplication facts, rather than the multiplicative inverse of dividing the left hand side. Although structurally similar, Item 32 was much more difficult than Item 28. This may be attributed to the fact that the solution to the equation was one half, rather than a positive integer.

*Table 4.8: Items categorised as Multiplication in descending order of difficulty*

| Description | Expressions | | | Equations | | |
|---|---|---|---|---|---|---|
| | Item Number | Item | Rasch difficulty estimate | Item Number | Item | Rasch difficulty estimate |
| Expressions requiring simplification by multiplying of terms (no brackets) or equations where the initial relationships are expressed as multiplications and which therefore require solution by one-step division.<br>*Mean difficulty: -1.472* | 13 | $4r \times 5t \times 3$ | -0.34 | 32 | If $10y = 5$,<br>$y = \ldots?$ | -0.34 |
| | 9 | $2ab \times a$ | -1.08 | 28 | If $4y$ =20,<br>Then $y$ =? … | -2.33 |
| | 4 | $4 \times 5b$ | -3.27 | | | |

Item 39 [$ax = 5$] was also structurally the same as Items 32 and 28 and so might be considered as being in the same category. However, the solution to Item 39 could not have been found by students simply recalling multiplication facts. Instead they needed to recognise that the inverse procedure of dividing was needed, and that the solution was not unique (see above discussion of Category 3: Semi-literals), and involved non-integers. The expressions in this category were relatively less difficult than many items, except for Item 13, an unfamiliar example because it contained three terms to be multiplied.

### *Differences between categories*

Differences in level of difficulty between Categories of items were established by finding the mean item difficulty for all items in each category, and then testing for significance between the means, using *t*-tests. The average difficulty for the categories of Brackets and Fractions differed significantly from the average difficulties for the Addition and Subtraction, and Multiplication categories. (Brackets compared with Addition and Subtraction: $t$=20.68, $p$<0.01; Brackets compared with Multiplication: $t$=20.15, $p$<0.01; Fractions compared with Addition and Subtraction: $t$=20.19, $p$<0.01; Fractions compared with Multiplication: $t$=9.83, $p$<0.01). Table 4.9 lists the categories, the number of items, the range of difficulty estimates, and the mean difficulty estimate for each category.

*Table 4.9: Categories of items, with numbers of items in each, difficulty range, and mean difficulty estimates*

| Category | Number of Items | Difficulty range (logits) | Category mean |
|---|---|---|---|
| Brackets | 9 | -0.27 to 3.56 | 1.043 |
| Semi-Literal | 6 | -3.02 to 3.09 | -0.235 |
| Multiplication | 5 | -3.27 to -0.34 | -1.472 |
| Fractions | 10 | -0.93 to 3.17 | 0.928 |
| Add/Subtract | 10 | -2.53 to 1.22 | -0.994 |

The differences between the mean difficulty estimates of the Semi-literal category and those of the categories of Brackets and Fractions were not significant (Semi-literal compared to Brackets: $t$=10.6, $p$=0.08; Semi-literal compared to Fractions: $t$=12.24, $p$=0.105), nor were the differences between the means of the Brackets and Fractions categories *(*$t$=1.09, $p$=0.43).

These data suggest items, such as those containing brackets and/or fractions, are more difficult than items that are less syntactically complex. Less difficult items are similar in structure to arithmetic expressions with which students would have been familiar. Items with fractions or brackets required students to interpret the notation, and syntax, and use more conceptually complex mathematical procedures.

### Items that did not fit the model

To fit the Rasch model, items must lie within the infit mean square values of 0.77 to 1.3. Three items did not fit the Rasch model of the survey responses. They were Item 5 [$2ab + 3b + ab$], Item 12 [$2/a \times 3/b$] and Item 30 [If $x/4 = 12$, what is $x$?] (Figure 4.1).

Response patterns to Item 5 and Item 12 were more haphazard than expected (underfit). Item 5 has a 44% greater variation in response patterns than expected (Infit mean square: 1.44). Item 12 had 35% more variation in response patterns than expected (infit mean square: 1.35). The response patterns thus indicate that a substantial number of students answered these items correctly, although the ability estimates of these students indicate that they would not be expected to do so. Also, a number of students with higher ability estimates failed to answer these items correctly, although they could have been expected to do so. Underfit of items to the model suggests that conceptual factors other than those of algebraic reasoning are operating.

Item 30 fitted the model better than expected (overfit), with an infit mean square value of 0.70. This indicates that the response pattern to the item had 30% less variation than expected. The model expects that students with ability estimates with the same value as an item difficulty estimate have a 50% chance of responding correctly. Hence, students with ability estimates less than an item difficulty estimate would be unsuccessful and those with ability estimates greater than the difficulty estimate for an item would respond successfully. A Guttmann response pattern such as this indicates that the item discrimination is too sharp. The item, however, is not necessarily testing a different construct to that for which the instrument was designed.

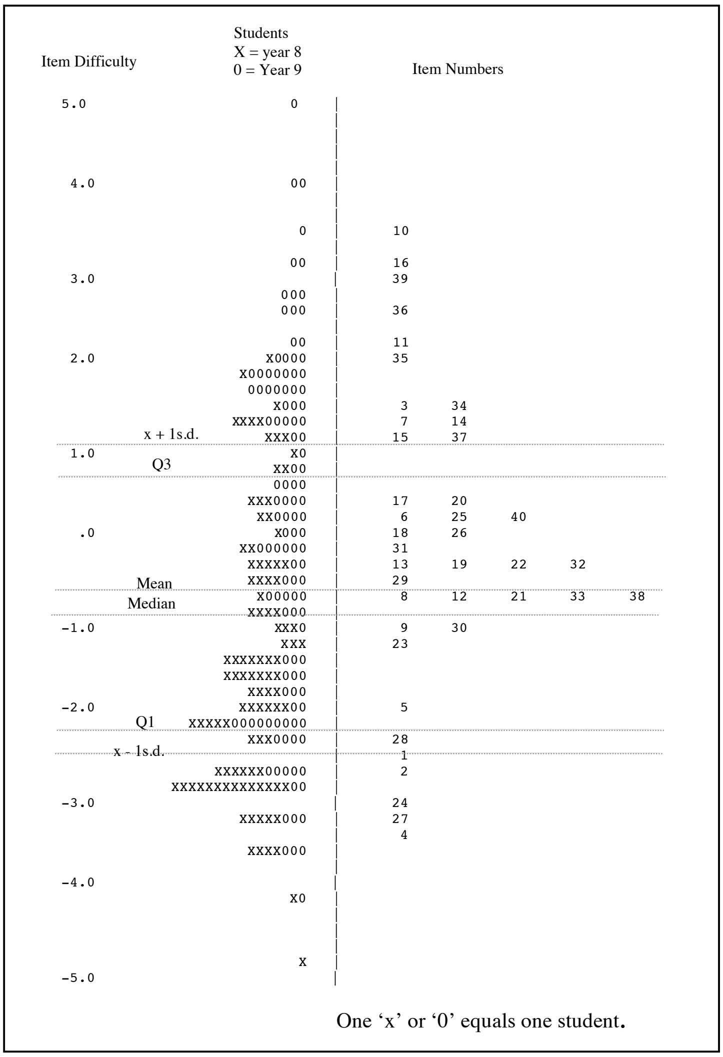
## Summary

Items were clustered about three significantly different means in a difficulty range of -3.27 to 3.56 logits. There was significant difference between mean difficulty estimates for Clusters 1 and 2, and between mean difficulty estimates for Clusters 2 and 3. These differences indicated that considerable conceptual change was necessary for students responding correctly to items in one cluster to being able to respond correctly to items in the next most difficult cluster (leaps of understanding). The differences in the mean difficulty estimates for sub-clusters 3a, 3b and 3c were also significant. The intervals between the sub-clusters thus indicate possible points where leaps of understanding would be needed for students to correctly respond to items within the large cluster.

The mean difficulty estimates for the various categories of items showed significant differences between two categories of items – those with brackets and those with fractions – and other categories consisting of items to be added, subtracted or multiplied, and the semi-literal items. Three items that did not fit the model were also discussed. Examination of responses to the two underfit items has the potential to provide useful clues to student understanding.

## PATTERNS OF STUDENT RESPONSES

Results from the Rasch modelling of item responses were complemented by the case estimates. In this instance, the case estimates were indicative of the abilities of the participating students as measured by their performance on the survey (test) items.



```
                          Students
                          X = year 8
Item Difficulty           0 = Year 9         Item Numbers

  5.0                              0  |
                                      |
                                      |
                                      |
                                      |
  4.0                             00  |
                                      |
                                      |
                                   0  |     10
                                      |
                                  00  |     16
  3.0                                 |     39
                                 000  |
                                 000  |     36
                                      |
                                  00  |     11
  2.0                          X0000  |     35
                            X0000000  |
                             0000000  |
                                X000  |      3     34
                           XXXX00000  |      7     14
            x + 1s.d.          XXX00  |     15     37
-------------------------------------------------------------------
  1.0                             X0  |
             Q3                 XX00  |
-------------------------------------------------------------------
                                0000  |
                             XXX0000  |     17     20
                              XX0000  |      6     25     40
   .0                           X000  |     18     26
                            XX000000  |     31
                             XXXXX00  |     13     19     22     32
             Mean            XXXX000  |     29
-------------------------------------------------------------------
             Median           X00000  |      8     12     21     33     38
                             XXXX000  |
-------------------------------------------------------------------
 -1.0                           XXX0  |      9     30
                                 XXX  |     23
                          XXXXXXX000  |
                          XXXXXXX000  |
                             XXXX000  |
 -2.0                       XXXXXX00  |      5
            Q1        XXXXX000000000  |
-------------------------------------------------------------------
                             XXX0000  |     28
          x - 1s.d.                   |      1
-------------------------------------------------------------------
                         XXXXXX00000  |      2
                    XXXXXXXXXXXXXX00  |
 -3.0                                 |     24
                            XXXXX000  |     27
                                      |      4
                             XXXX000  |
                                      |
 -4.0                                 |
                                  X0  |
                                      |
                                      |
                                   X  |
 -5.0                                 |

                                      One 'x' or '0' equals one student.
```

*Figure 4.7: Map of item difficulty and ability estimates, with students identified by years, quartiles and means (±1SD) marked.*

Student ability estimates (case estimates) and associated fit statistics are given in Table 4.2. The mean ability estimate was -0.64 with SD of 1.83 logits. Most ability estimates lay in the range 1.19 to -2.47 logits. Figure 4.7 above shows the mean ability estimate ± 1SD, together with quartiles (calculated from case estimates using Excel 2004 for Mac OSX, version 11.0) superimposed on the Quest-generated map (Adams & Khoo, 1996) of items and ability estimates.

Item difficulty estimates and ability estimates are reported on the same scale. Consequently three clusters of ability estimates align approximately with the three clusters of items. The median ability estimate (-0.93 logits) aligns nearly with the lower boundary of item Cluster 2 (-1.17 logits). The third quartile of ability estimates lies between item Clusters 2 and 3 (approximately 0.7 logits).

These clusters of student ability are identified, as whole groups, regardless of the school year, and described in detail in the following sections. First, the three groups are described; second, the groups are compared with the Clusters of items; and third, the groups are compared with respect to the Categories of items.

### Description of student ability groups

Student ability estimates form three groups: those less than -1.28 logits (Group 1: Low-ability); those between -1.1 and 0.93 logits (Group 2: Average-ability); and, those greater than 1.1 logits (Group 3: High-ability). The groups are numbered from the group of lowest ability estimates. The Low-ability group consisted of 102 students whose ability estimates range from -4.89 logits to -1.28 logits. The mean ability estimate for this group is -2.4 logits. The Average-ability group consisted of 68 students with a range of ability estimates from -1.11 logits to 0.93 logits and a mean ability estimate of -0.15 logits. The High-ability group included students with ability estimates ranging from 1.1 logits to 4.93 logits. This group had 52 students and a mean ability estimate of 2.00 logits (rounded). The data are summarised in Table 4.10, and illustrated in Figure 4.7.

*Table 4.10: Summary of ability group data with t-test of difference between group means*

| Ability Group | Group 1: Low ability | Group 2: Average ability | Group 3: High ability |
|---|---|---|---|
| Ability range (Logits) | -4.89 to -1.28 | -1.11 to 0.93 | 1.1 to 4.93 |
| Number in group | 102 | 69 | 52 |
| Group mean ability | -2.34 | -0.15 | 1.995 |

Differences in the mean ability estimates for each group are significant (p<0.005: High- and Average-ability groups: $t$=21.75, $p$<0.001; Average- and Low-ability groups: $t$=24.28, $p$<0.001).

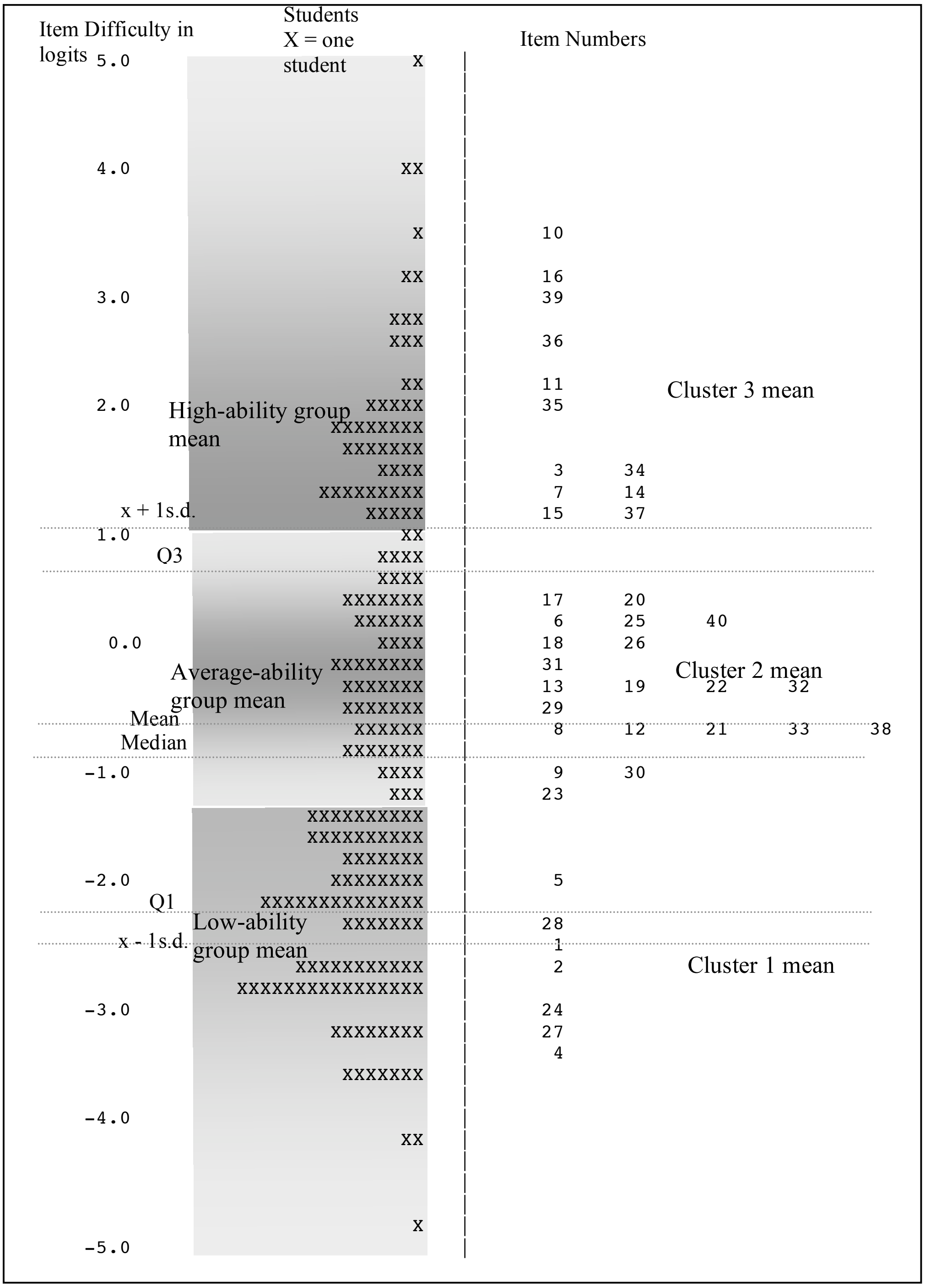


*Figure 4.8: Grouped student abilities. High-ability from 1.00 to 5.00 logits; Average-ability from -1.17 to 1.00 logits; Low-ability from -5.00 to -1.17 logits.*

two groups are concentrated near the boundaries of each of these groups with the Average-ability group. Figure 4.8 illustrates these distributions, with shaded sections on the Quest-generated map.

### Comparison of ability groups with item clusters

The three groups of ability estimates coincide with the three most significant clusters of items, described in the previous section. Figure 4.8 illustrates these data. Numbers on the far left of the map list the difficulty/ability estimates in logits. The graduated shading indicates the centres of the student ability estimates for the three groups of students, on the centre-left of the map. Each student is denoted with an “x”. The mean for the total number of participants is shown at -0.64 logits, as are the median and quartiles. Group means are also indicated. These means are at approximately the same points as the mean difficulties for the item clusters, shown on the right of the centre line. Item numbers are also listed.

The difference in the means between difficulty estimates for item clusters and ability estimates for the groups is not significant. The mean ability estimate for the Low-ability group was -2.34 logits. This compares with the mean difficulty estimate of -2.70 logits for the items in Cluster 1 ($t$=1.8, $p$=0.1). The mean ability estimate for the Average-ability group is -0.15 logits, and the mean difficulty estimate for items in Cluster 2 is -0.30 logits ($t$=1.72, $p$=0.1). The mean ability estimate for the High-ability group was 2.00 logits and the mean difficulty estimate for items in Cluster 3 was 2.09 logits ($t$=0.66, $p$=0.4).

### Comparison of ability groups with item categories

The categories of items most likely to be responded to successfully by each of the groups are illustrated in Figure 4.9. Individual student abilities, and the three significantly different groups of student abilities are mapped to items (identified by number) in Rasch difficulty order. The items are arranged into categories of brackets, fractions, semi-literal items, items requiring addition and subtraction, and multiplication. The mean ability estimates for each group is shown in the left-hand shaded areas, and the mean difficulty estimate for each category of items is marked, in each category column.

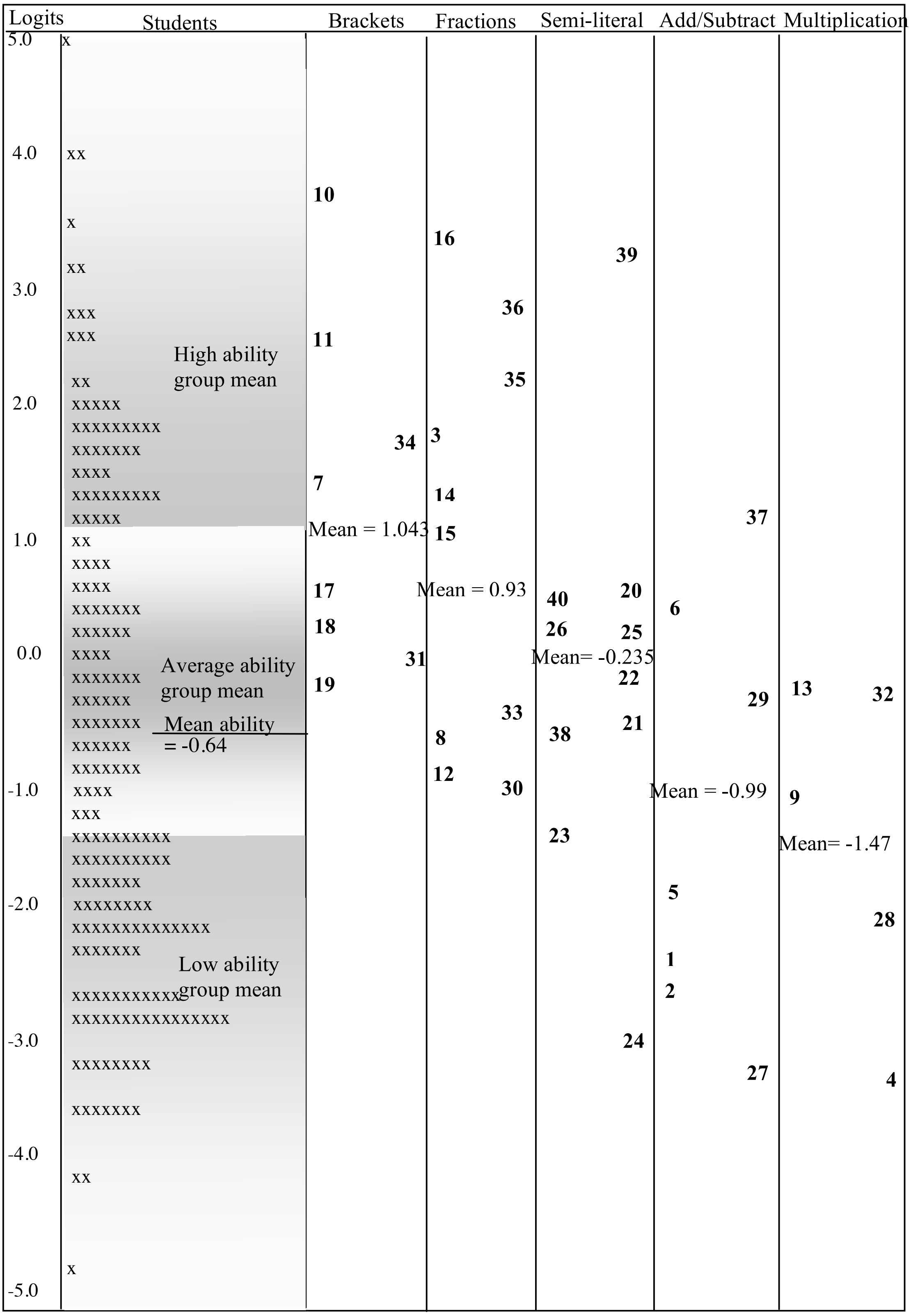


*Figure 4.9: Student ability groups mapped on to items arranged by difficulty and in categories.*

The Low-ability group has an ability range of -4.89 to -1.28 logits, and a mean of -2.34 logits. Only the upper bound of this ability range is approximately equal to the difficulty mean of the Multiplication category (-1.47 logits). The mean ability for the group is less than the difficulty means for all categories.

Some students, therefore, from the Low-ability group could be expected to be able to respond successfully to half of the items in the Addition and Subtraction, or Multiplication categories, and to Item 24 in the semi-literal category.

The upper bound for the Average-ability group (0.93 logits) is approximately equal to the mean difficulty for the Brackets category (1.04 logits), and greater than the mean difficulties for the other categories. However, ten items, nine in the Brackets or Fractions categories, and one in the Semi-literal category are of a difficulty greater than the upper ability of the Average group.

The mean ability for the High-ability group (1.995 logits) is greater than the difficulty means for all categories of items but is less than the difficulty estimates for the four most difficult items. No more than 12 students out of the 54 students in the High-ability group could be expected to respond successfully to these four items, two of which involve brackets, one is in fraction form, and the fourth, a semi-literal equation.

## Summary

In this section, the case estimates (student mathematical ability) established by the Rasch modelling of the survey responses have been described. Overall, the distribution of ability estimates is symmetrical. The mean student ability estimate of -0.64 logits was lower than the constrained item difficulty mean of 0.0 logits.

Within the distribution of case estimates three groups of abilities were identified. The boundaries of these groups coincided approximately with the boundaries of the item clusters. The differences between the means of these ability groups were statistically significant. The means of each of these ability groups did not differ significantly from the mean difficulty estimates for the three main clusters of items with which they aligned.

Comparison of the ability groups with the survey items arranged in categories showed the types of items that students in each of the groups could be expected to answer successfully. Patterns of distribution of the items according to type and difficulty estimates suggest that, overall, students found dealing with the items with brackets or with fractions the most challenging, and those requiring recognition and application of multiplication techniques the least. However, the range of difficulty in the categories of Brackets, Fractions and Semi-literal items means that some of these were accessible to students of ability near the overall mean (-0.64 logits)

## CONCLUSION

Analyses of results from the algebra survey conducted as part of the first phase of this research revealed several patterns. Statistics from Rasch modelling of the survey responses showed that the test was reliable and valid, measuring the one conceptual construct, that of basic symbolic algebra. Three items that did not fit the model have been identified and discussed.

When survey items were arranged in order of increasing difficulty estimates, three distinct clusters with significantly different mean difficulty estimates were apparent. The third cluster, of the most difficult items, also had three sub-clusters with significantly different means. When items were arranged according to syllabus topics (categories of items of related algebraic ideas), the mean difficulty estimates for each category indicated that items with brackets or fractions were significantly more difficult than items requiring manipulation by addition, subtraction or multiplication.

Student ability estimates also clustered around three significantly different means, which approximated the mean difficulties of the three main clusters of items. Comparison of the student ability groups with the item categories indicated types of items most likely to be answered successfully by students in each ability group. These comparisons suggest that differences in conceptual development might account for the existence of ability groupings of students, and the clusters of items.

The results suggest that: items in clusters, although they are of different types, share certain conceptual characteristics; items of different types (categories), might require different conceptual changes to be made by students, and that these are developmentally linked; and, ability clusters of students are associated with developmental stages of conceptual development (in algebra).

Examination of responses to the items and of explanations of their thinking given by a sample of students as they address the items should identify possible conceptual pathways. Reasons for erroneous responses and the consequent difficulty of items, and the subsequent lack of success for some, or many students, are needed for the development of a conceptual model of elementary algebra. These data, discussed in the following chapter, were obtained from the inspection of errors made in responses to the survey, and from interviews with students.

# CHAPTER 5: QUALITATIVE ANALYSIS OF SURVEY RESPONSES

Rasch-modelling of responses by students ($N$=222) resulted in clusters of items and students. Items clustered around three significantly different mean difficulty estimates. Students also clustered around three significantly distinct mean ability estimates. These have been labelled respectively *item difficulty clusters* and *student ability groups*. Increased item difficulty and increased student ability values imply a developmental progression.

Items in each of the clusters are psychometrically similar. Intervals between item clusters imply psychometric change; to deal with the items in each cluster requires new mathematical understanding on the part of students. As items become more difficult, students' mathematical understanding has to develop if the students are to deal successfully with those items.

An inference to be made from the grouping of students is that students in one group share common conceptualisations of the mathematics. Different groups of students therefore, understand the mathematics differently. Depending on the development of their mathematical understanding, students are able to respond successfully to items in increasingly difficult clusters.

The mean ability estimates for the three student ability groups aligned with the mean difficulty estimates for the three main clusters of items. The three student ability groups have been labelled Low-ability, Average-ability, and High-ability; and the clusters labelled as Cluster 1, Cluster 2, and Cluster 3. Three sub-clusters of items were identified in Cluster 3. The Low-ability group were aligned with item Cluster 1, the Average-ability group with item Cluster 2, and, the High-ability group with item Cluster 3.

In order to interpret the conceptual commonalities shared by items in the clusters, and students in each of the ability groups students' survey responses are analysed with respect to the types of responses, the errors made by students. This empirical analysis is supplemented by verbatim evidence given by a sample of students in interviews.

The first two sections of this chapter give an overview of the responses to the survey, first by describing types of responses to the survey, and secondly by discussing common errors made for individual items. The next three sections of the chapter discuss implications of the Rasch modelling of the survey responses. In these sections the three clusters of items are analysed, and also the three items that did not fit the model. The last section considers the responses based on three groups of student abilities.

## RESPONSE TYPES

Students' responses to the survey could be categorised as four types: non-attempts; incorrect answers; partially correct attempts; and correct answers. For the purposes of Rasch-modelling, the first three types of response were coded as incorrect. However, a more detailed analysis of response types provides information about concept development in the context of introductory algebra. The disinclination of students to make any recorded attempt on an item can contribute substantially to the 'difficulty' of an item, as does the proportion of students giving incorrect responses. This section describes each item in terms of the unsuccessful (non-attempts or incorrect responses) or successful (correct) responses made by all students. Items were marked as 'no attempt', 'incorrect' or 'correct'. The results are presented in Table 5.1.

*Table 5.1: Response type and numbers to survey items. Number of items = 40, total number of responses to each item = 222.*

| Item | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| No attempt | 2 | 1 | 27 | 7 | 4 | 9 | 32 | 38 | 35 | 24 |
| Incorrect | 50 | 49 | 152 | 26 | 65 | 136 | 142 | 74 | 63 | 186 |
| Correct | 170 | 172 | 43 | 189 | 153 | 77 | 48 | 110 | 123 | 12 |
| Item | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| No attempt | 35 | 37 | 48 | 79 | 58 | 77 | 37 | 36 | 40 | 40 |
| Incorrect | 158 | 77 | 76 | 95 | 111 | 129 | 112 | 103 | 86 | 109 |
| Correct | 29 | 108 | 98 | 48 | 53 | 16 | 73 | 83 | 96 | 73 |
| Item | **21** | **22** | **23** | **24** | **25** | **26** | **27** | **28** | **29** | **30** |
| No attempt | 20 | 20 | 52 | 33 | 23 | 46 | 30 | 39 | 60 | 59 |
| Incorrect | 96 | 104 | 44 | 6 | 118 | 91 | 6 | 19 | 58 | 45 |
| Correct | 106 | 98 | 126 | 183 | 81 | 85 | 186 | 164 | 104 | 118 |
| Item | **31** | **32** | **33** | **34** | **35** | **36** | **37** | **38** | **39** | **40** |
| No attempt | 89 | 65 | 63 | 120 | 95 | 137 | 119 | 78 | 104 | 88 |
| Incorrect | 42 | 59 | 53 | 57 | 96 | 62 | 51 | 38 | 101 | 53 |
| Correct | 91 | 98 | 106 | 45 | 31 | 23 | 52 | 106 | 17 | 81 |

Over all items the average number of non-attempts on any one item was 50; the average number of incorrect responses was 80; and, the average number of correct responses was 92 (out of 222 responses to each item). These responses are discussed in three sub-sections: non-attempts are discussed in the first; followed by a discussion of incorrect responses; and, the final sub-section discusses correct responses. Discussion of these data references clusters of items and ability groups of students.

### Non-attempts

From Table 5.1, it can be seen that the incidence of non-attempts increased for the last ten items. The number of non-attempts was greater than 84 for all of these items, which were printed on the last of the three pages of the survey. However, most students attempted some questions in the group of Items 29 to 40 (last page of the survey).

Numbers of non-attempts, were most commonly between 26 to 84, with the exception of the following items: Item 31 [Solve: $4(p + 3) = 32$], Item 34 [Solve: $x + (x + 2) = (x + 1) + 8$], Item 35 [If $x + x/3 = 4$, what is x?], Item 36 [If $63/x = 180$, what is x?], Item 37 [Solve: $5a - 4 = 2a + 8$], Item 39 [If $ax = 5$, then $x =$…?], and Item 40 [If $e + f = 8$, then $e + f - g =$…?], where the number of non-attempts was 89 or greater. Many students appeared to be unfamiliar with these equations, which were difficult for students to solve if they relied on arithmetic procedures, such as trial-and error substitutions, instead of using an algebraic procedure.

With reference to the Rasch model, Items 34, 35, 36, 37 and 39 had difficulty estimates greater than 1.18 logits. These items were all within Cluster 3 (the cluster of most difficult items). Items 31 and 40 had difficulty estimates of -0.12 and 0.2 logits respectively. These values are at or above the mean difficulty estimate of Cluster 2. Therefore, many items for which there were the greatest number non-attempts were those scored as most difficult. However, the two most difficult items did not attract a great number of non-attempts, but many incorrect responses.

## Incorrect responses

Incorrect responses are discussed in detail in the following section where the most common errors made by students are described. This sub-section gives a brief summary of items that were most frequently answered incorrectly.

Items where the number of incorrect responses was greater than 112, were Item 3 [$a/5 + a/10$], Item 6 [$5a - 2b + 3a + 3b$], Item 7 [$(a - b) + b$], Item 10 [$(x + y)^2$], Item 11 [$8p - 2(p + 5)$], Item 16 [$2/a^2 \times 5a/4$], and Item 25 [Take $n$ away from $3n + 1$]. All of these items had a difficulty estimates greater than the mean difficulty of 0 logits. Two items (Item 6 and Item 25) were in Cluster 2, the others in Cluster 3.

Most students attempted to answer two of the three most difficult items, Item 10 [$(x + y)^2$], Item 16 [$2/a^2 \times 5a/4$], but most did so incorrectly. Item 11 [$8p - 2(p + 5)$] and Item 7 [$(a - b) + b$] also attracted many incorrect responses, and relatively few non-attempts.

## Correct Responses

When the numbers of correct responses were mapped against the clusters of items it could be seen that 13% of students in the Low-ability group responded correctly to items in Cluster 2. Even fewer Low-ability students correctly answered any items in Cluster 3. The numbers of students in the Average-ability group correctly responding to items in Cluster 3 also dropped markedly. Although the numbers of students in the High-ability group correctly responding to items dropped from 87% in Cluster 2 to 67% in Cluster 3a, 32% only were able to give correct responses to items in the upper part of Cluster 3 (Clusters 3b and 3c). This indicates that items in sub-clusters 3b and 3c, presented conceptual difficulties to some students in the High-ability group.

Table 5.2 gives these results in terms of a percentage correct response by each ability group to items in a Cluster. This was calculated as the total number of correct responses to items in the cluster by students in an ability group out of the total number of responses by the students in each ability group, expressed as a percentage to the nearest whole number.

*Table 5.2: Average percent correct responses by ability groups to each cluster of items*

| Ability Groups | | | Clusters | | | |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3a | 3b & 3c |
| Group | Number in Group | Average ability estimate of group | Average difficulty estimates for clusters | | | |
| | | | -2.70 | -0.30 | 1.34 | 2.82 |
| HIGH | 52 | 2 | 97 | 87 | 67 | 32 |
| AVERAGE | 69 | -0.15 | 88 | 55 | 17 | 5 |
| LOW | 101 | -2.34 | 62 | 13 | 3 | 2 |

This can be illustrated by a specific example. Cluster 1 consisted of Items 4, 27, 24, 2, 1, 28 and 5, in increasing order of difficulty. Of the 52 students in the High-ability group, the numbers of students correctly responding to each item were respectively, 51, 52, 52, 51, 52, and 44. In all, there were 354 correct responses out of a total of 364 responses. As a percentage, this is approximately 97%. Similar calculations were made for each item cluster, addressed by each of the ability groups.

Correct responses by ability groups to categories of items were also examined. Percentages of participants in each ability group correctly answering the different categories of items are summarised in Table 5.3. The percentage of correct responses drops for all categories of items as one reads vertically from the group with the highest average ability estimate (High-ability group) down to the group with the lowest average ability estimate (Low-ability group).

*Table 5.3: Average percent correct responses by ability groups to each category of items*

| Ability Groups | | | Item Categories in increasing order of average difficulty | | | | |
|---|---|---|---|---|---|---|---|
| Group | Number in Group | Average ability estimate of group | Multiplication | Addition/ Subtraction | Semi-Literal | Fractions | Brackets |
| | | | Average difficulty estimate for each Cluster | | | | |
| | | | -1.47 | -1.32 | -0.17 | 0.93 | 1.12 |
| HIGH | 52 | 2 | 94 | 87 | 82 | 67 | 67 |
| AVERAGE | 69 | -0.15 | 73 | 66 | 51 | 33 | 30 |
| LOW | 101 | -2.34 | 35 | 40 | 17 | 8 | 4 |

Item categories are read from left-to-right, arranged according to average difficulty estimates. As the average difficulty for each category increases, the proportion of correct responses declines. There was one exception. The proportion of correct responses from students of Low-ability is greater in the Addition/Subtraction category than for the Multiplication Category, although the former is slightly more difficult. These numbers have contributed to the misfit of Item 5 [$2ab + 3b + ab$] to the Rasch model.

The percentage of correct responses to the Semi-Literal items by Low-ability and Average-ability students dropped sharply from the percentage of correct responses to items in the two less difficult categories. Percentages of correct responses from the High-ability group to this category of items dropped slightly. However, the percentage of correct responses by the High-ability group to items in the categories of Fractions and Brackets has dropped more sharply. This decline is even more marked in the case of responses from students in the Low- and Average-ability groups.

Change in the percentages of correct responses indicates that particular types of items require students to reconceptualise hitherto-understood (useful) mathematics. Also, the conceptual limitations of different ability groups are indicated by the change in percentage of correct responses to different categories of items.

## Summary

Items that attracted the greatest number of unsuccessful attempts had the highest difficulty estimates. However, magnitude of the difficulty estimates, particularly that of the two most difficult items, was also a result of the numbers of incorrect responses, rather than the numbers of non-attempts.

The items attracting the greatest numbers of non-attempts were equations, or expressions involving fractions. These data suggest that students will attempt items that, in the instances of equations, can be solved readily by simple number recall or trail-and-error procedures. In the instances of expressions needing to be manipulated, the items need to be examples presented in a previously encountered (familiar) form, or contain a visual cue to trigger a rehearsed or remembered response. The numbers, and types, of incorrect responses suggest that students operating on such visual cues work with imperfect, or oversimplified procedural rules. Rules, and misunderstanding of rules, can be identified by examination of errors made in response to items on the survey.

## ANALYSIS OF ERRORS

In this section, patterns of errors made by students in their responses to items in the survey are identified and discussed. Different incorrect responses were identified together with the frequency of occurrence of each error. The occurrence of typical errors or patterns of responses indicated commonly held misunderstandings across the ability range of the participants. Table 5.4 lists each survey item, together with its Rasch difficulty estimate (R) and the correct response. The most common errors, the frequency of those errors, together with the total number of incorrect responses (excluding non-attempts) for each survey item are summarised in the table.

Numbers of incorrect responses do not include numbers of non-attempts. Items with low difficulty estimates attracted few incorrect responses, but these numbers need to be read in conjunction with the numbers of non-attempts given in Table 5.1, and conclusions drawn from that data.

Errors made by students in responding to the most difficult and the least difficult items are discussed in the first two sub-sections of this chapter. A third sub-section discusses errors made in the various categories of items: Multiplication; Addition and Subtraction of terms; Semi-literal items; Fractions; and items containing Brackets (Figure 4.5, Tables 4.3 to 4.7). This sub-section also includes discussion of errors made as students attempted to solve equations representative of each category.

### The two most difficult items

Of the two most difficult items, one involved brackets and the other was a fraction to be simplified. As noted above, although Item 10 [$(x + y)^2$, difficulty estimate: 3.56 logits] and Item 16 [$2/a^2 \times 5a/4$, difficulty estimate: 3.16 logits] were the most difficult, most students attempted to provide a response. Only 22 out of the 222 participants offered no response to Item 10, a number considerably fewer than the average of 50 non-attempts per item. However, 77 students did not answer Item 16

### *Table 5.4: Summary of error-type and frequency by items*. **REPLACED WITH A3 VERSION**

The items are arranged by item number as in the survey. The ‘R’ value is the item difficulty estimate from the Rasch model. The item and the correct answer are listed. Only the most common errors are listed, together with the frequency. Total incorrect accounts only for incorrect responses, not for non-attempts.

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Item 1** | **R = -2.53** | **Item 2** | **R = -2.6** | **Item 3** | **R = 1.59** | **Item 4** | **R = -3.27** | **Item 5** | **R = -1.98** | **Item 6** | **R= 0.33** | **Item 7** | **R = 1.38** | **Item 8** | **R = -0.64** |
| *3m + 8 + 2m – 5 = 5m + 3* | | *5p – p + 1 = 4p - 1* | | *a/5 + a/10 = 3a/10* | | *4 x 5b = 20 b* | | *2ab + 3b + ab = 3ab + 3b* | | *5a – 2b + 3a + 3b = 8a + b* | | *(a – b) + b = a* | | *4ab/4b = a* | |
| Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors |
| 13 | 50 | 13 | 49 | 27 | 153 | 10 | 26 | 33 | 65 | 32 | 136 | 26 | 142 | 25 | 74 |
| Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency |
| 5m - 3 | 16 | (5 + 1 =) 6 | 20 | a/15 | 42 | 5b x 4 | 7 | 3a5b | 6 | 8a - 5b | 38 | a - 2b | 36 | 4a | 12 |
| 8m | 17 | 5p | 13 | 2a/15 | 33 | 4 x 5b | 5 | 6ab | 6 | 2a + 5b | 29 | $a - b^2$ | 16 | 4ab/4b | 11 |
| 5m3 | 6 | 6p | 5 | 2a/10 | 8 | | | | | 8a + 5b | 12 | a + b | 16 | ab | 11 |
| | | | | 2a + 15 | 8 | | | | | 8a - b | 9 | ab | 13 | | |
| **Item 9** | **R = -1.08** | **Item 10** | **R = 3.56** | **Item 11** | **R = 2.28** | **Item 12** | **R = -0.7** | **Item 13** | **R = -0.34** | **Item 14** | **R = 1.38** | **Item 15** | **R = 1.18** | **Item 16** | **R = 3.17** |
| $2ab\ x\ a = 2a^2 b$ | | $(x + y)^2$ | | *8p – 2(p + 5)* | | *2/a x 3/b = 6/ab* | | *4r x 5t x 3 = 60rt* | | *x/3 ÷ y/4 = 4x/3y* | | *3p/4 - p/8 = 5p/8* | | $2/a^2\ x\ 5a/4 = 5/2a$ | |
| Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors |
| 17 | 63 | 20 | 186 | 59 | 158 | 21 | 77 | 24 | 76 | 49 | 95 | 32 | 111 | 45 | 129 |
| Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency |
| 2ab x a | 23 | $x^2 + y^2$ | 86 | 6p + 10 | 40 | 6ab | 24 | 4r x 5t x 3 | 31 | xy/1 or xy | 15 | 2p/4 | 33 | $10a/4a^2$ | 52 |
| 3ab | 14 | $xy^2$ | 49 | 6p + 5 | 17 | 5/ab | 14 | 20rt x 3 | 9 | xy/12 | 10 | 2p/-4 | 16 | 10/4a | 10 |
| 2ab | 10 | $x^2y^2$ | 14 | 9p + 3 | 12 | 2b/3a | 6 | | | | | 2p - 4 | 12 | | |
| | | | | $8p^2 - 10$ | 10 | | | | | | | | | | |
| **Item 17** | **R = 0.46** | **Item 18** | **R = 0.13** | **Item 19** | **R = -0.27** | **Item 20** | **R = 0.46** | **Item 21** | **R = -0.58** | **Item 22** | **R = -0.34** | **Item 23** | **R = -1.17** | **Item 24** | **R = -3.02** |
| $(6xy)^2 = 36x^2 y^2$ | | *2(x + 4) + 3(x - 1) = 5x + 5* | | *2(x + 5) - 8 = 2x + 2* | | *Multiply x + 5 by 4 = 4(x + 5) or 4x + 20* | | *Add 4 on to n + 5 = n + 9* | | *Add 3 onto 4n = 4n + 3* | | *What can you say about m if m = 3n + 1 and n = 4 ? m = 13* | | *If a +b = 43, then a + b + 2 = …? 45* | |
| Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors |
| 20 | 112 | 60 | 103 | 39 | 86 | 24 | 109 | 20 | 96 | 7 | 104 | 18 | 44 | 4 | 6 |
| Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency |
| 36xy | 34 | 8x | 7 | 15x | 17 | 20x | 58 | 9n | 54 | 7n | 95 | m = 4n | 7 | 4b | 3 |
| 12xy | 18 | 2x + 8 | 7 | 2x + 13 | 9 | x + 20 | 7 | 4n + 5 | 10 | | | m = n | 5 | | |
| $6xy^2$ | 15 | 5x + 11 | 7 | | | 9 + x | 6 | | | | | m = 8 | 5 | | |
| $36xy^2$ | 11 | | | | | | | | | | | | | | |
| $12x^2y^2$ | 11 | | | | | | | | | | | | | | |
| **Item 25** | **R = 0.2** | **Item 26** | **R = 0.07** | **Item 27** | **R = -3.14** | **Item 28** | **R = -2.33** | **Item 29** | **R = -0.52** | **Item 30** | **R = -0.93** | **Item 31** | **R = -0.12** | **Item 32** | **R = -0.34** |
| *Take n away from 3n + 1; 2n + 1* | | *If p + q = 5, then p + q + r =…? 5 + r* | | *If x + 5 = 7, then x = …? x = 2* | | *If 4y = 20, then y = …? y = 5* | | *What is t if 2t - 23 =49? t= 36* | | *If x/4 = 12, What is x? x = 48* | | *Solve 4(p + 3) = 32; p = 5* | | *If 10y = 5, then y = …? y = 0.5* | |
| Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors |
| 12 | 118 | 14 | 91 | 1 | 6 | 5 | 19 | 17 | 58 | 11 | 45 | 21 | 42 | 9 | 59 |
| Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency |
| (3 + 1=) 4 | 99 | 5r | 34 | x =3 | 6 | y =16 | 9 | t = 13 | 20 | x = 3 | 32 | p = 7.25 | 11 | y = 2 | 23 |
| 3n | 7 | 9 | 23 | | | | | t = 6 | 9 | | | | | y = 5 | 20 |
| | | 7.5 | 10 | | | | | | | | | | | y = -5 | 8 |
| **Item 33** | **R = -0.58** | **Item 34** | **R = 1.51** | **Item 35** | **R = 2.17** | **Item 36** | **R = 2.64** | **Item 37** | **R = 1.22** | **Item 38** | **R = -0.58** | **Item 39** | **R = 3.09** | **Item 40** | **R = 0.2** |
| *(x + 3)/2 = 7; What is x equal to? x = 11* | | *x + (x + 2) = (x - 1) + 8; x = 5* | | *x + x/3 = 4; x = 3* | | *If 63/x = 180, what is x? x = 9/20 or 0.35* | | *Solve 5a - 4 = 2a + 8* | | *If r - 82 = 7, then r - 83 = …? = 6* | | *If ax = 5, then x = …? x = 5/a* | | *If e = f = 8 then e + f - g = …? 8 - g* | |
| Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors | Nº Different Errors | Total Nº Errors |
| 12 | 53 | 25 | 57 | 22 | 96 | 34 | 62 | 31 | 51 | 9 | 38 | 16 | 101 | 27 | 53 |
| Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency | Most Common Errors | Frequency |
| x = 4 | 22 | x = 4 | 6 | x = 6 | 42 | x = 11340 | 7 | 3a + 12 | 4 | 8 | 27 | x = 5 | 24 | 7 | 8 |
| x = 2 | 10 | | | x = 1 | 15 | x = 117 | 6 | | | | | x = 2.5 | 23 | 1 | 7 |
| | | | | | | x = 3 | 5 | | | | | x = 1 | 13 | 4 | 5 |
| | | | | | | | | | | | | x = 5 - a | 9 | | |

(see Table 5.1), indicating that this was a much less familiar expression than the mathematically similar Item 8 [$4ab/4b$]. This item was answered successfully by more students. The high difficulty estimates for both Items 10 and 16 are therefore a consequence of large numbers of incorrect responses. The number of different incorrect responses to Item 10 was 20. The most common response was to state that $(x + y)^2 = x^2 + y^2$ (n=86, out of a total 186 incorrect responses). On the other hand, Item 16 with a total number of incorrect responses of 129 produced 45 different incorrect responses. Many of these were incomplete attempts to simplify the expression by students relying on visual cues to prompt their cancelling of the same number. These visual cues were available in Item 8, although many errors in this item also resulted from students' incomplete simplification of the expression.

### The least difficult items

The least difficult items included a simple multiplication and a simple one-step equation. Of all item categories, that which included equations or expressions involving multiplication had the lowest mean difficulty estimate (Category mean: -1.472 logits, Figure 4.5, Table 4.8). Item 4 [$4 \times 5b$] was the easiest item (-3.27 logits). There were only seven non-attempts and 26 incorrect responses. Of those 26, there were ten different errors. The most common incorrect responses were to leave the expression unchanged or to reverse the digits in the expression, i.e., $4 \times 5b = 4 \times 5b$ (5 out of 26 responses) or $4 \times 5b = 5 \times 4b$ (7 out of 26 responses). Similar types of errors, where the expression was not modified at all or was incompletely dealt with, were also made in response to Item 16 [$2/a^2 \times 5a/4$], Item 8 [$4ab/4a$] and Item 13 [$4r \times 5t \times 3$]. Item 27 [If $x + 5 = 7$, then $x = \ldots ?$, difficulty estimate: -3.14 logits) and Item 24 [If $a + b = 43$, then $a + b + 2 = \ldots ?$, difficulty estimate: -3.02 logits] had 30 and 33 non-attempts respectively. There were only six incorrect responses to Item 27, all of which offered $x = 3$ as the response. Item 24 also attracted six incorrect responses, four of which were different. The most common response (3 out of 6 responses) was to state that $a + b + 2 = 46$.

## Errors in the categories of items

In this sub-section, some of the most common errors made by students are identified and discussed in the different categories of items. The categories, organised in order of increasing average difficulty estimates are expressions and equations that principally involve: multiplication, addition and subtraction of terms, semi-literal items, fractions, and brackets. A further examination of errors is discussed in terms of student responses to the equations in the survey.

### *Multiplication*

The Multiplication category of items (Table 4.8) had the lowest mean difficulty estimate (-1.472 logits). All items in this category were of a difficulty equal to or less than the mean difficulty estimate for all items in the survey (constrained at 0 logits). It should be noted that Item 39 (If $ax = 5$, then $x =$…?) was not included in this category, but was considered more appropriately placed as a semi-literal item. The most common error made by students for Item 28 [If $4y = 20$, then $y =$…?] was to state that $y = 16$ (9 out of 19 incorrect responses), indicating a misinterpretation of the notation $4y$ as incorrectly meaning $4 + y$. The errors for Item 32 [If $10y = 5$, then y =…?] were more numerous, as were the numbers of non-attempts. The most frequent errors were to state that $y = 2$ (23 out of 59 responses) or $y = 5$ (20 out of 59 responses). Many students had difficulty with the concept that multiplication could result in a smaller number.

### *Addition and subtraction of terms in expressions*

The Addition and Subtraction category of items (Table 4.7) had the next highest difficulty estimate (-0.994 logits) to that for the category of multiplication items. The items in this category included some of the least difficult expressions, with the exception of Item 6 [$5a$ -$2b + 3a + 3b$]. The errors indicate that the way in which a mathematically similar idea is expressed could influence responses by students.

Item 2 [$5p - p + 1$, difficulty estimate: -2.6 logits] and Item 25 [Take $n$ away from $3n + 1$, difficulty estimate: 0.2 logits] appear on the surface to require the same mathematical understanding. The difficulty estimates suggest that students found Item 25 more difficult, as they did for all the semi-literal items. One student only did not attempt Item 2. There were 13 different responses out of 49 incorrect responses. Of these, the most frequent error (20 out of 49 responses) was to remove the letter *p* to give the answer as either "5 + 1" or "6". For Item 25, there were 118 incorrect responses (23 non-attempts) with 99 of those offering the answer of either "3 + 1", or "4".

Both these responses indicate that students have difficulty understanding the mathematical meaning of the notation. There seems also to be connection with the tendency to *conjoin* terms. Where an expression such as "$5p + 1$" is simplified to "$6p$", the same logic would have the expression "$5p - p + 1$" simplified to "5 + 1" or "6". Interview data also suggest that students understand a single letter (e.g., *p*) as having a value of 1, although, in this example, the student obtained the correct answer:

> Question 2: it'd be $5p$ minus $p$ because they're both similar terms, would be $4p$, 'cause $p = 1$, $4p$ plus 1.

The need to close, by providing an unique answer, resulting in inappropriate conjoining of terms, identified by authors such as Matz (1982), was typical of incorrect responses to Item 1[$3m + 8 + 2m -5$] and Item 2 [$5p - p + 1$], and to those items in the semi-literal category (i.e., Item 20 [Multiply $x + 5$ by 4], Item 21 [Add 4 on to $n + 5$] and Item 23 [What can you say about $m$ if $m = 3n + 1$ and $n = 4$?] (Table 5.2).

Item 5 [$2ab + 3b + ab$], although mathematically similar to Item 6 [$5a -2b +3a +3b$], prompted a wide variety of errors. Out of 65 incorrect responses, there were 33 different responses. The two most common errors were to give the simplification as "$3a5b$" (6 instances out of 65) or "$6ab$" (6 instances out of 65). The introduction of the terms containing *ab* seems to have been a source of confusion for some students and so prompted a multiplication response.

Item 5, at a difficulty level of -1.98 logits, shares many characteristics with others in this group, but requires a more subtle appreciation of surface similarity, or dissimilarity, between the terms *ab* and *b*. The fit diagram from the Rasch model (Figure 4.2) shows that this item was often answered correctly by students who were in the lower ability range. These students had ability estimates less than the difficulty estimate for the item. However, students for whom ability estimates were at the same difficulty level as Item 5, were not as successful. Interviews suggest that students who used superficial (visual) procedures to manipulate the expression were more likely to answer Item 5 correctly.

One student (Low-ability) who gave the correct answer to the item explained:

> I'd circle the ones with the same pronumerals and then I'd do a different circle for the ones that were different to the ones I've already circled.

On being asked why the student would circle the pronumerals, the reply was:

> Because that's what I was taught in maths…

This student also answered Item 6 [$5a - 2b + 3a + 3b$, difficulty estimate: 0.33 logits] correctly, by manipulating it in the same way. The naivety of the strategy seems to have afforded some protection from confusion about like terms for this student, and for other students who also thought in a similar way.

The incorrect responses to Item 6 indicated that students most commonly tended to "lose track of the signs", as one student confessed during the interview. Of 136 incorrect responses, 38 responses were "$8a - 5b$", and 29 responses were "$2a + 5b$".

### *Semi-Literal Items*

Items in the Semi-literal category (Table 4.6) presented students with examples that were conceptually similar to some of the least difficult items in the survey (addition, subtraction and multiplication). Because of the form in which the items were written, the errors made indicated that students of Low- or Average-ability had little understanding of the mathematical meaning behind the items, or of the syntactical requirements of algebra.

Only students in the High-ability group read out the statements as requested and then gave at least one simplified or alternative version of the statement. Students in the lower two ability groups would read the question aloud and simply repeat the statement. Some students would change the order of the numbers or letters (e.g., Student reads the question, and then responds: "Add 4 onto *n* plus 5. It would be *n* plus 5 plus 4”). Many of the errors made were those where students conjoined terms inappropriately. Thus, the statement “Add 4 on to $n + 5$” was restated as “$9n$”.

Item 39 [$ax = 5$; difficulty estimate: 3.09 logits] contributed greatly to the average difficulty of this category of items. Only 17 students out of the total 222 students correctly answered the item. Many errors indicated that students could not go beyond a numerical response, struggling with the existence of both $a$ and $x$. Some students suggested that both letters had to represent the number 2.5, possibly (mis)interpreting “$ax$” as “$a + x$”. Others suggested that either $a$ or $x$ had to be one of the factors of five. Even when some students recognised, during the interviews, that other numbers (or any number) might be appropriate, they preferred to give an integer response, because “it was easier”.

The erroneous responses to these semi-literal items suggest that students are familiar with symbolic forms of expressions only (e.g., $4n + 4 +n$), and manipulate these according to a set of simple rules. Students appeared to be unfamiliar with mathematically similar items stated in words and symbols, the manipulation of which required an understanding of concepts represented by the symbols. Students also needed to connect the words in the question with conventional algebraic symbolisation. The expression or equation could then be manipulated in the manner to which the students had been trained. Where this connection could not be made, students’ errors indicated the limits of their understanding.

### *Fractions*

All items in the Fraction category (Table 4.5) were at a difficulty level approximately equal to the mean ability level of the participants (-0.64 logits), or greater (mean difficulty of faction items: 0.928 logits). This category consisted of expressions or equations that included one or more terms with a vinculum (i.e., Items 12, 8, 15, 14, 3 and 16, which were expressions, and Items 30, 33, 35 and 36, which were equations). Items requiring addition or subtraction of fractions attracted a number of incorrect responses as well as a considerable variety of common responses and demonstrated well-documented misconceptions about arithmetic operations on fractions (e.g. Stacey, Helm, & Steinle, 2001).

Item 3 [*a*/5 + *a*/10, difficulty estimate: 1.59 logits] attracted 153 incorrect responses, of which there were 27 varieties. Of all the incorrect responses, 42 stated that the answer was "*a*/15", and 33 gave the result as "2*a*/15". The next most frequent numbers of responses of "2*a*/10" and "2*a* + 15" were each given by 8 students. The responses to Item 15 [3*p*/4 – *p*/8] demonstrated similar thinking on the part of participants.

The relatively high level of difficulty of the items involving fractions (Items 3, 15 and 14) may be attributed to the difficulties and misconceptions students have with operating on arithmetic fractions.

> It's just that, um, my fraction skills are absolutely terrible to be frank; but usually I just do the like terms again in fractions. Sometimes...[5] I can't really do this on a calculator, but we can use calculators. I get really confused with these ones. Yep, I've got a couple of them right before, just by chance...

The attitude towards fractions expressed by this student of near average ability (-0.17 logits, average ability: -0.64 logits) was echoed by other students. In these items (Items 3, 15 and 14), the letters may still be ignored and arithmetic procedures suffice to manipulate the numbers. When asked what he would do with Item 3 [*a*/5 + *a*/10], a student answered with the following, common response:

> Well, firstly what I'd do is put it into one fraction. So *a* plus *a* is just 2*a*, I think. That's right. Then I'd go to five plus ten.

Although this was a common response, more students simply added the denominators to give the answer as "*a*/15". The reasons for this are not clear, except that the students mentally detached the letter from the expression and replaced it once they added the denominator of the fractions. This still does not explain why they did not apparently add the numerator to get "2*a*". However, some students seem to confound adding 1 + 1 with multiplying 1 × 1.

Another student used a similar strategy, of treating the numerator and denominator as unconnected numbers, when responding to Item 15 [3*p*/4 – *p*/8], at the same time, detaching the *p* from the numerals.

> .... You just go three over four because *p* minus *p* is nothing, so it would probably be three over minus four for the answer.

---

[5] '...' indicate long pauses in the interview responses.

This response also seems to depend on the same generalised logic applied to expressions such as Item 2 [$5p - p + 1$] discussed earlier. The student gave the answer to Item 2 as "$4p + 1$". Similar responses to Item 15 [$3p/4 - p/8$; response: $2p/4$], Item 25 [Take $n$ away from $3n + 1$; response: 4], and Item 28 [$4y = 20$; response: 16] indicate a fragile, confused memory of the conventions of algebraic notation. Students with these understandings would sometimes "get the answer right, just by chance".

Other students were also confused about the correct notation; so "$3a/10$" became "$3_{a/10}$" (three and one tenth $a$), whilst others seemed to generalise, incorrectly, the notion of *like terms*.

> S (Student): 'Cos I'm not that well at fractions and those sort of things. The addition and subtraction ones aren't too bad, but Question 3 [Item 3] is just $2a$ plus 15.
>
> I (Interviewer): Why would you do that?
>
> S: Because, um, five and ten are like terms, 'cos you know what they are and the two *a*'s they're… [the student's explanation ceased].

Dealing with quarters and eighths appears to have been easier than dealing with fifths and tenths, although both Item 3 and Item 15 only required students to find common denominators by doubling strategies. However, the procedure applied to the contexts of Item 3 and Item 15 was unable to be generalised to a situation that required students to use multiples other than doubles.

> S: the first one I get, um, $3a$ over 10, the next one I get $5p$ over 8
>
> […]
>
> S: Because that is half of that, I need to double [3] that to make it simpler, but if you double the bottom one you have to double the top one as well, so I double the five to make it ten and I double the $a$ to make it $2a$, which means $2a$ plus $a$ over ten, which equals $3a$ over ten… [student continues with an explanation of Item 15]
>
> I: So what happens if I give you something like [writes] $t$ over seven plus $t$ over three?
>
> S: I don't know. I'm not quite sure, um. I get a bit confused. I'm not really sure. I can't really remember how to do it. I think I've done it before, but I can't remember.

Others could add fractions, such as in Item 3, albeit incorrectly, but not subtract fractions as was required in Item 15.

> … I haven't done minus fractions. I just think you just do $3p$ minus $p$ and then four minus eight.

The more able students, those in the High-ability group, provided explanations that were procedurally sound, but lacking in any articulation of conceptual understanding. When asked "why?" most responded with "dunno", "to make it easier", "I don't know why. That's just what I've been taught to do", or:

> because both of the bottom numbers have to be the same, to be able to be a proper fraction, adding a fraction, and minusing. I don't know why.

### *Brackets*

Items in the brackets category (Table 4.4) were all of a degree of difficulty approximately equal to, or greater than, the mean ability estimates for the group of participants (category mean: 1.043 logits; average ability estimate: -0.64 logits). Accordingly only those students of average ability or better could be expected to complete these items successfully. Part of the difficulty with items such as those in this category lies in the tendency for students to conjoin terms inappropriately.

Incorrect responses to Item 11 [$8p - 2(p + 5)$], Item 18 [$2(x + 4) + 3(x - 1)$] and Item 19 [$2(x + 5) + 8$] indicated that students' understanding of the role of brackets was fragile. For Item 19 (difficulty estimate: -0.27 logits), the most common incorrect response given was "$15x$" (17 out of 86 incorrect responses), followed by "$2x + 13$" (9 out of 86). Item 18 (difficulty estimate: 0.13 logits) yielded 60 different incorrect responses out of a total of 103 incorrect answers. The most common responses given (7 out of 103) were: "$8x$", "$2x + 8$", or "$5x + 11$". There were 158 incorrect responses to Item 11 (difficulty estimate: 2.28 logits). Of those, 40 students gave the answer as "$6p + 10$". The next-most-frequent incorrect response given was "$6p + 5$" (17 out of 158). A similar pattern of incorrect responses appeared for Item 7 [$(a - b) + b$] (Table 5.4). These responses suggest that brackets indicate to students that one should "do it first", and that students' tend to conjoin terms in order to "do" something, and also to close on the operation in the brackets.

The absence of a numeral outside the brackets was also a concern for students who were more familiar with expressions having positive integers greater than "1" outside brackets. These numbers indicated to the students that they had to "do the brackets first", or "expand" the brackets. However, they could not make similar meaning out of expressions such as that in Item 7 [$(a - b) + b$]. It would also appear that the complete absence of numerals made this item more difficult.

> It looks different because there's not numbers anymore…and… I don't know.

Another student, also from the Average-ability group, admitted being confused, although at the same time there was recognition of the structural similarities with other items (18 and 19), as well as the implied *one* in front of the brackets.

> Oh yeah, I guess it does sort of look like that [Item 19], because if there's nothing there [indicating the front of brackets], it's one. So, it'd probably end up being… Yeah, it sort of looks similar to the… [indicates Item19], but I still wouldn't know what to do.

Other students, such as one from the upper end of the Average-ability group, recognised the implied *one* outside the brackets and made an attempt to arrive at a solution in general terms:

> You have to pretend there is an invisible one and a times in front of the bracket. So you go one times $a$ equals $a$ minus one times $b$, so it basically does not really matter. So you make it $2b$. Get rid of the brackets after you have done that, plus $2b$. $a$ minus $2b$, I think it would equal because $a$ can't be changed because there is nothing similar.

### *Equations*

Although not identified as a separate category, it is appropriate to discuss the errors made by students as typifying approaches to equations in general, regardless of their mathematical form. Most students relied on arithmetic strategies – either recalling number facts, or using trial-and-error substitutions. Where items required students to use algebraic procedures (e.g., Item 39 [If $ax = 5$, $x =$…?], or equations with the unknown on both sides) the number of non-attempts increased.

For simpler items, such as Item 28 [If $4y = 20$, then $y =$…?, difficulty estimate: -2.33 logits], there were 39 non-attempts but only 19 incorrect responses. Of these, nine students gave the solution as "16", indicating that they understood "$4y$" to mean, incorrectly, "$4 + y$". In contrast to this, responses to Item 32 [$10y = 5$] consisted of 65 non-attempts and 59 incorrect responses. Of those incorrect responses, "$y = 2$" (23 out of 59 responses) and "$y = 5$" (20 out of 59 responses) were the most common.

Incorrect responses to Item 30 [$x/4 = 12$, difficulty estimate: -0.93 logits] also suggested that students *reacted* to the numbers "4" and "12", rather than read the equation with meaning. The most common incorrect response was "$x = 3$" given by 32 students out of the 45 who answered incorrectly.

On the other hand, incorrect responses to Item 35 [$x + x/3 = 4$, difficulty estimate: 2.17 logits] indicated that students tended to multiply only the right hand side of the equation by three, rather than find a common denominator first. Interview responses suggest that students have learnt to "do the opposite to the other side" as a procedure for solving equations. The procedure resulted in the most common incorrect response of "$x = 6$", given by 42 students of the 96 who answered incorrectly. The incorrect response "$x = 1$" was given by 15 students.

Many students made no attempt to provide answers to Items 34, 36 and 37. Item 34 [$x + (x + 2) = (x – 1) + 8$] with 120 non-attempts produced 25 different answers out of a total of 57 incorrect responses. Item 36 [If $63/x = 180$, what is $x$?], with 137 non-attempts and 62 incorrect responses, yielded 34 different errors, none of which was repeated more than seven times. Out of a total of 51 incorrect responses to Item 37 [Solve $5a – 4 = 2a + 8$], there were 31 different solutions. The most common incorrect response was "$3a + 12$" which occurred four times. All other responses occurred with a lesser frequency.

To solve equations by trial-and-error strategies requires students to have a confident number sense, and be able to hold in working memory several bits of data that can only be connected in a final step. This cognitive load challenged many of the students in the Low- and Average-ability groups.

## Summary

The detailed examination of errors considered the numbers of attempts and non-attempts made by students to each of the 40 survey items. Particular incorrect responses to each item were identified together with the frequency of their occurrence. The absolute numbers of non-attempts, incorrect responses and correct responses indicated that the items with the greatest Rasch difficulty estimates were not those that most students avoided. The numbers of non-attempts for certain items, such as expressions with fractions or equations with the unknown on both sides of the equation, suggest that students may be unfamiliar with such items, or avoid them because of their appearance.

Analysis of incorrect responses revealed some common misconceptions, or limiting concepts, in responses to particular categories of items. Frequently occurring responses could be considered as typical, such as the responses to items requiring addition, or subtraction, of algebraic fractions. Other items attracted many different responses, some of which were quite idiosyncratic, thus suggesting the existence of widespread, differing, and confused mathematical ideas.

The errors suggest that students' structural understanding of algebraic expressions is not robust. Instead, understanding, or the ability to use a procedure, seemed to rely on the students' recognising an expression presented in a particular, and familiar, format. Explanations given by students during interviews were often simplistic procedures, based on a limited range of examples that they then tried to generalise to unfamiliar contexts. However, without understanding the mathematical meaning embodied in the structure of an expression, these students were unable to do so successfully.

## ANALYSIS OF ITEM CLUSTERS

Error analyses indicated that for some items, and for some categories of items, there were common misconceptions associated with the item difficulties, or student ability, as described by the Rasch model. In this section, these results are interpreted and discussed within a framework of the item difficulty clusters (i.e., Items of least difficulty, Cluster 1; average difficulty, Cluster 2; and, greatest difficulty, Cluster 3. Table 4.3, Figure 4.5). Similarities and differences of possible cognitive pathways within each cluster are described. Individual items and student comments from interviews are used as supporting illustrations.

The explanations of students about their thinking as they carried out the various algebraic processes provide insight into how their peers might also think. There is no claim made about the typicality of these thinking processes, although some are, it would seem, common to the groups of participants involved in this study. The comments of the students are drawn from students in each ability group (Table 4.10, Figure 4.8).

The discussion in this section begins with a description of possible cognitive steps needed to respond successfully to items in each cluster. The cognitive steps are those that appear to be the least mathematically sophisticated necessary to obtain a correct answer to the item. They are often those described by students in the lower ability groups who successfully responded to the items; and, might not be the most efficient, mathematically generalisable, nor algebraically elegant. Hence, in cognitive terms, such procedures require least mathematical conceptual development. From the interview data, these strategies appear to allow students to operate in a way that is most familiar to the individual concerned, although it may be argued that such strategies demand a greater cognitive load than do more elegant approaches.

Rasch modelling of the survey responses from 222 students resulted in three clusters of items with significantly different means (Table 4.3, Table 4.5). These are numbered from the cluster of least difficult items, Cluster 1 to the cluster of the most difficult items, Cluster 3. This cluster includes smaller, less significantly different Clusters 3a, 3b and 3c.

Items in Clusters 1 and 2, and particularly those in Cluster 1, are often simple examples of other more difficult items in later clusters. Responses by students to the more difficult items are also discussed alongside the easier items where appropriate.

### Cluster 1

Cluster 1 consisted of items of the least difficulty. The difficulty estimates for items in this cluster ranged from -3.27 to -1.98 logits, with a mean difficulty estimate of -2.70 logits.

The most unsophisticated, or superficial, strategies could be used by students to answer these items. Many students relied on visual cues or a visual memory of these items, and then used strategies that were predominantly arithmetical, rather than algebraic. The items in this cluster may all be read meaningfully from left to right. The arithmetic symbols of plus, minus, times and the equal sign need to be recognised, but understood only as instructions to operate on the terms of the expression or equation. The operations can be carried out sequentially as the expression is read. Only one or two operations are needed for a solution. Suitable number facts need to be recalled.

The letters may be ignored, as suggested by Küchemann is the case with Item 24 [If $a + b = 3, a + b + 2 =...?$], or the focus can remain on the numbers with the letters "tacked on" if necessary. As one student described the procedure used to answer Item 4 [$4 \times 5b$], successfully:

> Just four times five *b* would be twenty *b*. Just times it by all that, times it by the number and then just add the letter on…kind of...

And another:

> I think of timesing the numbers first, then doing something with the letters.

Item 27 [$x + 5 = 7$] and Item 28 [$4y = 20$] required students to recall simple number facts and had a structure that would be familiar to students from primary school, (i.e., $\quad + 5 = 7$ or $4 \times \quad = 20$). The equal sign could be interpreted as "an answer is required". This particular understanding of the equal sign is prevalent, and often the only interpretation that students had.

One student (Low-ability group) explained that she counted on from five to reach seven in order to answer Item 27. Students in the Average-ability group explained their thinking in algebraic terms, but only when they appeared to have already worked out the answer arithmetically – which is possible in these simple cases.

Correct answers to Item 1 [$3m + 8 + 2m –5$] and Item 2 [$5p - p + 1$] required some understanding of the role of letters (or some other symbol) in expressions or equations although no meaning need be attached to the letters. Only surface similarities of the letters in the expressions need be recognised, and the letters operated on according to remembered rules. The answers need not be closed although the most common errors made with these two items were the conjoining of the terms in Item 1 to give an answer of "$8m$", or the confusion of signs. This latter error was the most common cause of incorrect responses to Item 6 [$5a – 2b + 3a + 3b$].

Students at all levels stated that like terms needed to be collected in order that expressions, such as those in Items 1, 2, 5 and 6, be simplified. However, there were some confusions: about the "likeness" of terms such as *ab* and *b*; whether or not the rule about adding and subtracting like terms might also apply to multiplication; and, the effect of the "absence" of a number in front of a term, such as $ab$ in Item 5 [$2ab + 3b + ab$].

## Cluster 2

Cluster 2 consisted of items of average difficulty, grouped around the mean of 0.0 logits. The difficulty estimates for items in this cluster ranged from -1.17 to 0.46 logits, with a mean difficulty estimate of -0.302 logits. The difference in the means between this cluster and Cluster 1 was significant ($t$=20.29, $p$<0.05). This cluster contained the greatest number of items, with representatives as both expressions and equations, from each of the categories of item types.

The items in this cluster may still be read meaningfully from left to right, but the information could be partly disconnected (e.g., Item 6 [$5a – 2b + 3a + 3b$]) or implicit in the structure of the expression/equation (e.g., Item 17 [$(6xy)^2$]). More than one operational step might be needed to give a correct response, but each step can be closed or finished before the next step, and hence provide some concrete evidence of progress to a reasonable solution (e.g., Item 19 [$2(x + 5) + 8$]; Item 29 [$2t – 23 = 49$]).

Item 29 [What is $t$ if $2t – 23 = 49$?] may be solved step-wise: $2t = 49 + 23 = 72$; $72 \div 2 = 36$; therefore $t = 36$. Item 23 [What can you say about $m$ if $m = 3n + 1$ and $n = 4$?], Item 38 [If $r – 82 = 7$, then $r – 83 =$…?], Item 33 [$(x+3)/2 = 7$], Item 32 [If $10y = 5$, then $y =$ …?], and Item 31 [$4(p + 3) = 32$] may also be solved by this multi-step procedure.

Equations in this cluster contained only one pronumeral, on one side of the equation. The pronumeral represented a single unknown, that could be found by trial-and-error number substitution (one student called this *guess-and-check*) because the answers were small, positive integers. If trial-and-error strategies were relied upon, items such as Item 35 [$x + x/3 = 4$] would become so much more difficult (see below, Cluster 3, for further discussion).

Substitution of numbers in expressions was, by the time these students undertook the survey, a familiar exercise, and might explain why Item 23 [What can you say about $m$ if m = $3n + 1$ and $n = 4$?] was the easiest item in this cluster. Such a procedure does, however, require an understanding of the role of the letters as representing numbers, and the ability to infer what is to be done from the question. The syntax of the question does provide a prompt, in that the designation of the value of $n$ immediately follows the description of $m$. The case of a verbal or visual prompt also occurs with Item 40 [If $e + f = 8$ then $e + f - g =$…?].

Equations, Item 32 [If $10y = 5$, $y =$…?] and Item 31 [Solve $4(p + 3) = 32$], were more difficult because they required a more sophisticated understanding of arithmetic relationships (i.e., In Item 32, multiplication does not always lead to a larger number). Students with the perception that multiplication had to result in a larger number than those being multiplied found this item difficult.

> S*:* It would be ten divided by two.
>
> I: Why do you say that?
>
> S: Because ten divided by two is five.
>
> I: … What is the question actually telling you?
>
> S: Saying that 10*y* is five.
>
> I: So, what does 10*y* mean?
>
> S: Ten times something…
>
> I: You are telling me it is ten divided by two
>
> S: But it can't equal anything. It's ten times something doesn't equal five. It has to be divided by to get a smaller number.

One student appeared to confuse fractions with negative numbers

> …10*y* equals five. So *y* would equal minus two. So ten times minus two is five, isn't it? Yeah, I think so…

On the script of the algebra survey the student gave the answer as "-0.5". Other students gave "$y = -5$" as the answer (8 out of 59 responses).

One other student, who correctly solved Item 28 [$4y = 20$] by recalling number facts from the four-times table, rather than use the inverse operation, commented on Item 32 which was included in the same set and immediately followed Item 28 in the interview:

> Well, I'd times it. So, I'd go, umm, what? Four what's equal 20? Four fives equals twenty. And then, I don't know what I'd do here because that [$10y = 5$] doesn't look right.

Two possible strategies used to solve Item 31 [$4(p + 3) = 32$] were: to transform the left hand side of the expression into $4p + 12 = 32$ (a form like that of Item 29), and then solve the equation by a series of closed steps; or, alternatively, to understand that ($p + 3$) could be held as an entity until 32 was divided by four, to give the answer of eight, and then $p$ found by subtracting three from eight. No student interviewed suggested this latter procedure.

> S: …Like you do the four times $p$, um, and then you do four times three. And you take away what the four times three is from the 32, and you see what that number… what number equals it when it is timesed by four.
>
> I: How would you find that out?
>
> S: Go through my four times table and see which one equals that.

However, when this student was asked, immediately following this conversation, to solve Item 28 [If $4y = 20$, then $y$ =?], the reply was, even after some prompting: “I can’t do it. I don’t think it is the same”. The student recognised no structural similarity between the two items. Her success on Item 31 was due to her remembering the procedure for solving similar equations that had been taught in recent lessons.

In the case of examples such as Item 8 [$4ab/4b$], many students responded to the surface features, demonstrating no real mathematical understanding, but simply “scratched them out” if the numbers (4) and/or the letters ($b$) “looked the same”. This type of response may explain why Item 16 [$2/a^2 \times 5a/4$] was scaled at a greater level of difficulty (3.17 logits). The responses to this item are discussed in greater detail in the following sub-section.

Items with a difficulty estimates between -0.34 and 0.00 logits have elements that illustrate a transition in thinking from that required to answer items with a difficulty level in the lower part of this cluster (difficulty estimates: -1.17 to -0.52 logits) to that needed to answer items in the upper part (difficulty estimates: 0.07 to 0.46 logits). These items are: Item 26 [If $p + q = 5$, then $p + q + r$ =…?], Item 18 [$2(x+4) + 3(x-1)$], Item 25 [Take $n$ away from $3n + 1$], Item 40 [If $e + f = 8$ then $e + f - g$ =…?], Item 6 [$5a - 2b + 3a + 3b$], Item 17 [$(6xy)^2$], and Item 20 [Multiply $x + 5$ by 4], (in order of increasing level of difficulty).

The items may still be read meaningfully from left to right, and answers obtained by a sequence of steps. The syntax of the expressions needs to be attended to, and simple inferences drawn. The answers to each item tend to be general (algebraic) and unclosed, although, once again, no meaning need be attached to the letters.

Students could not obtain a unique answer by conjoining the terms in an expression such as $p + q + r = 5 + r$ to obtain the closed "$5r$" (Item 26). This was the most common error. Out of 91 incorrect responses to this item, 34 gave the answer as "$5r$", which was 37% of incorrect responses and 15% of all responses (Table 5.4). This means that many students can execute the procedures required but fail to use the correct algebraic notation, possibly because they attach no appropriate meaning to the result. Many students in the Average-ability group who could correctly simplify expressions such as Item 1 [$3m + 8 + 2m - 5$] gave the response of "$5r$" to this item.

The most difficult items in Cluster 2, Item 17 [$(6xy)^2$] and Item 20 [Multiply $x + 5$ by 4], required careful reading, attention to syntax, and additional inferences to be made from "ambiguous" statements before a correct response could be given. Once this inference is made, the algebraic procedures for simplification of the expressions consisted of two or three steps.

However, relationships between each of the mathematical steps have to be held in the mind of the student. In the case of Item 17, many responses indicated that students could square one, or some of the terms, but not necessarily all. Such responses were "$36xy$", "$6xy^2$", "$6x^2y^2$", or even "$12x^2y^2$". Students' explanations of their understanding of what it means to square a number, or the role of brackets indicated the source of some of the difficulties that they experienced with this item and Item 20.

Items in Cluster 2 represented a spectrum of difficulty, and a range of algebraic concepts. The items represent a developmental progression of algebraic thinking – from the application of arithmetic to algebraic statements (evaluating expressions by substituting numerical values) to being able to read, interpret, and manipulate general algebraic expressions.

## Cluster 3

Cluster 3 consisted of items of greatest difficulty. The difficulty estimates for items in this cluster ranged from 1.18 to 3.56 logits with a cluster mean difficulty of 2.095 logits. Three smaller, but significantly different sub-clusters (3a, 3b, 3c) identified within this large cluster are discussed individually in the following.

### *Sub-cluster 3a*

The difficulty estimates for items in Sub-cluster 3a ranged from 1.18 logits to 1.59 logits with a mean item difficulty of 1.377 logits. This was significantly different to the mean of Cluster 2 ($t$=18.76, $p$<0.05). In items in this sub-cluster, arithmetic signs act as relational signals, rather than as operators indicating that some arithmetic action be taken. Of particular importance are: the vinculum, signifying division relationships (Item 15 [$3p/4 – p/8$], Item 3 [$a/5 + a/10$], and Item 14 [$x/3 \div y/4$]); the equals sign, signifying expressions of equivalent value (Item 37 [$5a - 4 = 2a + 8$] and Item 34 [$x + (x + 2) = (x - 1) + 8)$]); and, the use of brackets (Item 7 [$(a – b) + b$] and Item 34). The vinculum and the brackets provided visual cues to students to treat the terms as *entities* (mathematical objects that could be manipulated as a whole). At the same time, elements in these terms also needed to be treated individually for students to arrive at a complete solution.

The equations may be solved using trial-and-error substitutions although students needed to recognise that they were dealing with related and equivalent expressions on both sides of the equals sign.

> All right, when I see Question 37 [$5a – 4 = 2a + 8$]…well, I see it as four minus. Well, four minus $5a$ will be the equivalent of two plus, $2a$ plus eight…

This particular student went on to explain his thinking process in more detail when asked to solve Item 34 [$x + (x + 2) = (x – 1) + 8$].

> Well, in this one I simplify both sides first. $2x$ plus two equals $x$ minus seven…so if $2x$ plus two equals $x$ minus seven…doesn't make any sense […] Oh, wait a minute, it's plus seven. It can work […] Yeah. I did it by guess-and-check I think. I don't know how to work them out properly.

In items that were equations, students recognised that the letter symbols stood for an unknown number, even if they did not always do so when dealing with expressions. One student (of Average-ability) thought that letters in expressions were:

> …Representing something, representing something that is not a number. Like representing a thing or an amount of something. Like, they could be representing anything from a book to a … to anything, to representing people. *a* and *b*… they all represent something, that's not a number.

But in equations:

> ...Oh, these are the complicated ones. *x* plus five equals seven, then *x*...
>
> In this one, *x* represents a numeral instead of something other than a numeral. This actually represents a numeral now…

The relatively high level of difficulty of the items involving fractions (Items 3, 15 and 14) may be attributed to the difficulties and misconceptions students have with operating on arithmetic fractions. In these items (Items 3, 15 and 14), the letters may still be ignored and arithmetic procedures suffice to manipulate the numbers. In many instances, noted in the errors, students deleted the letter altogether or used other, more idiosyncratic notations. These have been described in the previous section of error analysis.

The expressions with brackets presented many difficulties to students. These included a tendency to conjoin terms inside the brackets. Students interpreted the arithmetic signs of plus or minus as instructions to "do" something, particularly as most students expressed the limiting interpretation of brackets as indicating that some arithmetic procedure be "done first". The absence of any numeral outside the brackets (as in Item 7) also caused confusion; many students seemed to have met only examples with numerals outside the brackets. The brackets could be ignored in Item 7 and the correct answer still be obtained.

Others confounded ideas of one and zero.

> It looks different because there's not numbers anymore…and… I don't know.

Some students resorted to using numbers, although not successfully:

> It's like *a* minus *b* plus *b*. You can pretend they're numbers and work it out like that.

However the brackets could be ignored and the correct answer still obtained:

> I: What are the brackets doing there?
>
> S: Not much… I don't know.
>
> I: Why can you just ignore them?
>
> S: Because there's nothing going with the brackets. Like there's no two in front of it, or another number after it.

Lee and Wheeler (1987) commented that students often resort to numerical examples. When the results conflicted with their algebra, students preferred to accept the numerical validation, but rarely made attempts to reconcile these results with the algebraic result.

### *Sub-cluster 3b*

The difficulty estimates for items in Sub-cluster 3b ranged from 2.17 logits to 2.64 logits with a mean item difficulty of 2.363 logits. The three items in this cluster required students to: read the item as a whole, as the meaning is not immediately clear from a left to right reading; understand algebraic syntax; consider the mathematical relationships between each term; and, hold related arithmetic ideas in their working memory.

Item 35 [$x + x/3 = 4$, what is $x$ equal to?] and Item 36 [If $63/x = 180$, what is $x$ equal to?] required the use of arithmetic relationships and an understanding of number. These items may be dealt with arithmetically, but, in particular, Item 36 required the perception that $x$ has to be less than one, or found as the result of inverting the arithmetic statements. If trial-and-error methods were used in Item 35, the students needed to test each number as an integer as well as its relationship to three in order to give the answer of four.

The errors suggested that students who used a trial-and-error procedure found difficulty in holding open many of the computational steps. These students gave the answer as "$x = 1$" in many cases. On the other hand, students who attempted to use the algebraic procedure of inverse operations only inverted the division of $x/3$ rather than first finding the common denominator for the left hand side. These students gave the answer as "$x = 6$" (Table 5.4).

Although the solutions to this item and Item 36 are closed, with unique number answers, they are not easily checked. Some students managed to obtain the correct solution of “0.35” for Item 36. This could only be done through their understanding of, and correct use of, arithmetic inverse computations. Other students could neither carry out the arithmetic, nor understand the inverse procedure, but did demonstrate a sense of number, although this did not lead to a correct response. The following exchange is illustrative of this thinking. The student (Average-ability) was trying to explain how he could arrive at a solution for Item 36 [$63/x = 180$], which he failed to answer in the survey.

> …It’s just 180 times $x$. First to find it out you’ve got to go 63 divided by $x$ equals 180. So 180 times $x$, I mean 180 times $x$ equals 63, I think … So I’m doing 180 times $x$ equals 63. So it’s, it’s around like a third, around a third. So 63 divided by a third. Well, it has to be lower than zero, because it will go into that more than what that number is to get 180. So it has to be around a third, I think.

It is also worth noting the student’s conception of a fraction being less than zero, rather than a number between zero and one.

Item 11 [$8p – 2(p + 5)$] required students to distribute the negative two to the five (and $p$) in the brackets. It is this feature of the item that makes it more difficult than similar items, Item 18 and Item 19. Most students successfully found -2$p$ because of the visual cue and proximity of the -2 and the $p$, but then found +10 by multiplying the two (not -2) with the +5 (see errors, Table 5.4). Once again, there is the need not to close on each step, but to hold, incomplete in working memory, some of the arithmetic relationships, before deciding on the final answer.

### *Sub-cluster 3c*

The difficulty estimates for items in Sub-cluster 3c ranged from 3.09 logits to 3.56 logits with a mean item difficulty of 3.273 logits. Items in this cluster required meaning to be attached to the letters and other symbols. Also students needed an understanding of arithmetic principles such as how numbers, or their representatives, are connected and how, in general, numbers behave. Meaning of the item cannot be obtained through a simple left-to-right reading, nor are there any surface cues. The answers to each of these items are general, and not easily verifiable by arithmetic means.

Item 39 [If $ax = 5$ then $x =$…?] required students to move from simple arithmetic, trial-and-error strategies that resulted in unique solutions, to understanding that the answer was a general result based on knowledge of the behaviour of numbers. Thus students needed to apply their understanding of mathematical inverses. An intermediate step to understanding that $x$ could equal any number represented by the expression $5/a$ appeared to be an understanding that solutions to equations could be other than integers. Many students gave the answer as "$x = 5$" (24 students) implying that "$a = 1$". Some few (13 students) gave the answer as "$x = 1$". Interviews suggested that students opted for these integers as logical alternatives, but may have chosen "$a = 1$", reasoning that as $a$ preceded $x$ it represented the number 1.

> I don't know what I would do for that...There are so many different things that you could write. Well, not so many, but there's two different ways you could write it. You could…most likely to be one times five or five times one. But you don't know which one $x$ is going to be.

Other students justified this thinking as "the easy thing to do", or that:

> the only factors of five are one or five, so one number would have to be one and the other would have to be five, if it's $a$ times $x$.

When prompted, this student acknowledged that the answer did not have to be a whole number, but said:

> Like, if it would be a whole number then it's better just to use a whole number…because one and five are factors of five, so it just makes sense.

Students in the High-ability group were more likely to get close to the correct response. One student suggested five divided by $x$ equalled $a$, but could not, in the interview, restate this as "$x$ equal to five divided by $a$". Perhaps this indicated that the student was fixed on finding out a value for $a$ so that a value for $x$ could eventually be found.

Another student demonstrated the necessary cognitive shift from specific numbers to general statement:

> …Um, it could be many different, like…because there's more than…Oh, it would have to be one and five. Or it could be decimals, but depending on which one, it could be one times five or five times one…[…]. $x$ equals … five divided by $a$!

Item 16 [$2/a^2 \times 5a/4$] provided none of the surface clues that made Item 8 [$4ab/4b$] easier (by 3.81 logits). Students needed to understand the significance of the vinculum as denoting division as well as acting as a grouping symbol. The procedure of physically "scratching out" symbols that "look the same" could not work in this context, without the expression being transformed (factors found). Most frequently, students simply multiplied the terms together without simplifying the expression at all (52 out of 129 responses). Many of the other incorrect responses were various incomplete attempts to simplify the expression, such as "$10/4a$" (10 out of 129 responses, Table 5.4).

Students also failed to see the significance of the "×" sign, except to multiply the terms and so "put the terms together". They did not read the sign as uniting all the terms in the expression without there being a need to carry out any arithmetic operation. Alternatively, students who were unfamiliar with this particular structure of an expression had no conceptual understanding of the procedures they were used to follow in other examples similar to Item 8 [$4ab/4b$].

From the responses to Item 16 and Item 8 it was clear that students acted according to simple rules and surface features of the expressions. It appeared that many students were unfamiliar with questions that required them to carry out more than one step in order to simplify an expression; or, having closed after one step of the procedure, they saw no need to look for possibilities for further modifications. Rigid adherence to a routine in which terms of the expression were multiplied so that everything was "put together" before simplification, masked some of the clues that could have helped in simplification. That groups of students did this more frequently is discussed in the following section where the responses of different ability groups are discussed.

Item 10 [$(x + y)^2$], like Item 16, provided no cues as to how it might be rewritten. To transform the expression required meaning to be given to the expression as well as the application of a multi-step procedure that was abstract, general and unverifiable. The most common response was "$x^2 + y^2$" (Table 5.4). Students *knew* that squaring meant "to multiply a number by itself" as stated by many in response to a specific question in the interview. This response often appeared to be more a well-rehearsed definition than a mathematical statement to which students attributed deeper meaning.

Although Item 10 and Item 16 were the most difficult, many students attempted them (Table 5.1), and students interviewed seemed very confident in handling them. However, their explanations demonstrated that they relied on simplistic interpretations and overgeneralised rules. Item 39 [If $ax = 5$ then $x =$…?] presented the students with a general context which they seemed unable to associate with items of very similar structure, although this item was presented with items such as Item 28 [$4y = 20$] in the interview.

## Summary

Clusters of items, resulting from Rasch-modelling of the survey responses, included many examples that were algebraically different, as indicated by the different categories identified. The model suggests that the items in each cluster are psychometrically similar. The conceptual development needed to respond successfully to items in each cluster is similar, although the algebraic bases may be different. These have been described in this section, drawing on data from error analysis and interviews with students.

The means of each cluster were significantly different, and suggest that important conceptual shifts in the thinking of students were needed for successful responses to items as clusters of items became more difficult. These changes involved students in shifting their thinking from simple arithmetic responses (as for items in Cluster 1), through an understanding of arithmetic relationships (items in Cluster 2), to being able to understand general relationships between numbers and operations on numbers (Cluster 3).

The easier items, such as those in Cluster 1 and many in Cluster 2, could be answered successfully by students recognising a familiar example and using a rehearsed procedure. The items in these clusters did not require syntactical awareness of algebraic notation on the part of students. In some cases, the items could be answered using procedures that were, in fact, not at all algebraic, or even mathematical. Items that did not fit an expected representation were less likely to be answered successfully. This was most noticeable when students had to respond to items where words and symbols expressed algebraic ideas that students were more used to seeing in conventional symbolic representations.

Items in Cluster 2 and Sub-cluster 3a represented transitional stages in conceptual development, where: visual cues needed to be supplemented by more abstract, mathematical (either arithmetic or algebraic) understanding; or, arithmetic strategies needed to be developed to a level where students possessed sound number sense, and efficient number fact recall. The most difficult items demanded that students be able to interpret algebraic statements by making inferences based on their knowledge of notation and syntax, as well as their ability to conceptualise the mathematics exemplified in simpler examples.

The responses from students on the survey, and in the interviews, suggest that for many their experience of algebra was limited to the manipulation of simple expressions presented symbolically; or the solution of equations that needed only the recall of number facts, or simple trial-and-error strategies. The items in the most difficult cluster were elaborations, or less familiar representations, of conceptually similar items in the less difficult clusters. This suggests that most students had a limited experience of the variety of forms that the same mathematical idea can take. Also, many procedures used to manipulate expressions or solve equations tended to be unsophisticated, or simplistic, and of limited generalisation. The most difficult items required students to operate beyond familiar experiences and apply general principles based on conceptual, not procedural knowledge.

## ITEMS THAT DID NOT FIT THE RASCH MODEL

Analyses of the algebra survey results revealed clusters of items according to their difficulty estimates together with groups of students according to ability estimates. It is useful at this stage to consider the three items, Item 5 [$2ab + 3b +ab$], Item 12 [$2/a \times 3/b$], and Item 30 [If $x/4 =12$, what is $x$?] which did not fit the model, in terms of the overall response patterns of students in each ability group (Figure 1.1) and the common errors made (Table 5.4).

According to the model, *misfit* of items occurs because responses to items are outside the expected range. In the instances of *underfit* (Item 5 and Item 12), students responded correctly to items that they could not be expected to answer correctly, because the items were scaled with a higher difficulty estimate than the ability estimate of the students. *Overfit* items (Item 30), although measuring the same construct as other items in the survey, discriminated more precisely than expected.

Response patterns to those underfit items, Item 12 [$2/a \times 3/b$], and Item 5 [$2ab + 3b + ab$] are first discussed. A third sub-section addresses response patterns to Item 30 [If $x/4 = 12$, what is $x$?].

### Response Patterns to Item 12

Responses to Item 12 have to be considered with respect to responses to Item 3 [$a/5 + a/10$], Item 15 [$3p/4 - p/8$] and Item 14 [$x/3 \div y/4$]. Study of scripts from the survey shows that many students correctly answered Item 12, but not Item 3 or Item 15, both of which fitted the model. Item 14, although also requiring the simplification of an algebraic expression with fractions (this time a division of fraction), was so much more difficult, and also fitted the model. Errors for Item 14 provided no clear pattern that could indicate any particular common misconceptions, unlike the patterns of erroneous responses to Item 3 and Item 15 (Table 5.4). The question then arises about the nature of Item 12, Item 3 and Item 15 and the possible reasons for the misfit of Item 12.

The pattern of responses from the three different ability groups suggests reasons. For the participants in each group, the numbers of correct responses to the items were considered: students who correctly answered Item 3 only, Item 15 only, Item 12 only; students who correctly answered both Item 3 and Item 15; and, students who correctly answered all three items. The results are set out in Table 5.5 and illustrated in Figure 5.1.

Proportionally fewer students in the Low-ability group correctly answered any of the Items 3, 15 or 12. However, a far greater proportion of all correct responses from this group to those three items was to Item 12, compared to that from the other two ability groups.

One student of Low-ability correctly answered Item 3, but not Item 15. One other student correctly answered Item 15, but not Item 3. Twenty-eight students in the Low-ability group however, did answer Item 12 correctly.

*Table 5.5: Numbers of students by ability group correctly answering Items 3, 5 and/or 12*

| Ability Group | | Item 3 only correct | Item 15 only correct | Item 3 and 15 correct | Neither item 3 nor item 15 |
|---|---|---|---|---|---|
| 1 Low-ability | Item 12 correct | — | — | — | 28 |
| | Item 12 Incorrect | 1 | 1 | — | — |
| 2 Average-ability | Item 12 correct | 7 | 7 | 3 | 25 |
| | Item 12 Incorrect | 4 | 4 | 3 | — |
| 3 High-ability | Item 12 correct | 3 | 3 | 23 | 15 |
| | Item 12 Incorrect | 2 | 2 | 5 | 0 |

A higher proportion of students in the Average-ability group correctly answered Item 12 only. Some students in this group also correctly answered both Items 3 and 15, and three students answered all three items correctly.

Of those in the High-ability group, a greater number of students than in the other ability groups correctly answered all three items. On the other hand, proportionately fewer students in the High-ability group than was the case in the other two groups correctly answered Item 12 only,

Although students in the Low-ability group did not answer Item 3 or Item 5 correctly, more than half of that group answered Item 12 correctly. More students in the Average-ability group were able to answer Item 3 and Item 5, but a greater proportion was able to answer Item 12 correctly. Figure 5.1 below, illustrates these data.

Students in the Low-ability group explained that they simplified Item 3 and Item 15 by adding the denominators, and in most cases, the numerators as well. Their generalised rule was “what you do to the top, you do to the bottom”. By applying this rule to the visually similar expression in Item 12, they obtained the correct answer.

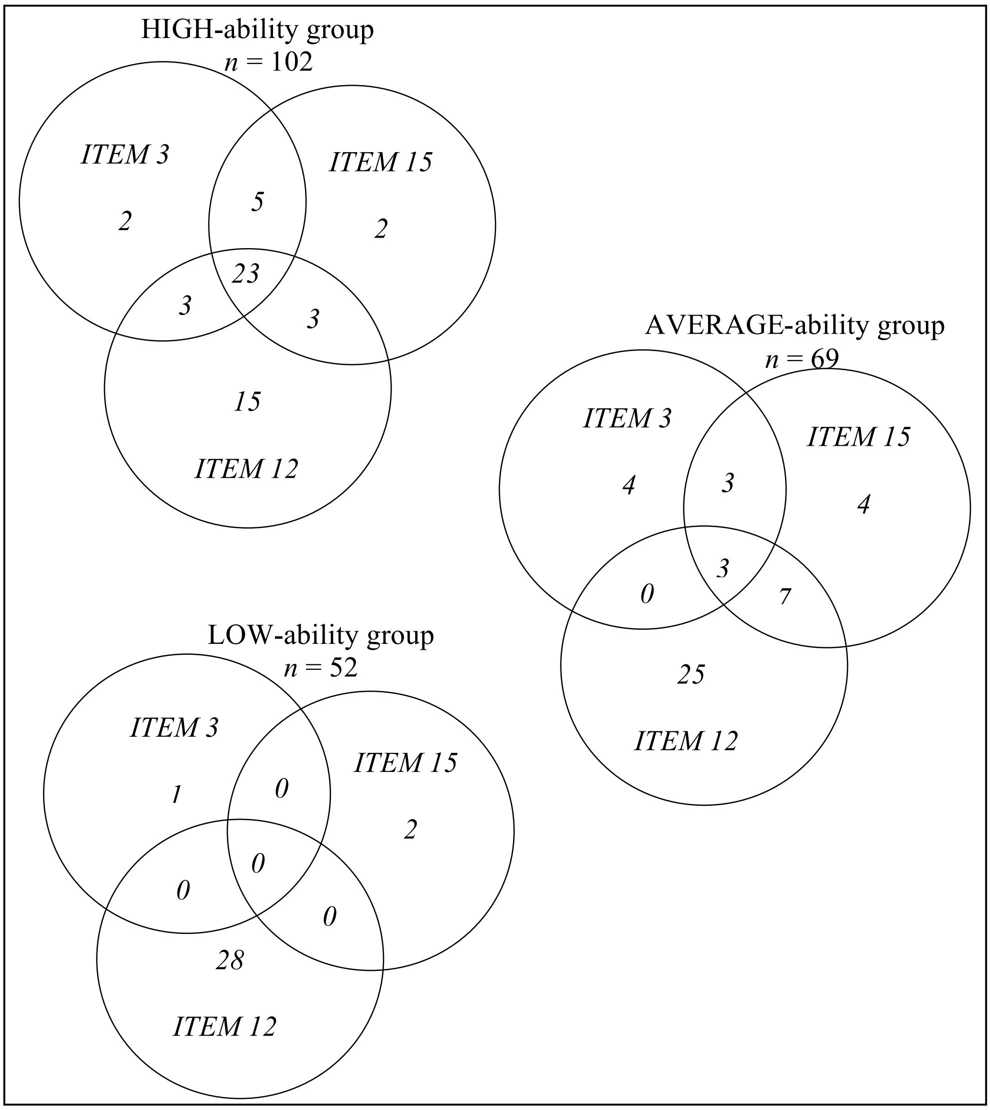


*Figure 5.1: Numbers of student responses to Items 3, 15 and 12.*

Students in the Average- or High-ability groups were apparently aware that Item 12 was mathematically different to Items 3 and 15. However, they were uncertain, or confused, as to the appropriate procedure required to simplify the expression. Hence, the response patterns that were more haphazard than predicted by the model.

## Response Patterns to Item 5

Responses to Item 5 [$2ab + 3b + ab$] can be considered with respect to responses to Item 1 [$3m + 8 + 2m - 5$], Item 2 [$5p - p + 1$] and Item 6 [$5a - 2b + 3a + 3b$] within each of the three ability groups. Each of these four items can be simplified by the addition and/or subtraction of like terms. Item 5 makes the matter more complex by having terms with letter symbols multiplied, and one term with a single letter symbol, but terms need only be added. Item 6 on the other hand requires that terms be either added or subtracted, although each term only has a single letter symbol. Likewise, Item 1 and Item 2 have terms with only single letter symbols, although in both cases, terms need to be added or subtracted.

Numbers of students in each of the ability groups presenting a correct response are given in Table 5.6. These numbers are also expressed as a percentage of the total number of students who participated in the survey ($N = 222$).

*Table 5.6: Number of correct responses to items 1, 2, 5, and 6 [Set 1] by ability groups, also given as percentage of total number of participants (N = 222)*

| | Item Number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Ability Group | **1** [$3m + 8 + 2m - 5$] | | **2** [$5p - p + 1$] | | **5** [$2ab + 3b + ab$] | | **6** [$5a - 2b + 3a + 3b$] | |
| | Number of students | % of all students | Number of students | % of all students | Number of students | % of all students | Number of students | % of all students |
| 1: Low<br>($n = 102$)<br>(46%) | 60 | 0.27 | 61 | 0.27 | 57<br>[35] | 0.26 | 18 | 0.08 |
| 2: Average<br>($n = 69$)<br>(31%) | 59 | 0.27 | 59 | 0.27 | 52<br>[≥ 59] | 0.23 | 20 | 0.09 |
| 3: High<br>($n = 52$)<br>(23%) | 51 | 0.23 | 52 | 0.23 | 44<br>[52] | 0.19 | 40 | 0.18 |

These percentage figures show that students in the High-ability group (the group with the highest mean ability estimate) responded with approximately the same success to all four items, and that most of these students did correctly answer all items. There is a drop in the numbers of correct responses to Items 5 and 6.

Students from the Average-ability group answered Item 1 and Item 2 with the same degree of success, although to a slightly lesser extent than did students from the

High-ability group. Their success rate to Item 5 was also less than for Items 1 and 2, but dropped considerably when responses to Item 6 are considered.

This was also the case with the students in the Low-ability group. Students in this group demonstrated an overall lower rate of success to all four items, but this level was maintained in the correct-response rate to Item 5. This suggests that a greater proportion of students in the lower ability group than that of other two, higher ability, groups were able to deal successfully with Item 5.

The figures in square brackets (Table 5.6) under the raw numbers of correct responses to Item 5 are estimated numbers of students in each ability group who could have been expected to answer Item 5 correctly. According to the Rasch model, ability estimates of the same value as item difficulty estimates mean that students of that ability have a 50% chance of correctly responding to that item. Students of ability estimates greater than the difficulty estimate for an item have a greater chance of successfully responding to that item. As Item 5 has a difficulty estimate below that for the lower bound of the Average-ability group (-1.11 logits), it could be expected that all Average- and High-ability students would respond correctly to Item 5.

The number of students in the Low-ability group with an ability estimate equal to or greater than that for Item 5 (-1.98 logits) was 35, or 34% of the Low-ability group (Figures 4.7 & 4.8). These students could be expected to answer Item 5 correctly. Instead, 57 out of the 102 students in the Low-ability group (more than 50%) gave the correct response.

Using a similar "rule of thumb" from the Rasch model, it could be anticipated that at least 59 students of Average-ability would respond correctly (as for Items 1 & 2), as would nearly all students of High-ability. The numbers for the Average- and High-ability students have been assumed to be equal to, or greater than, the numbers successfully responding to Items 1 and 2. In these latter instances, the actual responses are fewer than anticipated, strikingly so in the case of those students of High-ability.

The misfit of Item 5 might thus be attributed to certain (non-algebraic) approaches taken by the students in the various ability groups, particularly those of Low-ability. High-ability students, and students at the upper end of the Average-ability group appeared, because of their responses, to be more aware of subtleties of notation and syntax. However, with fragile understandings, these students have tended to make confused and incorrect responses more often than might be indicated by the item difficulty estimate.

## Response patterns to Item 30

Item 30 [If $x/4 = 12$, then $x = ?$] fitted the model better than expected – an example of "overfit". The mathematical concepts represented by the item are consistent with the construct being measured, but the item itself discriminated the responses too sharply.

Of the 222 participants, 59 gave no response, and 45 students responded incorrectly. The most common incorrect response was "$x = 3$" (32 responses out of 45 incorrect). Students who reacted to the numbers in the equation seemed to have seen the numbers "4" and "12" and recalled the "3" as in the statement "$3 \times 4 = 12$", but without attaching any meaning to the equation.

An other possible, and related, chain of reasoning was that, as $x$ was divided by four, one had to multiply. The epigrammatic statement "you do the opposite", uttered by students rarely stated the object of the "doing". Hence, all that they thought needed be done was some multiplication. The meaning of the equation was disregarded.

Hence, Item 30 seems to have distinguished students who could read the equation with meaning, paying attention to the structure. Other students, seeing the symbol for "division", could do one of two things: cue multiplication, but remain focused on the "4" and "12"; or, avoid the item.

## Summary

Two items answered successfully by students in the Low-ability group tended to confuse students with more developed algebraic understanding. These two underfit items are of particular importance as they suggest that conceptualisations of mathematical ideas are held by different groups of students in unorthodox ways. Low-ability students responded to these items by using an informal procedure that required no mathematical understanding, but which happened to yield a correct answer.

Students in the Average- and High-ability groups, trying to use procedural rules, but with a limited experience of different representations, offered incorrect responses more often than the difficulty and ability estimates of the model would suggest.

The single overfit item was consistent with the construct of the model as a whole. However, responses to the item suggest that its structure distinguished students who could attend to the meaning of an algebraic statement, from those students who attended only to familiar arithmetic cues.

## MATHEMATICAL THINKING OF STUDENTS IN THE ABILITY GROUPS

Examination of the survey results in terms of the clusters of items indicated stages of conceptual development where learned procedures, insight and ideas needed to be adapted or elaborated in order to deal with more difficult items. This section considers possible conceptual differences in mathematical thinking between students in each of the three ability groups identified by Rasch modelling of the survey data.

As in the previous section, the quantitative evidence of the survey and the model are supplemented by qualitative data from interviews with 31 participating students. Interview data provided insight into the differences in mathematical understanding between the three ability groups. Differences included number sense, arithmetic and algebraic strategies, understanding of the role of letters, and particularly the treatment of terms in expressions. Interviews also revealed that articulation of concepts about the meaning of the equals sign, the treatment of like terms in expressions, and the role of brackets, were expressed in similar ways by all students, although apparently acted on differently.

Students in the different ability groups were able to answer items in the clusters the means of which corresponded with the mean ability estimates for each group. Thus students in the High-ability group could be expected to respond successfully to almost all items in the survey, whilst those students in the Low-ability group could be expected to answer successfully only those items in Cluster 1 (the cluster of least difficulty). Conceptual development of students underpins their ability to move from one cluster of item difficulty to the next.

This section considers general characteristics of understandings that appear to be typical of students in each of the ability groups, as indicated by interview responses. The survey responses and the response-patterns of the Rasch model are discussed with respect to each of the three ability groups: Low-ability, Average-ability, and, finally, High-ability.

## Low-ability

The Low-ability group consisted of 102 students. The ability estimates ranged from -4.89 to -1.28 logits with an average ability estimate of -2.34 logits. Seven students from the group were interviewed. In all cases they tended to talk about single items only from each set. They demonstrated poor number sense, and inefficient arithmetic strategies, such as rote recall of multiplication tables, and/or counting on (five out of seven students) when solving equations; or, resorting to inappropriate use of halving or doubling strategies (three out of seven students).

Some examples from the interviews illustrate these points

> S: [After reading Item 27, $x + 5 = 7$, silently]…$x$ would equal two.
>
> I: How do you know?
>
> S: Because I added two on to five.
>
> […]
>
> I: Why did you choose two?
>
> S: Um…because I went five, six, seven, and I put it in the $x$ spot and then … $x$ equals seven.

One student explained a procedure for solving the equation $4y = 20$ as:

> The way I'd work it out would be "what other number equals 20?". So four times one, four times two … eight. Three times four…three fours are 12. Four fours are 16. Five fours are 20. So $y$ is 5. That's the way I'd work it out.

However, such a strategy failed when the student tried to answer Item 32 [$10y = 5$] by seeking facts that accommodated the numbers in the equation, but without sense being made of the equation itself:

> … Would it be 10 divided by…? Ten fives are 50, so ten divided by 50 equals five. So $y$ would be 50.

There was evidence of a reliance on calculators at all times, even to the point of using it to count on by ones:

> Ummm…[Now speaking about Item 29, $2t$ - 23 = 49] I would add. I would put the 23 and keep adding one until I got to 49. Or minus … or something like that.

Another student expressed the need for a calculator when dealing with Item 18 [$2(x + 4) + 3(x – 1)$] and Item 19 [$2(x + 5) - 8$].

Students tended to act on the surface features of items according to simplistic, non-mathematical procedures that they had learnt, such as circling like terms.

> Well, first I circle the $3m$ and $2m$ [Item 1], I circle the sign in front of it and I circle the eight and the five, and then I do it all together.

All students in this group described the procedure of cancelling in fractions as merely the "crossing out" of the same letters or numbers.

> S: …And then for eight [Item 8: $4ab/4b$] I just cross them off like how we did in class. And then just write down the answer, which would be $a$.
>
> I: Can you tell me more about this "crossing off"?
>
> S: Well, when they are the same, when the top and bottom…. You can cross them off and then get left with the answer… It's what you do.

Although this led to a correct simplification of expressions such as that in Item 8, the procedure is limited to situations where numerator and denominator of a fraction, written as one term, contain the same letters and/or numbers. Hence many students, in all of the ability groups, who used such a strategy without understanding the underlying mathematics, found Item 16 [$2/a^2 \times 5a/4$] very difficult to simplify completely.

Only three students could state that letters represented numbers. Five students either did not have any idea of what the letters stood for, or thought they could stand for a "thingummy", as one student so eloquently expressed it, or served to "put a complication in". Survey data show that this group had the greatest tendency to conjoin terms in items where relationships between terms were expressed as addition or subtraction.

Many students in this group failed to complete the survey. The greatest number of non-attempts for items occurred with items on the last (third) page of the survey.

These items were mostly equations (Table 5.1). One reason might be that inefficient strategies coupled with poor recall of number facts was so time-consuming or frustrating that many students gave up or ran out of time. The reason for many incorrect responses, other than non-attempts, can be attributed to the students mis-remembering procedures.

### Average-ability

The Average-ability group consisted of 69 students. The ability estimates ranged from -1.17 to 1.00 logits with an average ability estimate of -0.15 logits. Fourteen students from this group were interviewed. These students tended to answer each individual item in each set, except for Set 4, which dealt with fractions. Three students avoided giving any responses to this set of items, claiming "it doesn't make sense". Others confessed that their fraction skills were "absolutely terrible".

Most students in this group demonstrated sound number sense and efficient simple number fact recall. Five students admitted needing a calculator when solving equations, using guess-and-check procedures, and when working with fractions.

> I'd probably go just 63 [Item 36, $63/x = 180$] and I'd start, I'd probably start, usually I start with um, any number. Say I start with maybe a three, and then if it's absolutely… A number which is very low then, maybe I'd try 63 over ten. If it was a number which was too high then I'd work my way back between something that has to be between three and ten. So it takes a while, but I usually I find that the most reliable way for me for doing it.

Four students stated that they had no idea what letters represented in algebraic expressions, whilst seven students indicated that letters stood for numbers. Some students said that letters represented *anything* when used in expressions, but numbers when in equations. This type of response reflected students' experience of finding the numerical value of a letter symbol in an equation. When a numerical value cannot be found, as in Item 39 [$ax = 5$], eleven of the students in this group attributed one of the factors of five to $x$ (1 or 5), or simply stated that each letter took the value of *2.5*, instead of writing or stating that "$x = 5/a$". One of these students reached the conclusion that $x = 5/a$ with prompting.

In statements similar to those of students in the Low-ability group, eight of the fourteen students in this group described cancelling merely as "crossing out", although one of these students was prompted to state that the operation of division was involved. One other student volunteered that cancelling was a division procedure. Many of these students seemed to understand that, because of the structure of the items in this set (Set 4, fractions), division was involved, but that division meant that one "crossed out" or "took away" the same letter or number, as a physical rather than a mathematical act.

The students in the Average-ability group were able to solve equations that had the unknown on one side, by "reversing the equation", "doing the opposite", or "backtracking" when a verifiable numerical answer was possible. They had difficulty dealing with equations with the unknown on both sides of the equal sign as in Item 34 [$x + (x + 2) = (x - 1) + 8$ and Item 37 [$5a - 4 = 2a + 8$].

> I'd probably… These ones, I'd use, like, two simple digits, like five times… Like two or something like that just to start off with and then see if [it] equals this one and, if it didn't, then I'd just be like … And then, see if this would work. I usually do the easier ones first and then [unclear] I go like, what the answer is and what it looks like to see if I should go up higher or lower or… I dunno.

Students in this group also found difficulty with equations where the answer was not readily verifiable, as in Item 36 [$63/x = 180$], or unique, as in Item 39 [$ax = 5$]. In the cases of Items 34, 37 and 36, most students used trial-and-error procedures, and were hampered by their lack of access to a calculator. Item 36 also posed problems for those students who "reversed" the equation incorrectly. This indicated that students did not understand the mathematics behind reversal, but used a procedure suggested by the arithmetic operations they identified in the equation.

> For the next one [Item 36], 63 over something is 180….63 divided by something is 180 so I'd probably do 180 times 63, so that I could divide it by 63 to get 180 again. So then $x$ would equal 63 times 180, I'm not really too sure. Or maybe it's divided by…I don't know actually…
>
> […]
>
> $x$ would have to be less than one because then, instead of dividing it would like timesing [*sic*], so, I'd probably divide 180 by 63, and then … I don't know, um, yeah.

The following exchange illustrates many of these aspects of thinking by students in this group: Use of inverse operations; an efficient recall of number facts; and, a sound number sense – in simple situations at least. The early part of the conversation suggested that the student might be thinking algebraically, but the latter part of the conversation gives the lie to this as the student struggled to deal with the need to provide a non-unique expression as an "answer".

S: [Reads] If *ax* equals five, *x* equals what? [long pause for thinking] Oh, *x* is 2.5.

I: Why?

S: There's equal parts of those two [*a* and *x*]. It doesn't say 2*a* and *x*, it just says *ax*.

I: But what do *a* and *x* stand for?

S: Any numbers timesed.

I: Why did you decide they were the same?

S: Because I just did. I have no idea why. It's just how I worked it out.

I: Go back to 28. What did you do in question 28? [Item 28: 4*y* = 20].
How did you go about finding *y*?

S: I divided 20 by 4, and I got *y*.... Hang on. Would that be like, say, if that was ...
if *a* is ten, that'd be like ten times *x*, or ten times *y*...?

I: It is. But *a* is not ten, so how would I go about writing it as *x* equals something?
S: *x* equals... I'm not sure what it equals. I have no idea.

Few students in this group tended to conjoin terms when the items required them to simplify expressions that were typical of textbook examples (symbolic). However, many more did so when dealing with items that used words to describe algebraic ideas, such as those in the Semi-literal category.

### High-ability

The High-ability group consisted of 52 students. The ability estimates ranged from 1.1 to 4.93 logits with an average ability estimate of 1.995 (2.00) logits. Eleven students from the High-ability group were interviewed. Five of these students responded to all items in each set as presented to them, while six students gave a broad, general response to the first set of items. This global response (required by the introductory question of "What goes on in your head when you meet these expressions/equations?") was given by two students only as responses to the other sets of items. Both these students were at the upper end of the range for this group.

The students in this group displayed sound number sense and reliable recall of number facts. Two students did, however, state at points in the interview that they would use a calculator, but then proceeded to state, unhesitatingly, the number facts needed. Most students described their equation-solving procedure in terms of "reversal", "doing the opposite" or "backtracking". Five were able to apply these procedures to Item 34 [$x + (x + 2) = (x -1) + 8$] and Item 37 [$5a - 4 = 2a + 8$] where the unknown was on both sides of the equation. This contrasts with students in the Average-ability group who also used similar descriptions of their equation-solving procedures, but who could not apply the procedures to these items. Thus, phrases such as "doing the opposite" seem to have taken on a broader, more conceptually oriented meaning for some students in this highest ability group. Other students used trial-and-error to solve the equations in Items 34 and 37, but recognised that there needed to be a more efficient method.

> S: Well, in this one [Item 34], I simplify both sides first, $2x$ plus two equals $x$ minus seven… so if $2x$ plus two equals $x$ minus seven that means that … [writes]. Doesn't make any sense.
>
> I: Why not?
>
> S: Because $2x$ plus two can't equal $x$ minus seven.
>
> I: Why not?
>
> S: Because we are saying two lots of a number plus two can't equal that same number minus seven.
>
> I; Why not?
>
> S: It's not equal, because, two lots of a number plus two are not equal to one lot of a number minus seven.
>
> [Note the refusal to admit negative numbers as possible solutions, as well as an explanation that simply repeats the premise.]
>
> I: OK. So you are trying to find a number in your head that will fit that. Is that what you are trying to do? Are you trying to find a number that would suit that?
>
> S: Oh, wait a minute, it's plus seven, it can work … Yeah, I did it by guess-and-check, I think. I don't know how to work them out properly.

All of these students recognised that letters represented numbers in algebraic expressions as well as equations. One student avoided responding to Item 39 [$ax = 5$]; three gave a factor of five as the answer, because they felt that "it's better just to use a whole number… it's just easier"; three were prompted to see that the answer was $x = 5/a$; and, two students volunteered the correct answer. Five students in this group volunteered that cancelling was a division procedure. Three came to that conclusion after questioning that encouraged them to rethink their answers.

## Summary

The three ability groups identified through the Rasch modelling were identified from clusters of ability estimates, the means of which were found to be significantly different. Each group could be distinguished by different conceptual approaches articulated by the students during interviews, and demonstrated by the responses on the survey.

Students in the Low-ability group relied on remembered, simplified procedures that depended on visual clues in each item. They also relied heavily on arithmetic strategies, although their recall of number facts and application of procedures was often inefficient and inaccurate. Items tended to be read from left to right. Thus, meaning that had to be found by interpreting the syntax of the expressions was frequently disregarded. Many failed to understand that letters in algebra had mathematical meaning.

Students in the Average-ability group displayed a sounder number sense than those in the lowest group and more efficient number fact recall. These students still relied on rehearsed procedures and were uncomfortable when having to deal with unclosed answers to items, or items where the structure or syntax required them to make inferences. These students were able to solve simple equations with the unknown on one side, and some, given their sound number sense, were able to solve other equations by trial-and-error substitution methods. This group demonstrated an understanding of mathematical relationships, provided that they were situated in an arithmetic context.

Students in the High-ability group demonstrated their understanding of algebraic notation and syntax, and used algebraic strategies rather than trial-and-error where appropriate. However, most struggled with abstract items requiring inferences to be made from the notation used, and also when a complete response drew on deeper understanding of mathematical processes.

In order to deal effectively with increasingly formal algebra, students needed to have increasingly sophisticated concepts of arithmetic structures and relationships, and to be able to apply these to the general cases of algebra. Despite many students across the range of ability expressing their understanding in very similar ways, their responses to the survey indicated different conceptual development.

## CONCLUSION

Student ability estimates clustered around three significantly different means. These means approximated the means of the three item clusters. This indicates that the mathematical approaches used by the groups of students are those which allow a successful response to items in a cluster of items that have the same, or a lesser, difficulty estimate than the ability estimate of the group. The same sorts of approach would not, however, allow success in items in a cluster of greater difficulty. The procedures used by students in each of the ability groups, and articulated by a sample of students from each group, served to explain the response and error patterns and the possible reasons for the difficulty levels of clusters of items.

Students who relied on informal, almost amathematical, procedures, based on simple arithmetic manipulations could only deal successfully with the easiest items in the survey. These were students in the Low-ability group. Students in the Average-ability group appeared to use better-understood, but often informal, rule-based procedures that allowed little flexibility in their response, because of the associated lack of underlying conceptual understanding.

Students in the Average-ability group, and students in the lower end of the High-ability group, displayed inconsistent responses, indicating a confused, transitional state of concept development. An increase in ability was marked also by the students' improved number sense, ability to recall accurately particular procedures, and, for the most able, to be able to work with abstractions – expressions and equations that could not be closed.

The gaps between either item clusters or ability group clusters might therefore be seen as points where a change in understanding of hitherto accepted mathematical ideas is essential for further conceptual development. The types of items successfully responded to by the ability groups, and the ways in which the students described their thinking suggest a link with the SOLO framework.

Students in the Low-ability group often described their procedures in terms of physical acts, or acts on terms in expressions that looked alike. Such descriptions indicate that these students depend on similarity of appearance of mathematical objects, but without associating this with any deeper mathematical meaning. Thus, their thinking about algebraic items relies heavily on ikonic support, although they are able to use known arithmetic facts to answer simple items, with no more than one or two steps. Their number sense, requiring connections to be made between arithmetic facts, tended to be very poor. Less-able students focused on one aspect of more complex items, just as they often based their interview comments on a sole example from a set of items.

Although still relying on the appearance of items, students in the Average-ability group were more likely to depend on connected arithmetic understandings to deal with items, and be able to express their thinking in more formal, conventional ways. Students in the Average- and High-ability groups were able to respond successfully to items requiring many steps and some interpretation of mathematical syntax and notation. Thus, items in Cluster 2 were accessible to students of Average-ability, because they could be manipulated in a series of closed steps, and validated by arithmetic solutions.

The more able the students, the better their memory for appropriate procedures. However, throughout the interviews, it became apparent that most of the students found it difficult to articulate any conceptual basis for the procedures they used. More able students appeared to apprehend the reasons, but were unable to articulate explicit understanding.

Disconcertingly, explanations by students regardless of ability or item difficulty tended to be phrased in similar ways, often as standard, epithetical classroom instructions, of limited mathematical application. However, the different success rates of the students suggest that the understandings behind the utterances must be different in quality. A closer study of these utterances should provide insight into that qualitative difference.

The following two chapters analyse severally, and together, linguistic features of students' explanations of their algebraic thinking. This addresses research questions 4 and 5; those questions the answers to which provide data on linguistic features of students' language as they discuss their algebraic thinking.

# CHAPTER 6: ANALYSIS OF LANGUAGE USED BY STUDENTS

The spoken word is immediate, and transient. It conveys both knowledge, and attitude to that knowledge, by the speaker. A listener makes inferences about the speaker's knowledge not only from what is said, but also from the way it is said. Attending only to the content of students' statements, although necessary, is not always sufficient for teachers to be able judge the quality of students' understanding.

Students, demonstrating a range of ability on an algebra survey, tended to explain their thinking using similar words and phrases. Some of these were part of the mathematical register; others were of a less formal classroom register. Common usage across a range of abilities suggests that students understand these expressions differently. Hence, attention needs also to be paid to the pragmatics of students' utterances.

*Pragmatics* is the study of the use of language by participants in discourse to convey knowledge and negotiate social relationship. How they use language in classroom discourse indicates the extent to which students might understand the ideas about which they speak. Thus, analysis of verbal responses by students as they explain their thinking about various algebra items can provide insight into the security of their understanding. Linguistic features such as: verbosity of explanations; use of pronouns; and, the type, modality and tense of explanations, convey information about students' confidence and security of understanding.

The previous chapter used interview data to interpret patterns of item difficulty and student ability arising from Rasch modelling of the survey responses. This chapter explores the developmental significance of linguistic features through a pragmatic analysis of the interviews. Two questions to be answered by this analysis are: "What linguistic features of students' conversations characterise ability?" and "What linguistic features change as item difficulty changes?" (i.e., as student knowledge becomes insufficient to deal with particular algebra concepts).

Interviews were analysed with respect to the ability groups of the students, and also with respect to the difficulty of each item set. Therefore, the first section of this chapter revisits the main features of ability groups and item clusters identified by the Rasch-modelling of survey data that were detailed in Chapter 4. The five subsequent sections describe and discuss the analyses of the interviews. Linguistic features that form the focus of the analysis are: the verbosity of students' explanations; the frequency with which students used the pronouns *I, you* and *it*; the type of response whether attitudinal (affective), specific or general; the modality of the explanations as indicated by hedges and shields; and the tense used by students.

## ABILITY GROUPS AND ITEM CLUSTERS

When the data from the algebra survey were Rasch-modelled, three clusters of item difficulties and three groups of student abilities were identified. These groupings have been used to organise the analysis and discussion of interview data. The main features of the ability groups and the clusters of item difficulties are summarised in this section.

### Ability Groups

It was apparent from the interview transcripts that particular individuals responded in idiosyncratic ways when asked to explain their thinking about particular items or sets of items. However, within any one group of individuals, these differences might be balanced, or reinforced, by the responses of others. Accordingly, responses were analysed as group responses. The decision of which individuals to include in a group was made on the basis of the ability groupings determined from the Rasch modelling of the survey (Chapter 4 and 5). Table 6.1 summarises the features of the three ability groups: the number of students in each group; the ability range of the groups; and, the average ability of the students in the group who were interviewed. The average ability estimate for all students in each group is also given in parentheses in the right-hand column.

*Table 6.1: Summary of students interviewed (n=31) in ability groups, with ranges of the ability estimates, and averages*

| Ability Group | Number of Students | Ability Range (logits) | Ability Average (logits) |
|---|---|---|---|
| HIGH | 10 | 1.27 to 3.6 | 1.87 (2.00)* |
| *High-3* | *3* | *2.58 to 3.6* | *3* |
| *High-7* | *7* | *1.27 to 1.65* | *1.4* |
| AVERAGE | 14 | -0.94 to 0.93 | -0.01 (-0.15)* |
| LOW | 7 | -2.86 to -1.28 | -2.18 (-2.4)* |

*Average ability estimate for all participants in the nominated cohort (Chapter 4)

The High-ability group comprised two sub-groups, delineated by a gap of 0.9 logits. One sub-group (referred to as High-3) consisted of the three students with the highest ability estimates (3.6, 2.87 and 2.58 logits). The second sub-group (referred to as High-7) consisted of the remaining seven students with ability estimates ranging from 1.27 to 1.65 logits. This sub-grouping of the students in the High-ability group has been used in some of the analyses. Other analyses use the larger cohort of High-ability students.

## Item sets

Items for the interviews were grouped in sets according to common algebraic concepts. These sets mainly followed syllabus categories described in earlier chapters (Chapters 4 and 5). However, items could have been allocated to more than one category, and thus, interview sets differed from these categories by the reallocation of some items, or the inclusion of some items in two sets.

Linguistic features of interview responses were analysed in terms of their occurrence in the sets of items presented to the students. The average difficulty of each set was determined from item difficulty estimates from the Rasch model (Chapter 4). Table 6.2 summarises these data in terms of the set number (as presented at the interviews), the identification number of each item in the algebra survey included in the interview set, a brief description of items in the set, and the average difficulty estimate for each set.

*Table 6.2: Summary of survey items arranged in sets for interviews, with average difficulty for each set*

| Set | Item Numbers | Item Description | Average difficulty |
|---|---|---|---|
| 1 | 1, 2, 5, 6 | Expressions to be manipulated by the addition/subtraction of terms | -1.66 |
| 2 | 4, 9, 13 | Expressions to be manipulated by multiplication | -1.56 |
| 3 | 20, 18, 19, 7 | Expressions to be manipulated with brackets to be expanded, terms to be added/subtracted | 0.425 |
| 4a | 3,15 | Expressions with fractions to be added/subtracted | 1.385 |
| 4b | 8, 12, 16 | Expressions with fractions to be multiplied/divided | 0.82 |
| All 4 | 3, 15, 8, 12, 16 | Expressions with fractions, to be manipulated by addition/subtraction or multiplication/division | 1.046 |
| 5a | 27, 29, 37, 34, 31 | Linear equations to be solved, one or several steps, unknown on both sides, brackets | -0.23 |
| 5b | 28, 32,39 | Linear equations – multiplication, one literal | 0.14 |
| All 5 | 27, 29, 37, 34, 31, 28, 32, 39 | Linear equations to be solved – one or two steps, addition/subtraction/division/brackets | -0.08 |
| 6 | 30, 33, 36 | Linear equations to be solved – fractional forms | 0.34 |
| 7 | 17, 10, 7, 11 | Expressions with brackets – including powers of 2 | 1.92 |
| 8 | 21, 22, 25, 26, 40 | Items adapted from Küchemann (1981) | -0.09 |

Set 4 and Set 5 each consisted of sub-sets of items. The average difficulty is given for the entire set, and also for the sub-sets. One other set of items, Set 9, was also used during interviews to establish any background linguistic features that might differ when students spoke about ideas other than algebra. Set 9, by its very nature, could not be Rasch-modelled, but it is referred to in analyses that follow.

The results of the analysis of interviews are described and discussed in the following sections. Firstly, the interview responses are analysed in terms of *verbosity*; the numbers of words with which students responded to the interview questions. Subsequent sections examine: how the use of pronouns changed with student ability and item difficulty; how types of responses changed; changes in modality; and lastly, changes in tense.

## VERBOSITY

Verbosity refers to the number of words used by students in the course of the interviews. The number of words with which students answered interview questions varied according to the willingness, or reluctance, of individual students to engage in a conversation with the interviewer. The numbers of words spoken have been taken as a possible indicator of the students' confidence in the subject matter. The use of overlong and elaborate explanations might also serve to mask a refusal, or lack of knowledge (Rowland, 2000).

General patterns of speech behaviour, such as verbosity, might be expected to vary according to ability or concept (item) difficulty. The first sub-section describes changes in verbosity with respect to the ability groups; the second, verbosity with respect to the difficulty of the sets of items. A third sub-section discusses the results.

### Verbosity and ability groups

Although general patterns of verbosity emerge from examination of ability groups, it is useful to consider the tendency of individual students to be either verbose, or laconic, within each of these groups. The words used, by each individual student, are shown as standardised scores in the scatter plot, Figure 6.1. In each ability group, numbers of words used set-by-set, and during the entire interview, were counted. Vertical lines at approximately ±1 indicate approximate ability estimate thresholds between Low- and Average-ability groups, and Average- and High-ability groups (see Table 6.1 for details). Horizontal lines indicate one standard deviation from the mean numbers of words used.

A wide variation in the verbosity of the students, particularly those in the Average-ability group, is illustrated by the scatter plot. Half of this group was equally or more verbose than any of the students in the Low- or High-ability groups. The High-ability group tended to be less verbose than the other groups, although the number of words used by students in this group rose as the ability of the students approached the threshold with the Average-ability group.

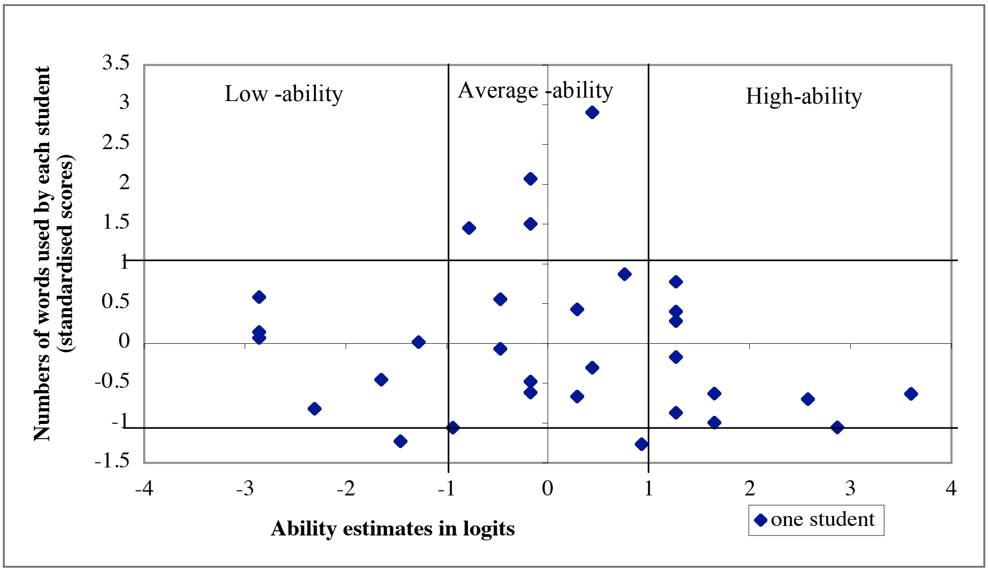


*Figure 6.1: Numbers of words as standardised scores used by individual students ordered according to Rasch ability estimates.*

The cluster of three students with the highest ability estimates (High-3) can be seen to the far right of the figure. The remaining seven students (High-7) form a cluster near the vertical line representing the threshold between the High- and Average-ability groups at 1.0 logit. Verbosity of students in the Low-ability group tends to increase towards the lower end (left hand side of the figure) of the ability range.

Table 6.3 details: ability ranges for each group; the number of students interviewed as absolute numbers, and as a percentage of the interview participants; and, the overall number of words contributed by each of the ability groups to the interviews as raw numbers, as a percentage of the total number of words in the interviews, and as an average number of words per student in each group. The last row gives totals overall, and the overall average number of words per student.

*Table 6.3: Average numbers of words used by ability groups*

| Ability Group | Ability Range | Number of Students (% of total interviewed) | Number of words | Percentage of total words | Average words/student |
|---|---|---|---|---|---|
| HIGH | 1.27 to 3.6 | 10 (32%) | 17942 | 28% | 1794 |
| AVERAGE | -0.94 to 0.93 | 14 (45%) | 33384 | 52% | 2385 |
| LOW | -2.86 to -1.28 | 7 (23%) | 13216 | 21% | 1888 |
| | Overall | 31 (100%) | 64542 | 100% | 2082 |

The proportion of words (expressed as a percentage) contributed by the different ability groups differed from the proportion of students in each of those groups. Any one particular student might talk more, or less, than another according to individual personality and reaction to the interview context. By aggregating individual responses into responses by groups of students, it was assumed that individual differences would be compensated.

If it is assumed that each student contributes approximately the same number of words to an interview, then the proportions of words contributed by the total number of students in each group should be equal to the proportion of students in the group. It could, then, be expected that the average number of words per student would be the same for each ability group. That this is not the case is shown in Table 6.3.

The High-ability group contributed fewer words than might be expected from the proportion of students. This was also the case with students in the Low-ability group. Students in the Average-ability group contributed more words than might be expected from the number of students. The average number of words per student for the High- and Low-ability groups was approximately the same, whereas the average number of words per student in the Average-ability group was higher.

Differences of the actual numbers of words to those expected are significant (*p*<0.001, Appendix D2). Table 6.4 summarises expected and actual numbers of words contributed by each of the ability groups.

*Table 6.4: Comparison of expected number of words contributed by ability groups and actual numbers of words*

| Ability Group | Number of Students | Proportion of Students | Expected number of words by group | Actual number of words by group | Proportion of words by group |
|---|---|---|---|---|---|
| HIGH | 10 | 0.32 | 20820 | 17942 | 0.28 |
| AVERAGE | 14 | 0.45 | 29148 | 33384 | 0.52 |
| LOW | 7 | 0.23 | 14574 | 13216 | 0.20 |
| Totals | 31 | 1.00 | 64542 | 64542 | 1.00 |

Because it was numerically the largest group of students interviewed, the Average-ability group could be expected to contribute more words overall to the interviews. If the assumption that ability (and hence assuredness of understanding) has no influence on verbosity, then the number of words contributed by the group of average students should approximate the proportion of students in the group. This was not the case (Table 6.4). The proportion of words contributed (52%) was greater than the proportion of students (45%) in the group. On the other hand, both the High- and Low-ability groups contributed proportionately fewer words than could be expected given the proportion of students in the interview sample. Hence, verbosity can indicate that students are operating in a conjecturing environment, where their insecurity of understanding might be expressed by unnecessarily long responses.

### Verbosity and item sets

If over-long responses indicate the degree of confidence students have in their understanding of algebraic concepts, then verbosity of students' explanations could change as students are asked to talk about items of different difficulty, with respect to their individual conceptual status. Figure 6.2 illustrates changes in verbosity of each of the ability groups as students answered each item set. The sets are ordered left to right in ascending order of average difficulty (Table 6.2). Verbosity is expressed as an average number of words per student in each of the ability groups, for each item set.

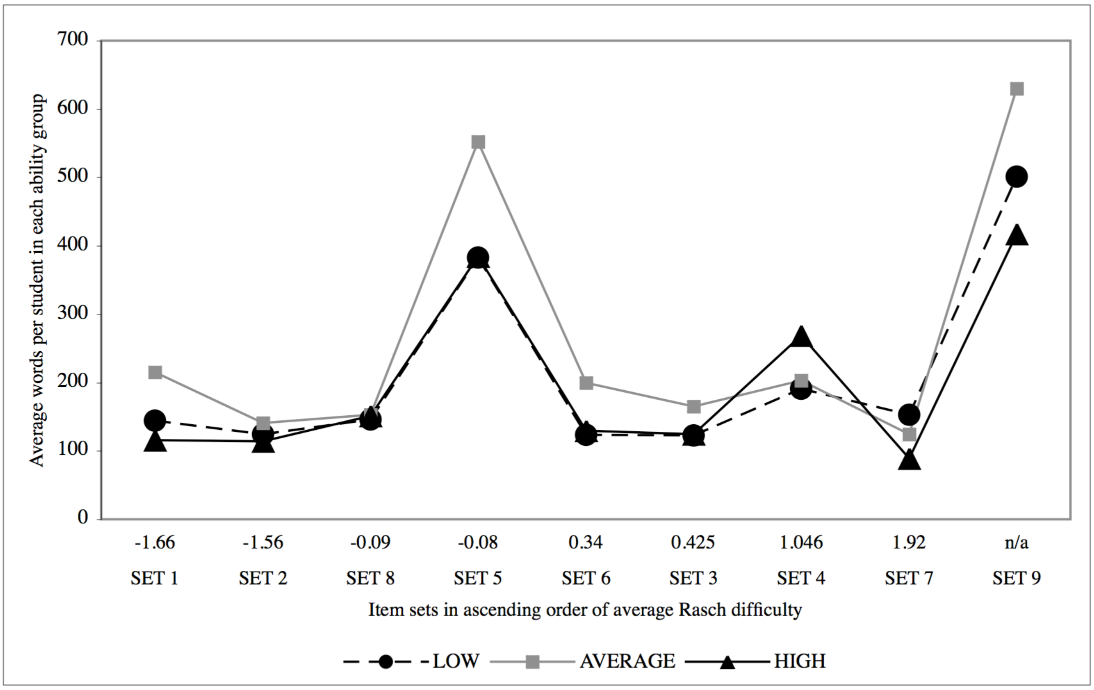


*Figure 6.2: Verbosity of students in ability groups by item-sets in order of average difficulty*

Reading from left-to-right, the verbosity of each group tended to follow the same pattern, spiking at Set 5 and Set 9, both of which sets consisted of more questions than other sets. Set 5, the greatest number if items, consisted of simple linear equations to be solved. Students in each of the ability groups seemed to be more comfortable discussing equations, and attempted to answer each of the items, thus finding more to talk about.

When answering Set 9 (background questions) all groups became more verbose. The Low-ability group contributed more than usual, the High-ability group the least. The proportion of words contributed by each group approximated the proportion of students in each group. The Low-ability group (20% of interviewed students) contributed 21%, Average-ability (52% of interviewed students) contributed 54%, and High-ability (28% of interviewed students) contributed 25% of the words in Set 9.

The Average-ability group shows greater variation in verbosity as different sets of items are answered. The verbosity of all ability groups dropped as they addressed items in Set 6. There were fewer items in this set of equations, but these were also more difficult equations than those in Set 5.

The pattern of verbosity of students in the Low- and High-ability groups shows little difference between the groups, except with responses to items in Set 4. Students in the Average-ability group tended to be more verbose in their responses to almost all other sets (see Figure 6.2), except for Set 4. Items in this set consisted of expressions requiring manipulation of fractions. Students in the High-ability group tended to answer all or most items, whilst students in the Low- and Average-ability groups often declined to respond, or chose a single item to discuss – briefly. Average-ability students tended to focus on the multiplication and division examples, principally Item 8 [$4ab/4b$], but were reluctant to deal with the items requiring addition or subtraction of fractions [Item 3, $a/5 + a/10$, Item 15, $3p/4 - p/8$].

Students in the High-ability group addressed each item in the set, although not always correctly. Hence, they found more to say about these items, often generalising, although they struggled with the mathematical ideas, evidenced by the increased verbosity and other linguistic features discussed in later sections.

The Low-ability group became slightly less verbose than expected when addressing items in Set 6, and more verbose when talking about Set 7. Set 6, equations with fraction notation, requiring understanding of algebraic syntax, and arithmetic competence, challenged these students. On the other hand, Set 7 consisted of items with brackets, which most students felt able to address, at least in part. The students' apparent confidence when responding to items in Set 7 seems to be at odds with the high average difficulty estimate. This number is largely influenced by the inclusion of the most difficult item [Item 10, $(x + y)^{2}$, difficulty estimate 3.56 logits], which most students attempted, but incorrectly (Chapter 5, Table 5.4, Error Analysis).

Differences in verbosity between ability groups for each set of items were significant, except for Set 2 (expressions requiring multiplication) and Set 3 (expressions with brackets). Overall, the difference between the actual and expected verbosity for each ability group across all sets of items was significant ($p<0.001$ Appendix D1).

## Discussion

Changes in the verbosity of responses appeared to reflect the ability of students and the difficulty of sets of items. Verbosity is discussed, firstly as patterns of linguistic behaviour by the students in ability groups, and secondly, as patterns of verbosity emerge when sets of items are considered.

### *Ability Groups*

Relative lack of verbosity by students in the Low-ability group could be attributed to the fact that many of these students responded to a single item only in any set, or avoided answering any at all ("I don't know"). Even when Set 5 (equations) is considered, the Low-ability students answered fewer items than students in the other ability groups (Table 6.8, in the section discussing Pronouns).

Students in the High-ability group exhibited a lack of verbosity similar to that of the Low-ability students. This could be attributed to their tendency to identify and articulate some general, mathematical features existing between items in the sets, rather than their being able only to give limited responses to single items. The High-ability students often felt that their rule-like generalisation was sufficient explanation of their thinking; prompting failed to elicit further elaboration. When High-ability students did address all individual items in a set, they did so without the marked hesitations and digressions shown by the Average-ability group.

Verbosity of students in the Average-ability group was the combined result of frequent occurrences of hedges, (discussed in the section on Modality), false starts, and a tendency to address separately each item in a set. These students did not articulate any general pattern that might exist between items in a set, but responded to the items as if they were textbook exercise examples, indicative of multistructural thinking. No connection was apparently made between one item and another.

It is the convoluted nature of the explanations given by the Average-ability group, the numerous false starts, and the focus on individual items that provide an indication of their level of algebraic confidence and, hence, understanding and consequent success in the survey. The use of convoluted and unnecessarily long explanations is inferred by the listener to be a shield for lack of substantial knowledge on the matter discussed. For example, the response of a student in the Average-ability group to an item in Set 8 is typical of responses from this group of students, although not so typical of responses to Set 8.

> This one? [Reads very softly, Item 26: If $p + q = 5$, then $p + q + r = ?$] Three things came into my head. One of them was I'd probably, like... My first immediate reaction was $5r$ but then I thought about it and I might even... Another thing I might have tried would be $p$ plus $q$ then I'd, like, say half of five is two and a half, and then to the two and a half plus two and a half... Plus two and a half... two and half plus two and a half equals five and then I would say. So it is seven and a half.

### *Sets of items*

The nature and difficulty of items in the sets appeared to influence the verbosity of responses. Except for items in Set 1 (simplification of expressions requiring addition and subtraction of terms), students appeared to be more comfortable when dealing with equations rather than expressions.

The rise in the number of words contributed by each of the groups when answering questions about items in Set 5 is due, in part, to the greater number of items in the set. Students also appeared to be more confident in responding to each item in Set 5, and consequently had more to say. Set 5 was of average difficulty (-0.08 logits), and hence many items should have been accessible to students in the Average-ability group, and all accessible to students in the High-ability group.

Set 8 consisted of Küchemann items (Table 6.2), and the interviewer's request that students "read the items aloud and tell me how they might be rewritten" did not encourage verbose responses. Students in the High-ability group read the items aloud and rearranged the items, simplifying where necessary, without hesitation. Some students in the Average-ability group read the items aloud, and often made several attempts to rearrange the items, suggesting several possibilities, as in the example quoted above.

Students in the Low-ability group tended to read the items aloud (as instructed) but then simply restate each one, or give no further response.

> S: Add four on to $n$ plus five. It would be $n$ plus five plus four.
>
> I: Any others?
>
> S: Add three on to $4n$, $4n$ plus three. Take $n$ away from $3n$ plus one, $3n$ plus one take away $n$.
>
> I: Could you write that another way?
>
> S: $3n$ minus $n$ plus one.
>
> I: Any others?
>
> S: Um…no. If $p$ plus $q$ equals five then $p$ plus $q$ plus $r$ equals… I don't know how I'd write that. If $e$ plus $f$ equals eight, then $e$ plus $f$ minus $g$ equals… I don't know how to write that either.

Verbosity of response might therefore be taken as an indicator of confidence engendered by the familiarity of appearance of an item, but not necessarily of success or deep understanding. Overlong responses indicate a cognitive struggle to articulate ideas that are apprehended at an intuitive level only. This becomes apparent when the responses of the High-ability group to Set 4 are considered. Whilst students in the Average- and Low-ability groups declined to respond to many items in Set 4 (fractions) or responded selectively to one or two items, the High-ability group tended to respond to all items, although some struggled to do so correctly.

Set 4 consisted of two subsets, one of which had an average item difficulty of 1.385 logits (Table 6.2). This difficulty estimate approximated the average ability estimates for the High-7 group (1.4 logits, Table 6.1). Thus, the students in this sub-group could be expected to struggle with the algebraic ideas in the items, expressing that struggle in overlong statements where false starts indicated that they recognised errors but were unable to identify the nature of the error. This could explain the atypically increased verbosity of the High-ability students and the comparatively lower verbosity of the Average- and Low-ability students in responses to Set 4. The students in these latter two groups either did not respond to any of the items, or chose to answer a single item only.

However, terseness, where a rule is simply quoted, need not guarantee understanding. Students tended to be less verbose, and confident, when responding to Items in Set 3 and Set 7, although the average difficulty estimates for both sets was greater than the overall average ability estimate (Table 6.2). Both sets consisted of expressions with brackets. The relatively high difficulty estimates for individual items in these sets indicates that many students lacked a robust understanding of the algebraic significance of brackets. The students' confidence, and hence concise responses to these sets, might be explained by the fact that they either answered selected items, or provided a generalisation, rule or observation such as, "It's brackets".

The brief responses by students in the Low-ability group are the result either of their quoting a rehearsed rule, such as "[brackets] mean do it first", or of their selecting one item only from a set of items. The low success rate of students in this group to items in the survey suggests that their understanding of the rules they remember and use is incomplete or inaccurate.

This can be contrasted with the similarly brief responses of students in the High-ability group, who quoted rules or simple generalisations rather than answer every item in a set of items. Their success in answering the survey suggests that they have a sound understanding of the mathematical implications of the rules they use, and a flexible understanding of the mathematics symbolised by notation such as brackets. Attempts to answer all items as disconnected entities, inevitably leads to a greater verbosity, as was the case with the Average-ability group of students.

Therefore, verbosity of a response could indicate a lack of confidence, insecurity of understanding or a complete lack of knowledge. It might also indicate that students are operating in a conjecturing environment – struggling to understand and articulate new ideas. Other features of student utterances, such as use of personal pronouns, provide further evidence of knowledge security, or insecurity.

## USE OF PRONOUNS

Pronoun use is influenced by power relationships that necessarily exist in a classroom. When talking to a teacher, or interviewer, students will use *you* as a substitute for the more formal, and archaic, *one*. A teacher, in response, uses *you* in the personal sense of specifically addressing the student. Often teachers will use the inclusive *we* when discussing general principles and rules accepted by the community of mathematics practitioners, thereby instructing, and including, students.

The use of *you* by students in the context of this study indicates the utterance of a general principle or instruction. Students who use the pronoun *I* tend to describe or explain a personal understanding or approach to a problem. The use of the third person pronoun *it* (or *they*) implies vagueness.

Use of personal pronouns by the students interviewed is discussed in three sub-sections. The first analyses how students in each of the three ability groups use the pronouns *I* and *you*. The second sub-section analyses changes in the use of the pronouns *I, you,* and *it* as students talk about items of different relative difficulties. Sets 5 and 6 (linear equations to be solved) provide the context for this analysis. A third sub-section discusses the results.

### The use of the personal pronouns *I* and *you* by ability groups

The pronouns *I* and *you* comprised a small proportion of the total words uttered by students during the interviews (4,378 instances out of a total of approximately 64,542 words). Words or, more often, phrases that students uttered to indicate uncertainty about their correct responses to items in the interview invariably contained the pronoun *I,* e.g., “I haven’t done these for a while” (see following section on Response Types: affective responses). These instances of the use of the pronoun *I* have not been included in the data used in this section. Only data from utterances focused on mathematics have been included.

The numbers of times the pronouns were used by each of the ability groups for each set of items are given in Table 6.5, as averages per student in each ability group. Trends in the use of the pronouns can be seen by reading horizontally for each of the ability groups. The last column gives the overall average frequency for *I* or *you* per ability group.

On average, the High-and Low ability groups used *I* with the same frequency, although the students in the Low-ability group used *you* more often than did students in the High-ability group. The High-ability students used *you* more frequently than *I* when discussing items in Sets 3, 4, and 8, some of the more difficult sets of items. The Low-ability group, on the other hand used *you* more often than *I* when talking about Sets 2, 3, 4, 5 and 7.

*Table 6.5: Average numbers of pronouns (*I *or* you*) per student in each ability group, for each set of items*

| | *I* and *you* by Item Sets | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ability Group | SET 1 | | SET 2 | | SET 3 | | SET 4 | | SET 5 | | SET 6 | | SET 7 | | SET 8 | | SET 9 | | Averages | |
| | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* | *I* | *you* |
| HIGH | 4.9 | 2.8 | 4.7 | 4.0 | 3.2 | 4.3 | 6.6 | 12.2 | 17.5 | 12.0 | 3.8 | 3.8 | 2.9 | 2.1 | 2.8 | 3.2 | 15.7 | 11.7 | 46.4 | 44.4 |
| AVERAGE | 9.6 | 3.1 | 7.2 | 3.5 | 5.4 | 5.1 | 9.3 | 7.5 | 24 | 15.7 | 9.4 | 2.4 | 3.5 | 3.9 | 6.2 | 2.5 | 23.5 | 21.8 | 74.6 | 43.8 |
| LOW | 5.9 | 3.7 | 4.6 | 5.1 | 4.1 | 6.3 | 6.9 | 7.7 | 13.1 | 14.4 | 6.0 | 3.3 | 2.9 | 4.8 | 3.0 | 2.1 | 17.3 | 18.1 | 46.4 | 47.3 |
| Overall Average | 7.3 | 3.2 | 5.1 | 4.0 | 4.4 | 5.1 | 7.8 | 9.1 | 19.5 | 14.2 | 6.8 | 3.1 | 3.2 | 3.5 | 4.4 | 2.6 | 19.6 | 17.7 | | |

When scores for each ability group are aggregated across all sets of items the difference in usage of the pronouns is significant (Table 6.6, $p<0.001$; Appendix D3).

*Table 6.6: Overall comparisons of expected and actual* I *and* you *frequencies between ability groups*

| | Actual | | | Expected | |
|---|---|---|---|---|---|
| Ability group | Total *I* | Total *you* | Total pronouns | *I* | *you* |
| HIGH | 621 | 561 | 1182 | 659 | 523 |
| AVERAGE | 1374 | 918 | 2292 | 1278 | 1014 |
| LOW | 446 | 458 | 904 | 504 | 400 |
| Totals | 2441 | 1937 | 4378 | | |

Students in the Low-ability group used *I* less often than expected, but *you* more often. The students in this group tended to quote various forms of rules or procedures as explanations for their thinking, hence using *you* frequently. However, the success of these students on the survey would suggest that the rules or procedures they remember and quote are mathematically limited.

The Average-ability group used *I* more often than expected, and, consequently, *you* less frequently than expected. The students tended to explain what they, individually, would do in responding to items in each set. The consistency with which these students used *I* suggests a developmental link with algebraic understanding.

The High-ability group used *I* less frequently than expected, but not for all sets of items, and used *you* more frequently than expected (See Table 6.5 and Figure 6.3 below). The lower-than-expected use of *I* by the High-ability group might be attributed to their confidence and security of mathematical knowledge, particularly when these students have tended to use you more frequently when addressing more difficult items.

### *Set-by-set changes in I/You responses*

A set-by-set analysis of pronoun use explores the general patterns described above in further detail. The analysis uses whole number data derived from the averages recorded in Table 6.5, and illustrated in Figure 6.3

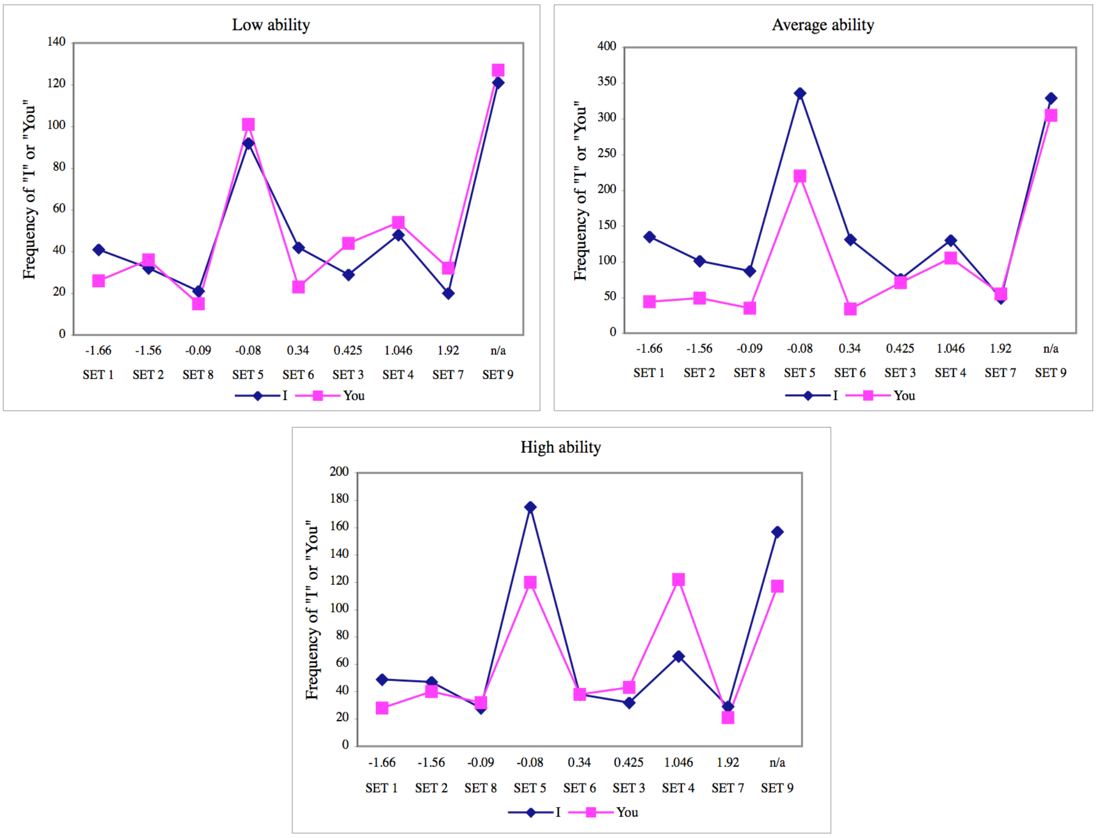


*Figure 6.3: Frequency of* I *and* you *pronouns by each ability group responding to each item set*

Item sets are ordered according to average difficulty. Within the groups, differences between the actual frequencies of *I* and *you* in sets 1 to 8 are significant for Average- and High-ability groups ($p<0.001$, see Appendix D4), and less significant for the Low-ability group ($p=0.02$)

Use of the pronoun *you* implies that the student is stating a generalisation – a rule, procedure, or concept. From the results (Table 6.5 and Figure 6.3), it would seem that the Low-ability group tended to generalise as much, or more than, the High-ability group. However, many of the statements uttered by students in the Low-ability group would appear not to be useful, as demonstrated by the students' lack of success on the survey. These types of statements are illustrated by this exchange from a Low-ability student describing, in general, a rule for dealing with items in Set 3 (brackets):

> S: You do the brackets first, so the brackets stand out and you just go… You do 'em first
>
> I: What do you mean, "You do them first"?
>
> S: Well, you add them first, and then you go what the number is inside, then you times it by the outside. [6]

It is notable that High-ability students used *you* to a far greater extent in responding to items in Set 4 (Fractions), and to a lesser extent in Set 6 (more difficult equations). In these instances, it can be inferred, from their success on the survey, that the High-ability students were able to make generalisations that are mathematically useful, such as the following response to the first two items in Set 4 [Item 3, $a/5 + a/10$; and Item 15, $3p/4 - p/8$]:

> S: … You've got to make the bottom, the denominator, you've got to find common denominators and that's… [long pause].
>
> I: Why would that be?
>
> S: Because, if you are adding one third and one half… If you are not going to use decimals, you've got to have the same… You've got to have a common denominator to add them together.

Despite the pauses and hesitations that imply a degree of uncertainty (possibly because the student is thinking in what is, to him, a problematic situation) the student expresses a developing understanding of the mathematics. From his survey result, he was able to apply the general rule quoted in this exchange.

Set 4 and Set 6 were two of the more difficult sets (see Table 6.2 for details). Students who were less successful in the survey avoided giving any answers, such as this response from a student in the Low- ability group to items in Set 6:

> Well 33 [Item 33, $(x + 3)/2 = 7$] just muddles me. I don't get that. The same with 36 [Item 36, $63/x = 180$].

---

[6] Note the vague pronouns, *it* and *them* – discussed in the following sub-section.

These students also struggled to explain their thinking, except in personal terms, such as this response to items in Set 4, where the use of the pronoun *I* is frequent:

> When I see the fractional symbol [Item 3, $a/5 + a/10$], I think divided by, like division. So, like, I would think… Because to work out an answer you have to multiply it again to get, like, make it not a division, you have to multiply it. Just do the opposite. So I'd, first one here, I'd times by five, and then times the *a* by five.

The use of *I* indicates insecurity – a lack of confidence in mathematical knowledge, and/or an associated lack of commitment to the statement made. All ability groups tended to use *I* more frequently when responding to items in Set 1. This set was, on average, the easiest set, and therefore might not be expected to have students respond using the pronoun *I* as often as they did. It was also the first set of items to be asked of the students. The presence of a higher- than-expected use of the pronoun might be more the result of students' nervousness or unfamiliarity with the situation than their insecurity of knowledge.

In the case of the High-ability students, the frequent use of *I* in responses to Set 5 reflects the uncertainty in solving some of the more difficult equations by students in the High-7 group. The two following examples are excerpts from responses by students attempting to answer Item 37 [$5a - 4 = 2a + 8$] and Item 34 [$x + (x + 2) = (x - 1) + 8$].

> I: What about [item] 37, or 34?
>
> S1: You'd … I don't know. I have no idea.
>
> I: What would you do to solve that equation?
>
> S1: I wouldn't solve it, because I can't understand how it is done.

and

> S2: […] All right, when I see Question 37… Well, I see it as four, minus…Well, four minus 5*a* will be the equivalent of two plus, 2*a* plus eight. I probably need to write it down but, I um, …

The second student, confessed, after a long struggle that, "Yeah, I did it by guess-and-check I think. I don't know how to work them out, properly". Thus the student indicated the source of the conceptual struggle in which he engaged. (Responses to equations in Sets 5 and 6 are further analysed in detail in the following section, and further highlight changes in pronoun use as students encounter items of greater, or less challenge.)

The use of *I* dominates responses from students in the Average-ability group, except in Sets 3 and 7. Both these sets consisted of items with brackets. The use of *I* by the students in the Average-ability group when responding to the other sets, however, may be taken as indicative of their insecure understanding. Set 6 (equations) was approached with confidence (see also in section Verbosity), but the difficulty estimates of the items (Item 30 [$x/4 = 12$], -0.93 logits; Item 33 [$(x+3)/2 = 7$], -0.58 logits; Item 36 [If $63/x = 180$, what is $x$?], 2.64 logits) in the set were near or beyond the ability range of the group (-0.94 to 0.93 logits).

Students were also not successful when responding to items in Set 8 in the survey. The insecurity of understanding is indicated by a more frequent use of *I* than the use of *you* (Figure 6.3). As noted elsewhere, Low-ability students tended to respond to the items in Set 8 by restating the question. High-ability students reworded the items as they might be rewritten in simplified terms.

### *The occurrence of* **I** *or* **you** *in Set 9*

Set 9 consisted of questions designed to provide background information on linguistic patterns when students were asked questions of a non-algebraic nature, or about non-mathematical topics (Appendix B2 for the interview protocol). Changes in the frequencies of *I* and *you* as students responded to questions in Set 9 could indicate a change in their confidence, as they no longer had to address algebraic ideas. The data in Table 6.5 and illustrated in Figure 6.3 illustrate changes in the use of *I* and *you* by all groups. All groups used these pronouns more often than in other sets (Sets 1-8), with the exception of Set 5. This is due to their providing longer responses to Set 9 items than for many of the other sets of items.

Average- and Low-ability students appeared to use *I* and *you* with approximately equal frequencies. The High-ability group tended to use *I* more often than *you*. The changes in proportions of the pronoun *I* with respect to the total numbers of pronouns used by each group in responses to Sets 1 to 8, all item sets (1 to 9) and Set 9 only, demonstrate the extent of the shift in pronoun use by each ability group. These data are recorded in Table 6.7

*Table 6.7: Proportion of the pronoun* I *used in comparison with all* I/you *pronouns for Sets 1 to 8, all sets and Set 9, by ability groups*

| Ability Group | Proportion of I / all pronouns | | |
|---|---|---|---|
| | Sets 1-8 | Sets 1-9 | Set 9 |
| HIGH | 0.5 | O.53 | 0.57 |
| AVERAGE | 0.63 | 0.6 | 0.52 |
| LOW | 0.5 | 0.49 | 0.49 |

The results indicate that the students in the Average-ability group were more able to generalise about topics in Set 9, than they were about algebraic concepts in Sets 1 to 8. On the other hand, the High-ability group of students was more inclined to answer background questions from a personal perspective than that which they adopted when speaking about algebra items.

The lack of a significant change in the use of the pronouns *I* or *you* by the Low-ability group of students suggests that these students might have learnt to use the impersonal *you* when speaking to a teacher, regardless of the topic for discussion. Alternatively, for the Low-ability students, the background questions appeared to be sufficiently similar in concept to those in Sets 1 to 8 not to prompt any change in the linguistic features of their conversation.

Many students answered the questions in Set 9 quite impersonally, avoiding any use of pronouns. This might imply an understood impersonal *you* or vague *it*, but these occurrences were not counted, because they were not explicit.

### The use of personal pronouns with respect to sets of items

The previous sub-section examined trends in the use of personal pronouns as students responded to each item set. However, there were many instances where students either declined to make any mathematical response to items, or addressed a single item only. Therefore, examination of responses by students to individual items could provide clearer patterns of pronoun use in response to conceptual challenges as items became more difficult.

The use of the pronouns *I*, *you* and *it* was examined in one particular context – as students responded to items in Sets 5 and 6. These sets consisted of equations to be solved (Table 6.2), and were the sets that produced the longest conversations, particularly Set 5. This was due, in part, to there being more items to discuss in Set 5, than there were in any of the other sets. Conversations were also longer because the students, on the whole, appeared to be most comfortable with these types of items.

Analysis of the frequency of the three personal pronouns in the responses by each of the groups to particular items gave a clearer pattern of changes in linguistic behaviour as students addressed items of different difficulties. The two sets of equations were chosen because the students' responses tended to be directed to the individual items, and most students attempted to respond to all items in the two sets.

The average difficulty of Set 5 (-0.09 logits) meant that it was the third easiest set, and near the mean difficulty of all items in the algebra survey (0.00 logits). The average difficulty estimate also approximated the ability estimate of the Average-ability group (-0.01 logits). Set 6 was one of the more difficult sets (average difficulty estimate: 0.34 logits), but also consisted of equations. The two sets provided a range of data that allowed the examination of the use of personal pronouns.

The students' responses are analysed in the first part with respect to the difficulty of the items. A second part focuses on responses of the ability groups to individual items in Sets 5 and 6.

### ***Personal pronouns and changes in item difficulty***

Table 6.8 details the frequency of occurrence of the pronouns *I*, *you* or *it* in responses by students (*n*=31) to items in Sets 5 and 6. Each item is identified by the survey number, together with the Rasch-difficulty estimate in logits, the number of correct responses to the survey item by the students interviewed, and the number of refusals to respond to the interview question. Each response by a particular student

*Table 6.8: Frequency of statements using predominantly the pronouns* I, it *or* you *by students responding to items in Set 5 and Set 6 of the interview ordered by item difficulty.*

| | Items in Set 5 and Set 6 in order of Rasch difficulty | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Item number | **27** | **28** | **30** | **33** | **29** | **32** | **31** | **37** | **34** | **36** | **39** |
| Rasch difficulty | *-3.1* | *-2.3* | *-0.9* | *-0.6* | *-0.5* | *-0.3* | *-0.1* | *1.22* | *1.51* | *2.64* | *3.09* |
| | Responses by all ability groups | | | | | | | | | | |
| No. Correct responses | 29 | 23 | 21 | 20 | 17 | 19 | 3 | 11 | 8 | 5 | 1 |
| No reply | 0 | 1 | 1 | 6 | 4 | 3 | 19 | 9 | 12 | 11 | 5 |
| *I* | 13 | 13 | 14 | 12 | 12 | 9 | 6 | 14 | 13 | 9 | 8 |
| *it* | 3 | 2 | 11 | 6 | 4 | 10 | 2 | 2 | 3 | 7 | 1 |
| *you* | 15 | 15 | 5 | 7 | 11 | 9 | 4 | 6 | 3 | 4 | 7 |

to a particular item was coded according to the most frequently occurring pronoun in that response. Thus, a response might be an "*I*-response", a "*you*-response", or an "*it*-response". (E.g., Item 31, difficulty estimate of 1.22 logits, attracted three correct responses in the survey. During the interviews, 19 students declined to answer, or to attend to the item; six students used *I* as the predominant, or only, personal pronoun in their responses; two most frequently used *it*; and, four students tended to use *you* most frequently in their responses to this item.)

The number of responses tended to decrease as the items became more difficult. Item 31 [$4(p+3) = 32$] was often avoided during the interview, unless the students were asked specifically about the item. On the basis of their earlier responses to the set, many of the less-able students were not asked to respond to Item 39 [$ax = 5$] and, hence, fewer responses of any type were recorded. This item, although one of the most difficult in the survey, attracted confident answers from the Average- and High-ability students who, in most cases, suggested possible values for $a$ and/or $x$ as being factors of five, or 2.5. These details are contained in the Error Table 5.4 in Chapter 5.

The occurrence of the pronoun *you* suggests that the student is, in some ways, making a general, propositional statement. The number of these second-person responses tended to decrease as items became more difficult. Item 29 [$2t - 23 = 49$] was the exception.

The use of *I* suggests a degree of uncertainty, the student merely stating what he or she would do in a particular circumstance. The occurrence of the pronoun *it* indicates vagueness – a position less assured even than that of a student who provides a personal opinion about a specific item by the use of *I*, rather than a general rule.

Items that attracted the greatest number of statements in the first or third person were those that required careful reading, e.g., Item 30 [$x/4 = 12$], or involved challenging arithmetic, e.g., Item 32 [$10y = 5$], and Item 36 [$63/x = 180$]. The two easiest items, Item 27 [$x + 5 = 7$] and Item 28 [$4y = 20$], were usually responded to in the first or second person.

Figure 6.4 and Table 6.8 show three peaks for the occurrence of the pronoun it. These peaks occur at difficulty estimates that are near the thresholds for each ability group. The first peak occurs at Item 30 with a difficulty estimate of -0.9 logits. This is at a difficulty level above the upper threshold of the Low-ability group (-1.28 logits) and slightly above the lower threshold for the Average-ability group (-0.94 logits). The second peak occurs at Item 32 with a difficulty estimate of -0.3 logits, a little lower than the mean ability estimate for the Average-ability group (-0.01 logits), but which is greater than the ability estimates for half of the Average-ability group. The third peak occurs at Item 36 with a difficulty of 2.64 logits, which is greater by 0.9 logits than the upper ability estimate for the High-7 ability group (1.65 logits) and greater than the lower ability estimate for the High-3 ability group (2.58 logits).

These figures suggest that when students meet items that are beyond their ability level, their statements are likely to be couched in vague terms, indicative of a struggle to make understanding explicit. If the sum of the numbers of first person and third person statements is considered (implying specificity or vagueness), the ability of students to make general statements, inferred from their use of the second person, decreases with the difficulty of the items. Vagueness of statements indicates students are dealing with ideas of which they have little understanding.

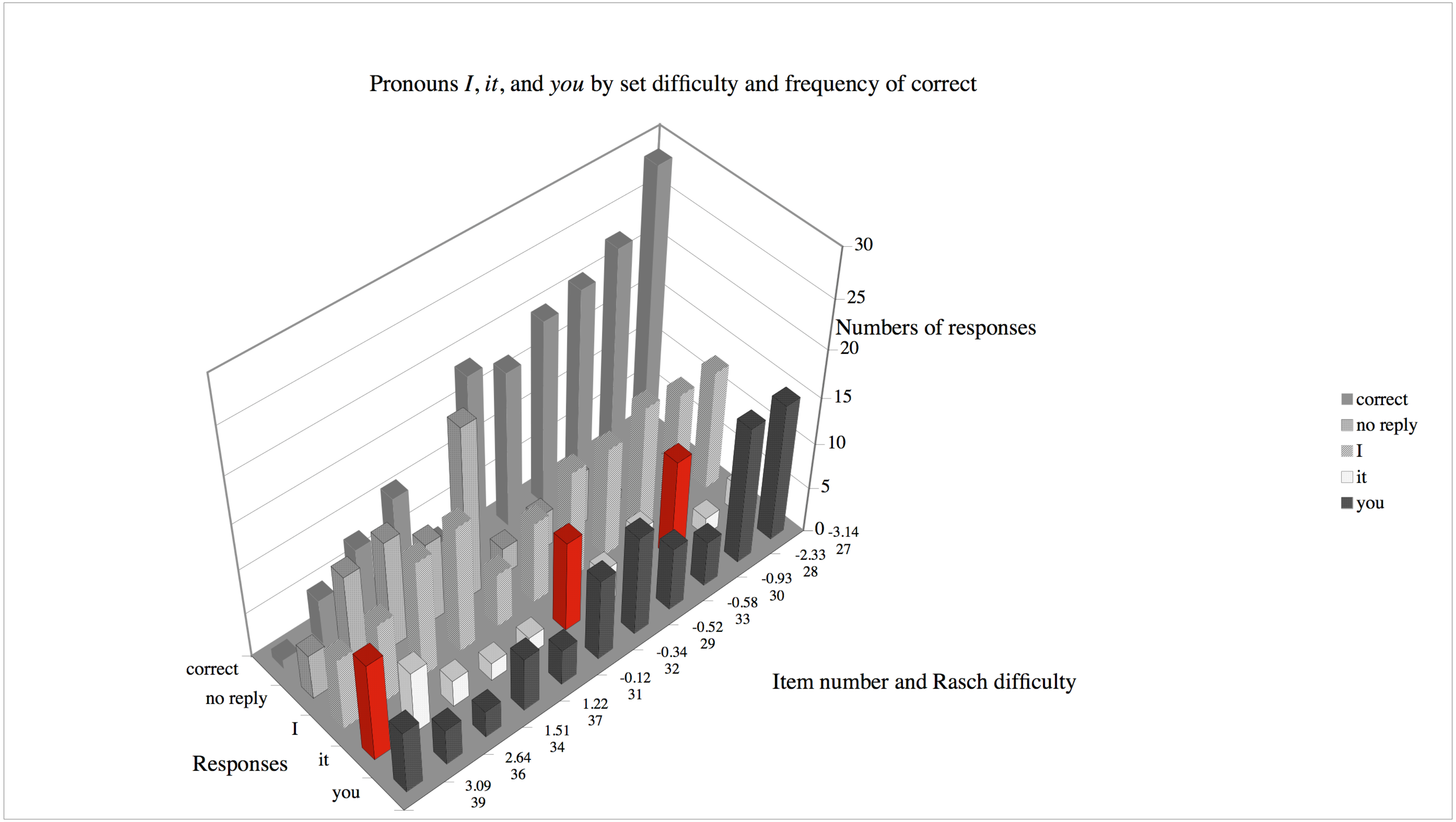


*Figure 6.4: Comparison of the frequency of occurrence of the pronouns* I, it *and* you *with the number of correct responses by the interviewed students to items in Set 5 and Set 6, arranged in descending order of difficulty.*

### *Personal pronouns and ability group responses to changes in item difficulty*

The use of the pronouns *I, you,* and *it* was also analysed with respect to the ability groups for each item in Sets 5 and 6. For this analysis, the High-ability group was split into two sub-groups delineated by a gap of 0.9 logits between the three highest ability estimates (High-3) and the remaining seven (High-7) (Table 6.1).

Table 6.9 outlines the frequency of statements consisting of the three personal pronouns by each of the ability groups in their responses to items in Sets 5 and 6. The frequency is expressed as a proportion of responses rather than as raw counts.

The proportions of responses have been calculated as the average number of statements using a particular pronoun (*I*-responses, *You*-responses, *it*-responses) for each student in the group (or sub-group). (E.g., in the Average-ability group, five students out of the fourteen in the group used *I* statements when responding to Item 27, and nine students out of the fourteen responded with *you* statements to that item.) The proportions are expressed as decimals to one decimal place.

Shaded areas in Table 6.9 indicate where students in the ability groups (or sub-groups) have a less than, or equal, chance of obtaining correct answers to the items. The Average-ability group has a wide spread of ability (1.8 logits, either side of the mean item difficulty estimate of 0.0 logits). An area where some students in the group could be expected to struggle, whilst others do not, is indicated by diagonal hatching in columns under Items 33, 29 and 32. There is also a similar transitional area for the Low-ability group at Item 30.

Of the High-3 sub-group, one student persistently used *I*. This seemed to be an individual idiosyncrasy of expression. The others used *you* almost exclusively. One only student in the sub-group avoided responding to Item 34 [$x + (x + 2) = (x - 1) + 8$], and none of the students used the third person pronoun *it*.

*Table 6.9: Proportion of* I, you *or* it *statements by ability groups in response to items in Set 5 and Set 6, items ranged in ascending Rasch order, ability in descending order*

| | | Items in Set 5 and Set 6 in order of Rasch difficulty | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ability groups [Range in logits] | Response | **27** | **28** | **30** | **33** | **29** | **32** | **37** | **34** | **36** | **39** |
| | | *-3.14* | *-2.33* | *-0.93* | *-0.58* | *-0.52* | *-0.34* | *1.22* | *1.51* | *2.64* | *3.09* |
| HIGH-3<br>[2.58 to 3.67] | Nil | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.3 | 0 | 0 |
| | *I* | 0.3 | 0.3 | 0.3 | 0.3 | 0 | 0.3 | 0.3 | 0.3 | 0.3 | 0 |
| | *it* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | *you* | 0.7 | 0.7 | 0.7 | 0.7 | 1 | 0.7 | 0.7 | 0.3 | 0.7 | 1 |
| HIGH-7<br>[1.27 to 1.65] | Nil | 0 | 0 | 0 | 0.1 | 0.1 | 0 | 0.0 | 0.3 | 0.3 | 0 |
| | *I* | 0.6 | 0.1 | 0.7 | 0.6 | 0.6 | 0.6 | 0.7 | 0.7 | 0.4 | 0.3 |
| | *it* | 0 | 0 | 0 | 0 | 0.1 | 0.1 | 0 | 0 | 0.3 | 0.4 |
| | *you* | 0.4 | 0.9 | 0.3 | 0.3 | 0.1 | 0.3 | 0.1 | 0 | 0 | 0.3 |
| AVERAGE<br>[-0.94 to 0.93] | Nil | 0 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.3 | 0.3 | 0.2 |
| | *I* | 0.4 | 0.5 | 0.4 | 0.4 | 0.4 | 0.1 | 0.5 | 0.5 | 0.4 | 0.3 |
| | *it* | 0 | 0.1 | 0.1 | 0.4 | 0.1 | 0.4 | 0.1 | 0.1 | 0.3 | 0.5 |
| | *you* | 0.6 | 0.3 | 0.1 | 0.1 | 0.4 | 0.3 | 0.1 | 0.1 | 0.1 | 0.1 |
| LOW<br>[-2.86 to -1.28] | Nil | 0 | 0 | 0 | 0.4 | 0.1 | 0.1 | 0.6 | 0.7 | 0.9 | 0.6 |
| | *I* | 0.4 | 0.6 | 0.4 | 0.3 | 0.4 | 0.3 | 0.1 | 0 | 0 | 0.3 |
| | *it* | 0.4 | 0 | 0.6 | 0 | 0.1 | 0.4 | 0.1 | 0.3 | 0.1 | 0.1 |
| | *you* | 0.1 | 0.4 | 0 | 0.3 | 0.3 | 0.1 | 0.1 | 0 | 0 | 0 |

Students in the High-7 sub-group tended to use *I* more frequently then *you*, and more frequently than students in the other ability groups. However, they did not tend to use *it*, until responding to the two most difficult items. These distinctively different linguistic behaviours reflect the gap in ability estimates of 0.93 logits between the sub-groups of the High-ability group. A similar gap (0.53 logits) also occurs in the between-the-item difficulty estimates on the Rasch model, between Item 3 [$a/5 + a/10$] and Item 35 [$x + x/3 = 4$] (Chapter 4, Figure 4.4).

The Average-ability group tended to use *I* and/or *it* most frequently, more frequently than *you*, except for Item 27 [If $x + 5 = 7$, then $x = ?$]. *It* began to be used more frequently as the items became more difficult.

The Low-ability group also tended to use *I* and *it* more frequently than *you*, particularly when responding to items within their ability (the least difficult equations, Items 27, 28 and 30). This group also gave more "nil" responses the more difficult the items became.

## Discussion

The use of different personal pronouns can indicate the degree to which students are secure in their mathematical understanding. When ability groups were analysed for the use of the pronouns *I* or *you,* the differences were most marked when the most difficult sets of items were considered.

Students in the Average-ability group tended to use *I* more often than expected. This could be associated with these students attempting to answer each item in a set individually, rather than addressing them as a group of items sharing common characteristics. These students appear to be unable to make any general propositional statements, implied by use of the impersonal *you*, about items in a set. This pattern can be seen clearly when responses by the Average-ability group to items in Sets 5 and 6 are analysed. Despite the difficulty estimates of some items being below the threshold of ability estimates for the group, *I* was used approximately 50% of the time. This frequency did not change markedly. Its use suggests that students in this group have a "singular " perspective, reflected by the tendency of students in the group to focus on one item at a time in each set.

Students in the High-ability group tended to use *I* less frequently, and to use *you* most often. This seems to be an indication of their security of knowledge. The sub-group High-3, encountered few items with which they would be insecure (i.e., the degree of difficulty of the items was either less than or approximately equal to the ability estimates of the students). None of these students was vague (as indicated by the use of *it)*, and two of the three students consistently used *you*. This was a small group, and therefore the use of *I* consistently by the third student gives greater weight to the occurrence of this pronoun than might occur with a larger sample.

The High-7 sub-group appeared to make a sharply defined shift from the use of *you* to the use of *I* when the difficulty estimates of items neared the threshold for this group (1.22 logits). Students in this sub-group also began to use *it* when dealing with items 0.8 logits above the sub-group's ability threshold. Such a shift in frequency of use suggests a connection between the ability of students to respond to items on the basis of general or principled understanding and their use of *you*. On the other hand, the use of *I* signalled students' insecurity of knowledge as items posed greater conceptual challenges (i.e., item difficulty estimates were higher than students' ability estimates).

When students responded to items that were of a difficulty estimate lower than their ability, they tended to use the pronoun *you*. When item difficulty approached or became greater than the students' ability estimates, the use of *I* and/or *it* became more frequent. The use of *I* and *it* occurred most frequently with the students in the Low-ability group. In particular, the use of *it* was often in association with general, rule-like statements. At one level of analysis these statements bore close resemblance to the statements made by the students in the High-ability group. The use of the pronoun *it*, instead of the general *you*, indicates that these statements were vague utterances rather than statements of firmly understood general principles.

The change in frequency of use of personal pronouns by students indicated a change in their security of knowledge. A shift from the general *you* to the more personal *I,* to the use of the vague *it,* occurred as items became more difficult. These shifts occurred earlier with the Low-ability group, and later with the High-ability group, reflecting the capacity of students in the groups to deal with items of increasing difficulty.

## RESPONSE TYPE

This section discusses the analysis of types of responses made by students in their interviews to items in Sets 1 to 7. The statements were classified as *affective*, *specific* or *general*.

Affective statements were those that expressed certain feelings about the items in the set, such as the negative statement "To be perfectly frank, I am no good at fractions", or a positive statement such as "These are easy". Specific statements were of two types: the first, where a student responded to one item only in the set, usually by providing an answer; the second, where a student responded to all items in a set, individually providing answers, but not making explicit any relationships between the items. General statements were of three types: the first, a description of the items (e.g., “They all have brackets”; “They all have letters”); the second, a statement of a rule or generalised procedure – not necessarily correct, or useful, (e.g., “…do it first”); and the third, a propositional statement that attempted to explain the mathematics (e.g., “[…because] you add the like terms”).

Data are analysed with respect to the ability groups and their responses to the algebra items in Sets 1 to 7. The frequencies of these responses are then compared to the frequencies of occurrence of the different types of responses given to items in Set 9. (Set 8 has been excluded because of its structure, which precluded any extensive free responses.)

Response types are discussed using examples from Sets 1 to 7 in the first two sub-sections, examining firstly, how responses change as the item difficulties change, and, secondly with respect to the ability groups. A third sub-section analyses response types by ability groups to items in Set 9. Finally, implications of these analyses are discussed.

### Response types and set difficulty

For this analysis, responses by the participants were aggregated and analysed with respect to the difficulty of the sets. Table 6.10 summarises the numbers of the different types of responses for each set of items. The item sets are ordered according to difficulty with the exception of the sub-sets of Sets 5 and 4. These sets were presented to the students as two parts.

Sub-set 5a comprised simple linear equations, and the second sub-set, 5b, linear equations of the form $ax = c$. Sub-set 5a had an average difficulty less than that for sub-set 5b, and less than the overall average difficulty for Set 5.

Sub-set 4a consisted of two expressions where algebraic fractions needed to be added or subtracted; sub-set 4b consisted of algebraic fractions to be simplified by multiplication and division (Table 6.2). Sub-set 4b had a difficulty estimate less than that for sub-set 4a. In Table 6.10, these sets are listed firstly with the average difficulty for the whole set, and then the sub-sets in alpha-numeric order (5a, 5b; 4a, 4b).

*Table 6.10: Numbers of affective, specific, and general responses to items in Sets 1 to 7*

| | | TYPE OF RESPONSE | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sets | Set difficulty | Affective | Specific | | General | | | Total responses | Total maths responses |
| | | | One only | All/most | Descriptive | "Rule" | Explanation | | |
| 1 | -1.66 | 13 | 7 | 11 | 4 | 16 | 10 | 61 | 48 |
| 2 | -1.56 | 14 | 17 | 11 | 8 | 5 | 9 | 64 | 50 |
| 5 | -0.09 | 26 | 20 | 30 | 14 | 11 | 18 | 119 | 93 |
| *5a* | *-0.23* | *16* | *13* | *20* | *6* | *8* | *7* | *70* | *54* |
| *5b* | *0.14* | *10* | *7* | *10* | *8* | *3* | *11* | *49* | *39* |
| 6 | 0.34 | 8 | 25 | 10 | 3 | 0 | 3 | 49 | 41 |
| 3 | 0.425 | 10 | 10 | 7 | 1 | 12 | 18 | 58 | 48 |
| 4 | 1.046 | 17 | 36 | 13 | 10 | 7 | 11 | 94 | 77 |
| *4a* | *1.385* | *11* | *17* | *4* | *8* | *5* | *9* | *54* | *43* |
| *4b* | *0.82* | *6* | *19* | *9* | *2* | *2* | *2* | *40* | *34* |
| 7 | 1.92 | 5 | 12 | 14 | 3 | 5 | 5 | 44 | 39 |
| | Totals | 93 | 127 | 96 | 43 | 56 | 74 | 489 | 396 |

Out of all responses, 19% expressed a student's attitude (affective responses); 46% focused on individual items (specific); and, 35% were generalisations. More than half of the mathematically focused responses were specific (56%). The following part discusses the affective responses. In a second part, specific and general responses are considered.

### *Affective responses*

Affective responses were those where students expressed either positive or negative attitudes to the items or sets of items. The uttering of affective statements that expressed either positive or negative attitudes often served to deflect the conversation from the mathematics to be discussed. Negative statements, such as "I hate fractions", served both as an apology for, and as a warning of possible errors the student might make. Positive statements, such as "These are easy", often substituted for more mathematically focused explanations. These statements comprised approximately one fifth of all types of responses, and appeared to occur with greater frequency as students responded to item sets of lower difficulty estimates. The first two sets of items were of the lowest average difficulties, and the students' tendency to make affective remarks might indicate an initial uncertainty about the interview situation itself, rather than the mathematics they were asked to explain.

However, Set 5, consisting of the greatest number of items, and items with which students appeared to be more comfortable, provoked the most conversations (See Verbosity, Figure 6.2) and the greatest number of affective responses. The more difficult sets of items provided less opportunity for students to talk and, hence, make attitudinal comments.

The one exception was sub-set 4a, which consisted of two expressions requiring the addition or subtraction of fractions. This group of items was the second most difficult group (1.385 logits), and provoked almost twice as many affective statements as the items in sub-set 4b (multiplication of fractions), with a difficulty of 0.82 logits. Set 4 also resulted in a considerable amount of student talk, second to that for Set 5.

Set 7 although on average the most difficult set of items, because of the inclusion of Item 10 $(x + y)^2$, provoked few affective responses. This could be attributed to students feeling confident in dealing with expressions with brackets, although not necessarily correctly.

### *Specific and general responses*

The frequency of occurrence of specific and general responses as the difficulty of item sets changed was also analysed. The numbers of these different responses are summarised in Table 6.9, above. Because comparisons are made between frequencies of occurrences in several contexts the two types of responses are discussed together in this part.

Figure 6.5 illustrates the trends in the occurrences of specific and general responses for Sets 1 to 7. Overall, more specific responses were offered than general responses. The difference between the actual and expected frequencies of response type was significant ($p<0.001$, Appendix D5). However, the greatest difference between the actual and expected responses was with Sets 1, 3 and 6. The expected frequencies are shown in Figure 6.5 by markers.

Only in Set 1 and Set 3 were there more general responses than specific responses. Set 1 consisted of expressions where terms were to be added or subtracted; Set 3 consisted of expression with brackets. Set 6 (equations of the form $ax = c$) attracted the least number of general responses.

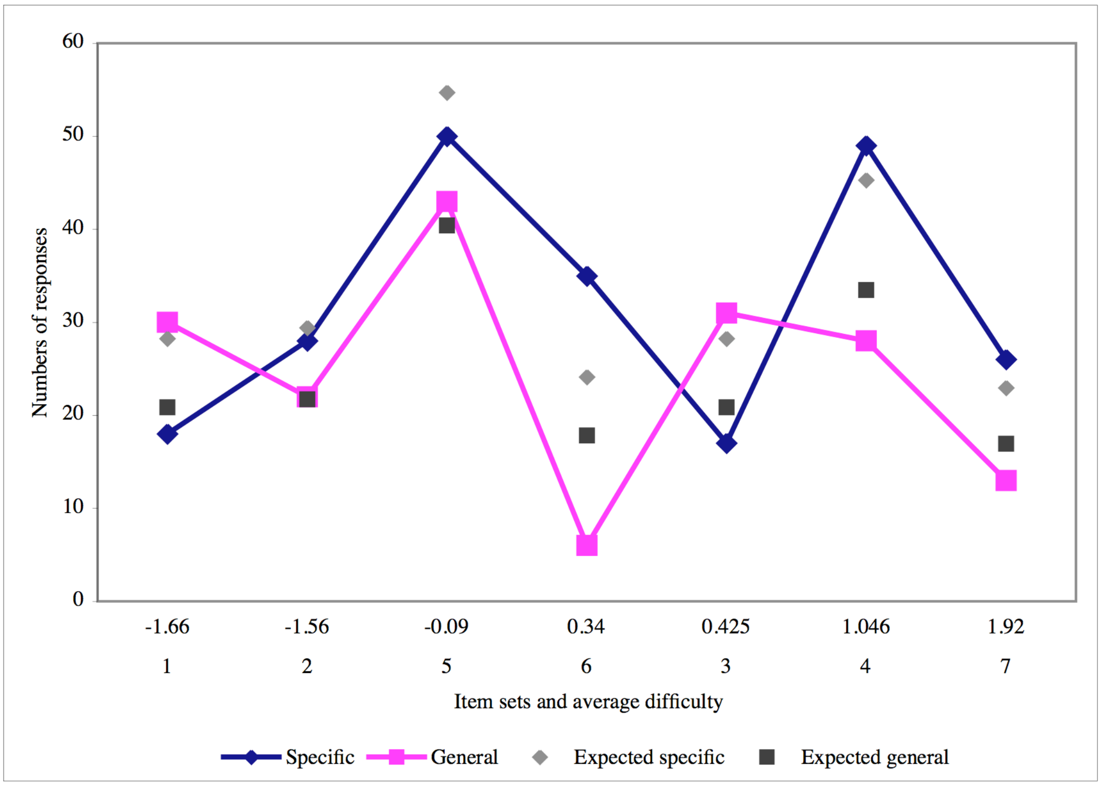


*Figure 6.5: Actual and expected frequencies of specific and general responses to Sets 1 to 7.*

When specific and general responses are examined for frequencies of sub-types, other patterns emerge. Figure 6.6 illustrates the patterns for specific responses that address either a single item in a set, or multiple items in a set (see also Table 6.10). The trend appears to be that the number of responses to single items increases as the sets increase in average difficulty. These more difficult sets, Sets 6, 3, 4 and 7 contained a variety of items, some having a much greater difficulty estimate than others (e.g., Set 6 difficulty estimates ranged from -1.58 to 1.64 logits; Set 7 from -0.54 to 2.56 logits). Students often chose to respond to one of the less difficult items only.

Set 1, Set 5, and Set 7 attracted more responses to multiple items than any of the other item sets. Set 5 also consisted of the greatest number of items (eight in total), all of which were simple linear equations to be solved. Set 7 consisted of four expressions with brackets, including two using index notation.

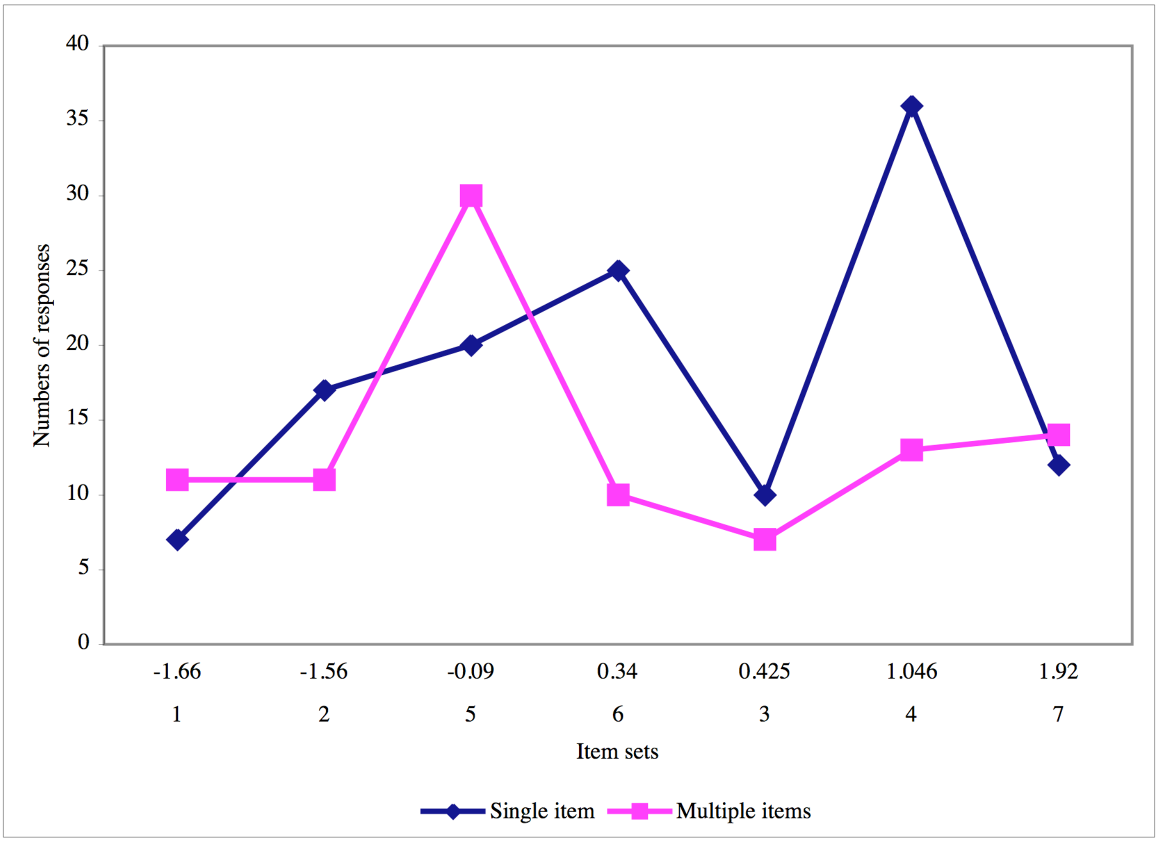


*Figure 6.6: Numbers of specific responses to single and multiple items in Sets 1 to 7*

Set 4 consisted of four algebraic fractions to be manipulated. Many students chose to answer Item 8 [$4ab/4a$] only, avoiding giving responses to the addition or subtraction items. Set 6 (equations), as shown in Figure 6.5, attracted few general responses, but many responses to a single item, usually Item 33 [$(x + 3)/2 = 7$, difficulty estimate: -1.58 logits]. Two of the items (Items 33 and 36) were those which attracted many vague (*it*) responses. Many students answering Item 33 only, used the pronoun *I* (see previous section, Table 6.8). Interview responses suggest that students solved this equation in a series of closed steps.

General responses were those that did not identify any one specific item. Most general responses tended to be explanatory, although the explanations were always procedural, describing how an answer was to be arrived at, but lacking any conceptual rationale for procedures. Set 5 (equations) and Set 3 attracted the greatest number of these general responses. Students tended to solve specific equations, in Set 5, but also explain, by describing in more general terms, steps to a solution.

> S1: Well what I try to do is rewrite it so that it is just numbers, because it makes it easier for me.

> I: Can you explain further?
>
> S1: Well, a lot of these, it is a lot easier if you put it in numbers. Well for me… Then it's just a matter of reversing a lot of these. If $x$ plus five equals seven, then seven minus five equals $x$.

The student has tried to explain his solution of the equation $x + 5 = 7$. This statement is considered to be general in nature, because of the use of *you* in the description of a procedure. However, it is also vague, as indicated by the frequent use of *it*. Uncertainty is indicated by phrases, "…what I try to do…", and "Well, for me".

Similarly, students would explain what they would do when simplifying expressions with brackets, as this student does when answering Item 19 [$2(x+5) + 8$] from Set 3.

> S2: You have to times everything, what's out of the bracket by what's in the bracket. So $2x$ plus ten and then $2x$ plus eight…
>
> I: Can you tell me a bit more about that?
>
> S2: No, because it's a minus in between there, and you just times what's on the outside by what's on the inside. Yeah.

The citing of a general rule was the least common general statement, except for Set 1, Set 3 and Set 7. For Set 1, students usually stated, "You add (put) like terms together". This statement, particularly the use of the word *put*, was manifest in various ways. Students in the Low-ability group tended to conjoin the terms so that the expression $3m + 8 + 2m - 5$ became $8m$ or $5m3$ (See Chapter 5, Error Analysis). Although the rule sounded more or less correct, even if informal, it did not necessarily translate into correct mathematics. This was apparent also for the Sets 3 and 7 that consisted of expressions with brackets. Brackets were described as meaning "you do it first". How this rule was applied has been discussed elsewhere.

Descriptive general responses tended to be more common than the stating of a rule, but less common than explanations, except for Sets 1, 6, 3 and 7. The descriptions, however, provided no mathematical insights, focusing only on surface features of items. In some instances descriptive statements were elaborated by a student stating a rule such as those described above.

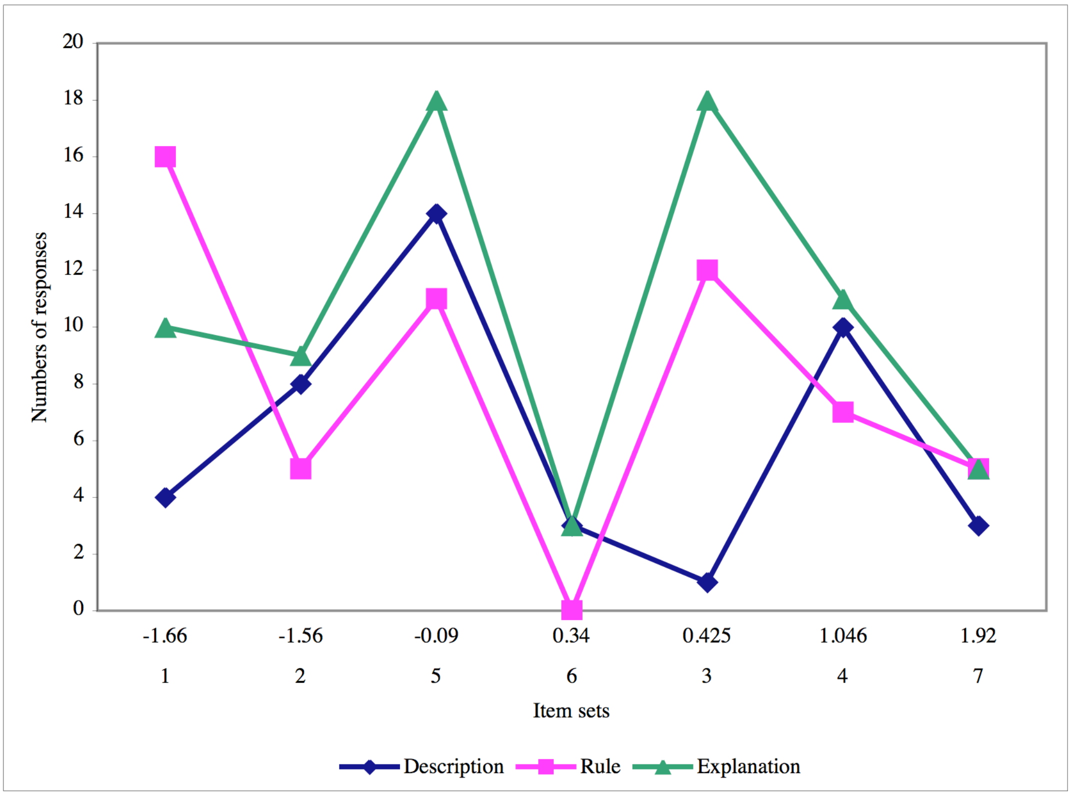


*Figure 6.7: Numbers of general responses (descriptions, rules and explanations) to Sets 1 to 7.*

The different types of general responses by students are illustrated in Figure 6.7. Overall there were fewer general responses, a little more than one third of all responses (Table 6.10, Figure 6.5). Most general responses were given for Sets 1, 5 and 3. Only for Set 1 did the number of general responses (26, not including descriptions) exceed that for specific responses (16). Responses to Set 5 consisted of 30 specific responses, and 20 general responses that were either rules or explanations; to Set 3 there were 49 specific responses and 18 general responses. Explanations were usually given by students in the High-ability group, descriptions by students in the Low-ability group (see following section for further discussion of responses by ability groups).

Throughout the interviews, students tended to respond by giving answers to specific items in each set, rather than articulating a generalisation based on common mathematical features of the items. Sets of less difficult items provided some students with the opportunity to address multiple items, but, as set difficulty increased, students tended to select individual items on which to base their responses. The number and type of items in each set influenced the patterns of responses, particularly whether a student would focus on one item, or multiple items. Descriptive generalisations appeared to be based on identification of common surface features of items in a set. Rule-like generalisations tended to be vague, explanations procedural.

By aggregating responses of each of the groups, patterns of changes in response types as set difficulty increased, showed no clear trends. When the responses by the individual ability groups were analysed, clearer patterns emerged.

### Response types by ability groups in Sets 1-7

Different ability groups tended to respond to items in the interviews in significantly different ways to that which could be expected ($p<0.001$, Appendix D6). Patterns of responses were most noticeable for the Average-ability and High-ability groups. In some instances, the response patterns of both the High- and Low-ability groups were similar. The numbers of different responses by each ability group, aggregated across the item sets, are summarised in Table 6. 11.

*Table 6.11: Numbers of response types by ability groups*

| | Response type | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Specific | | | General | | | |
| Ability group | Affective | One item | Multiple items | Description | Rule | Explanation | Total responses | Total math responses |
| HIGH | 12 | 25 | 35 | 7 | 26 | 39 | 144 | 132 |
| AVERAGE | 55 | 72 | 47 | 18 | 14 | 24 | 230 | 175 |
| LOW | 26 | 30 | 14 | 18 | 16 | 11 | 115 | 89 |
| Totals | 93 | 127 | 96 | 43 | 56 | 74 | 489 | 396 |

Numbers of students in each of the ability groups were different. The Average-ability group consisted of the greatest number of students (14), while the Low-ability group consisted of half that number; the High-ability group consisted of ten students. To illustrate the trends in response patterns the average numbers of response-types was found for each of the groups. These are summarised in Table 6.12.

*Table 6.12: Average number of responses by type per student for each ability group*

| | Response Type | | | | | |
|---|---|---|---|---|---|---|
| | Affective | Specific | | | General | |
| Ability Group | | One item only | All/ most items | Description | Rule | Explanatory |
| HIGH | 1.2 | 2.5 | 3.5 | 0.7 | 2.6 | 3.9 |
| AVERAGE | 3.93 | 5.14 | 3.36 | 1.29 | 1.00 | 1.71 |
| LOW | 3.71 | 4.29 | 2.00 | 2.57 | 2.29 | 1.57 |

The trends of these averages are illustrated in Figure 6.8. Response types are identified along the horizontal axis. Average numbers of responses per student are on the vertical axis.

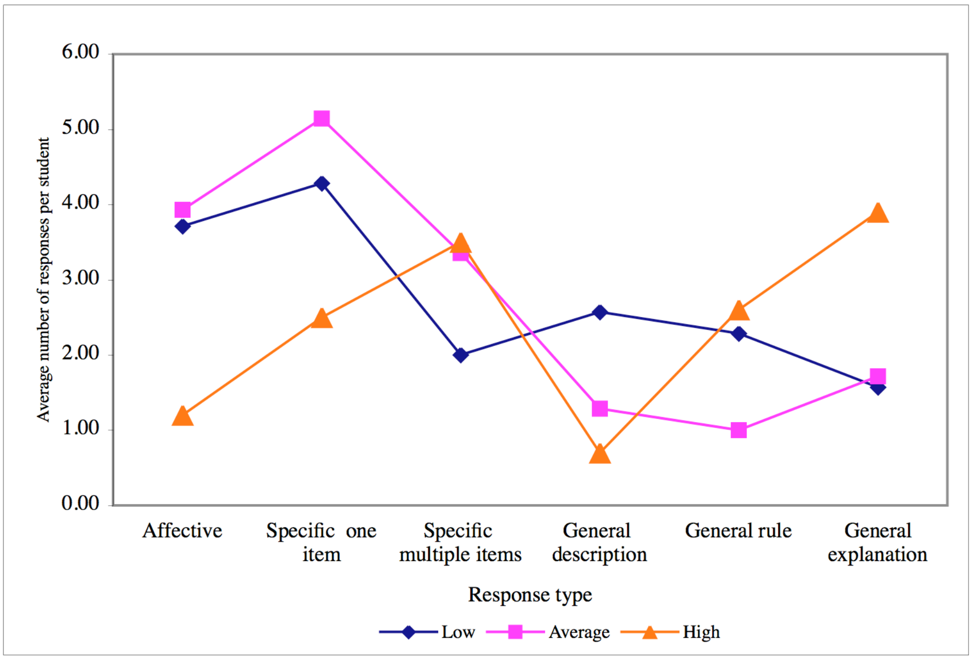


*Figure 6.8: Average number of responses per student to items in Sets 1-7 by ability group*

The High-ability group gave fewer affective responses per student than did either the Low- or Average-ability groups, and more explanatory, general statements. Average- and Low-ability students tended to make affective statements. Students in the Low- and Average-ability groups also gave more single item specific responses than students in the High-ability group. Students in the Low-ability group tended to avoid responding to all items in a set, even when prompted by being questioned about particular items.

The Average- and High-ability groups tended to respond to all or most items in a set as discrete entities (specific, multiple items) to the same extent. The Average-ability group tended not to provide general responses. The average number of general responses by the Low-ability group was similar to that for the High-ability group although these tended to be descriptions of superficial features of items (e.g., “They all have letters in them”) rather than statements of rules or an explanation. The High-ability group provided the greater proportion of explanatory, general statements.

The following parts discuss responses from each of the ability groups in further detail. Expected numbers of different response types and actual numbers are compared, and possible reasons for differences suggested.

### *High-ability group*

Responses from students in the High-ability group tended to vary significantly from the expected numbers in three of the response categories ($p<0.001$, Appendix D6). The pattern of expected numbers of responses compared with actual numbers of responses is illustrated in Figure 6.9.

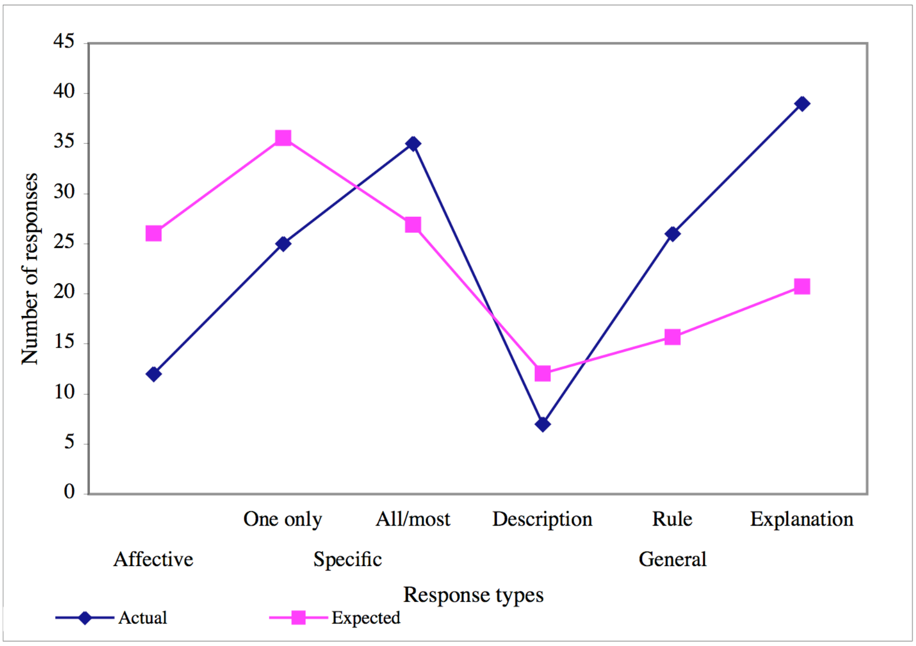


*Figure 6.9: Actual and expected response types from the High ability group to Sets 1-7*

Students in the High-ability group tended to make many fewer than expected affective statements and significantly more general statements of rules or explanations. Where the High-ability group tended to answer specifically, they usually answered all, or most, items in a set. The numbers of responses for the High-ability group are listed in Table 6.13 against each set of items in order of increasing difficulty. Sets 4 and 5 are listed as their sub-sets (4b, 4a; 5a, 5b).

*Table 6.13: Response types by High-ability group to item sets 1 to 7*

| | TYPE OF RESPONSE | | | | | |
|---|---|---|---|---|---|---|
| | Affective | Specific | | General | | |
| Set | | One only | All/most | Description | "Rule" | Explanatory |
| 1 | 1 | 1 | 6 | 0 | 9 | 4 |
| 2 | 2 | 2 | 3 | 2 | 2 | 7 |
| 5a | 3 | 4 | 5 | 1 | 2 | 4 |
| 5b | 2 | 0 | 2 | 3 | 2 | 7 |
| 6 | 1 | 5 | 3 | 0 | 0 | 2 |
| 3 | 1 | 2 | 1 | 0 | 4 | 9 |
| 4b | 1 | 5 | 5 | 0 | 1 | 0 |
| 4a | 1 | 5 | 3 | 1 | 3 | 5 |
| 7 | 0 | 1 | 7 | 0 | 3 | 1 |
| Total | 12 | 25 | 35 | 7 | 26 | 39 |

From the table it can be seen that students in the High-ability group tended to provide general statements, few of which were merely descriptive. In four instances, Sets 5a, 4b, 6 and 7, the numbers of specific responses were greater than the numbers of general responses.

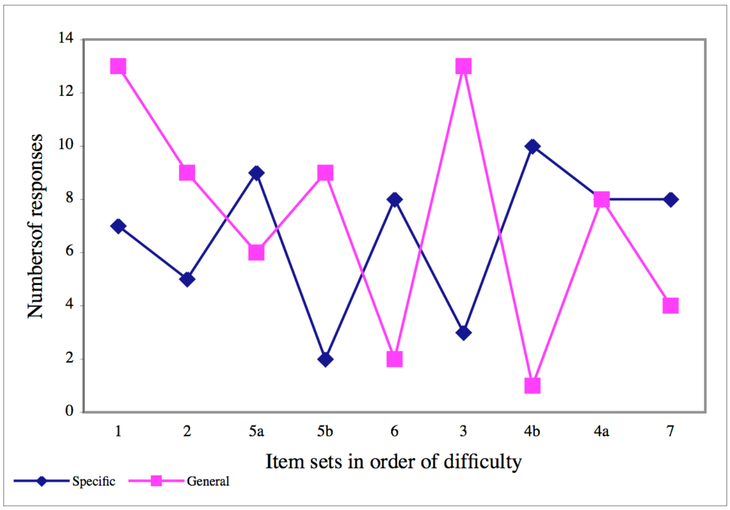


*Figure 6.10: Specific and general response patterns of the High-ability group to Sets 1 to 7*

The line graph in Figure 6.10 illustrates the rather complex response pattern for this ability group. Numbers of responses to specific items (one or many) decrease as the numbers of general responses (as rules, or explanations) increase. High-ability students provided either specific responses to items, or gave a generalised procedure or rule. Where these students could articulate a general rule, or describe a procedure in general terms, they did so. When they could not do so, they demonstrated understanding of generally applicable rules or procedures by addressing many items in a set. However, when students in the High-ability group were challenged by items in Set 6 and Set 4, this resulted in there being several single-item responses, and fewer general responses.

### *Average-ability group*

Differences between the expected and actual numbers of different response types for the Average-ability group were significant at *p*=0.01 level (Appendix D6). The differences are illustrated in Figure 6.11. Average-ability students did not offer as many general statements as would be expected. Their specific responses followed expected trends, although they gave more than expected single-item responses.

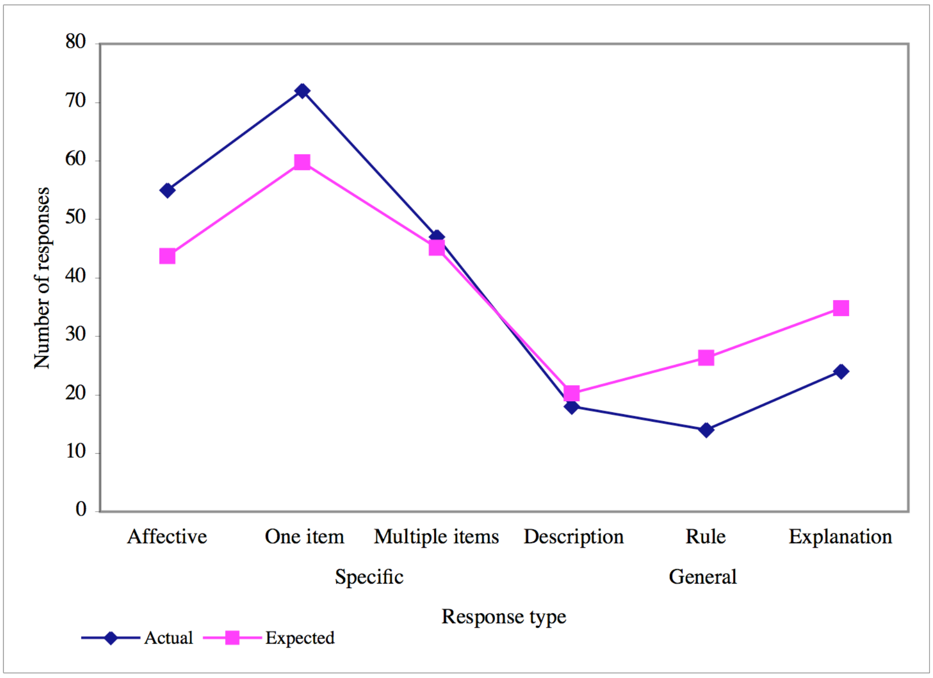


*Figure 6.11: Actual and expected response types from the Average ability group to Sets 1-7*

Numbers of the different responses by the Average-ability group of students are detailed in Table 6.14. Item sets are ordered by difficulty, with Sets 4 and 5 as two subsets.

*Table 6.14: Response types by Average-ability students to item sets 1 to 7*

| | | TYPE OF RESPONSE | | | | |
|---|---|---|---|---|---|---|
| | | Specific | | General | | |
| Set | Affective | One only | All/most | Description | "Rule" | Explanatory |
| 1 | 9 | 3 | 4 | 1 | 3 | 4 |
| 2 | 8 | 11 | 6 | 1 | 2 | 2 |
| 5a | 9 | 9 | 10 | 3 | 3 | 1 |
| 5b | 4 | 3 | 7 | 5 | 1 | 3 |
| 6 | 4 | 14 | 6 | 1 | 0 | 1 |
| 3 | 6 | 6 | 5 | 1 | 3 | 6 |
| 4b | 4 | 12 | 3 | 1 | 0 | 1 |
| 4a | 8 | 8 | 1 | 4 | 0 | 4 |
| 7 | 3 | 6 | 5 | 1 | 2 | 2 |
| Total | 55 | 72 | 47 | 18 | 14 | 24 |

Average-ability students gave many more affective responses than High-ability students, and more responses than expected (Figure 6.11). Such responses were most common at the beginning of the interview as students addressed the first two sets of items, and when they discussed Set 4a (addition and subtraction of fractions). Sets 2, 6 and 4b were most commonly answered with responses to a single item. These students offered few general statements, but the patterns of specific responses show a clear trend to single-item responses as sets of items become more difficult.

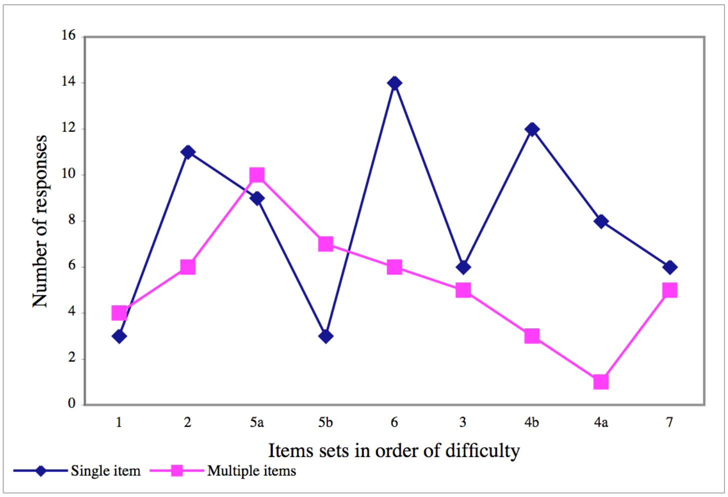


*Figure 6.12: Single- and multiple-item responses by Average-ability students*

Figure 6.12 illustrates these trends for both single-item and multiple-item specific responses. With the exception of Set 2, students in the Average-ability group tended to explain their thinking about multiple items in the less difficult sets (Sets 1 and all of 5). The more difficult sets (Sets 6, 3, 4 and 7) consisted of items with a range of difficulty estimates, many of which were greater than the upper ability estimate for the group, although the average difficulty estimates for Set 3 (0.425 logits) and Set 6 (0.34 logits) did fall within the range of ability estimates (approximately ±1.0 logits) for the Average-ability group.

However, when attempting to answer items in Set 4a (Item 3, $a/5 + a/10$; Item 15, $3p/4 - p/8$), some Average-ability students described, in general terms, what needed to be done to add (or subtract) algebraic fractions, but could not always apply the procedure successfully.

> S: Weeell, for the first two, [item] 3 and [item] 15 I guess I just… You got to make the bottom the same. You just can't plus different things like, because it would be different.
>
> [After some further conversation, the student continued to illustrate her understanding]
>
> S: Change the denominators, just so that they are the same and then try and work with the top, that's $3p$ minus $p$ that's $2p$.
>
> I: So you change the denominators. What do you mean when you work with the top?
>
> S: Just try and work out what it says, like $a$ plus $a$ is just $2a$ and $3p$ minus $p$ is $2p$.

Overall-response patterns for the Average-ability group suggest either that these students possess an implicit understanding of generalised procedures, or that they perceive some common features of items in a set. However, in the words of one student, they "don't know how to explain most of the stuff [they] know, [they] just know it".

### ***Low-ability group***

Students in the Low-ability group tended to make affective statements with approximately the same average frequency as students in the Average-ability group. They made single-item specific statements at the expected frequency, but fewer multiple-item specific statements.

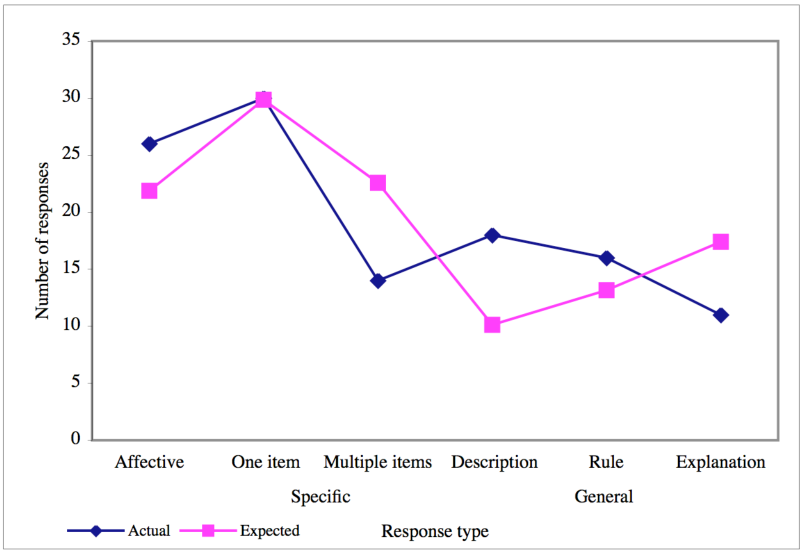


*Figure 6.13: Actual and expected response types from the Low ability group to Sets 1-7*

Unexpectedly, Low-ability students showed a greater tendency to make general responses, usually as overall descriptions of surface features of the items in a set. They also quoted rules a little more frequently than expected. Overall, the difference between actual and expected responses was significant at $p$=0.02 level (Appendix D6)

*Table 6.15: Numbers of response types by Low-ability students to item sets 1 to 7*

| | | TYPE OF RESPONSE | | | | |
|---|---|---|---|---|---|---|
| | | Specific | | | General | |
| Set | Affective | One only | All/most | Description | "Rule" | Explanatory |
| 1 | 3 | 3 | 1 | 3 | 4 | 2 |
| 2 | 4 | 4 | 2 | 5 | 1 | 0 |
| 5a | 4 | 0 | 5 | 2 | 3 | 2 |
| 5b | 4 | 4 | 1 | 0 | 0 | 1 |
| 6 | 3 | 6 | 1 | 2 | 0 | 0 |
| 3 | 3 | 2 | 1 | 0 | 5 | 3 |
| 4b | 1 | 2 | 1 | 1 | 1 | 1 |
| 4a | 2 | 4 | 0 | 3 | 2 | 0 |
| 7 | 2 | 5 | 2 | 2 | 0 | 2 |
| Total | 26 | 30 | 14 | 18 | 16 | 11 |

Low-ability group students made general statements in approximately the name numbers as the Average-ability group. In both cases, however, this was not as often as students in the High-ability group. However, Low-ability students tended to provide general descriptions of the items in the sets such as “they all have letters and numbers”, rather than cite rules or general procedures.

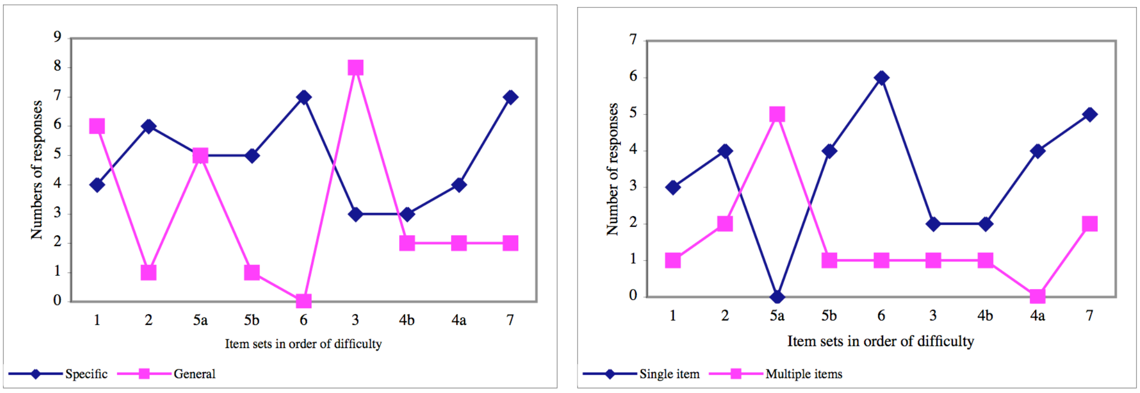


*Figure 6.14(a, b): Response patterns by Low-ability students to items in Sets 1 to 7.*

For most sets Low-ability students tended to focus on single items, except for Set 5a (linear equations). Students in this group also gave many rule-based responses to Set 3 (items with brackets), but few specific responses. Two response patterns in Figure 6.14 show diagrammatically the trends described in Table 6.15. Figure 6.14a, illustrates the numbers of specific and general responses to each set of items. Figure 6.14b, details specific responses to the sets of items – distinguishing single- and multiple-item responses.

With the exceptions of Set 1 and Set 3, Low-ability students tended to respond to individual items, usually only a single item from any one set (Figure 6.14a). The Low-ability group cited rules, with approximately the same average frequency as the High-ability group (Low: 2.29 responses/student; High 2.6 responses/student). Survey results suggest however that Low-ability students do not understand, nor apply, learnt rules in the same way that students from the High-ability group might do. In a previous section, (Pronouns) general statements of the High- and Low-ability students were found to be distinguished by students' use of different pronouns.

Figure 6.14b shows that students in the Low-ability group addressed multiple items only when discussing Set 5a (simple linear equations). The increase in the number of general statements for Set 3 is due to students citing rules such as "You do to the top what you do to the bottom", or "You find a number they both go into". In practice, revealed by survey data, these rules were understood in ways that did not result in correct responses to the items.

In general, frequency of occurrence of different response types was associated with different ability groups. Less-able students made more affective statements and usually addressed only one item in a set. Where less-able students offered general statements, these were either descriptive or vague rules based on a single item.

More-able students made fewer affective comments. Their specific responses were likely to address most or all items in a set. Average-ability students tended to make few general statements, while the most-able students were able to offer mathematically useful rules, or procedural explanations. When challenged by more difficult sets of items, High-ability students tended also to give specific responses only.

How these types of responses differed from responses when students were to discuss items that were not algebraic was also analysed. The results of this analysis are discussed in the following sub-section.

### Response types to Set 9 and Sets 1-7

The question as to whether students' types of responses would change when the focus of the interview shifted from algebra was investigated by presenting students with a ninth set of items. By comparing the occurrence of different types of responses to items in Sets 1 to 7 with responses to items in Set 9, it was hoped to ascertain whether the types of responses were influenced by the nature of the items, or were more generally how students framed their talk in an interview, regardless of any particular content.

Table 6.16 below shows the numbers of specific and general responses to items in Set 9, and in Sets 1 to 7, by each of the ability groups. The table also reports the different types of responses as a fraction of all responses made by each of the groups to Set 9 and to Sets 1 to 7. (E.g., the High-ability group gave 37 specific responses out of a total of 96 responses in Set 9. This means that 0.39 of the responses from the High-ability group were specific in nature, when they answered questions in Set 9.) This contrasts with 0.45 of their responses to Sets 1 to 7 being specific. When responding to items in either Set 9 or Sets 1 to 7, the High-ability group tended to make more general responses than specific responses. However, they made proportionately more general responses to Set 9 items.

*Table 6.16: Specific and general responses to items in Set 9 and Sets 1-7 within ability groups*

| Ability Groups | Set 9 Specific | Set 9 General | Set 9 Totals | Sets 1-7 Specific | Sets 1-7 General | Sets 1-7 Totals |
|---|---|---|---|---|---|---|
| NUMBERS OF SPECIFIC AND GENERAL RESPONSES | | | | | | |
| HIGH | 37 | 59 | 96 | 60 | 72 | 132 |
| AVERAGE | 71 | 68 | 139 | 119 | 46 | 175 |
| LOW | 25 | 34 | 59 | 44 | 45 | 89 |
| RESPONSES AS A FRACTION OF TOTAL GROUP RESPONSES | | | | | | |
| HIGH | 0.39 | 0.61 | 1.00 | 0.45 | 0.55 | 1.00 |
| AVERAGE | 0.51 | 0.49 | 1.00 | 0.68 | 0.32 | 1.00 |
| LOW | 0.42 | 0.58 | 1.00 | 0.49 | 0.51 | 1.00 |

The Average-ability group tended to use general or specific statements with approximately the same frequency in Set 9, but when talking about the algebra items in Sets 1 to 7 they tended to use specific statements (see Table 6.13). The Low-ability group gave a slightly higher proportion of general statements in responses to Set 9 questions than when responding to Sets 1 to 7. The Low-ability group made almost the same proportion of general and specific statements throughout responses to Sets 1 to 7.

Overall, the difference between the expected numbers of response types and the actual numbers was significant (*p*=0.01, Appendix D7). The greatest influence appears to be the increase in the numbers of general responses from students in the Average-ability group.

All groups tended to make proportionately more general statements when responding to items in Set 9, than when responding to items in Sets 1 to 7. Because of the greater number of students in the Average-ability group, there were numerically more responses than from the other groups. It is the difference between the actual and expected numbers of responses for each ability group that indicates changes. Line graphs in Figure 6.15 illustrate these patterns (see Appendix D7 for tables of values).

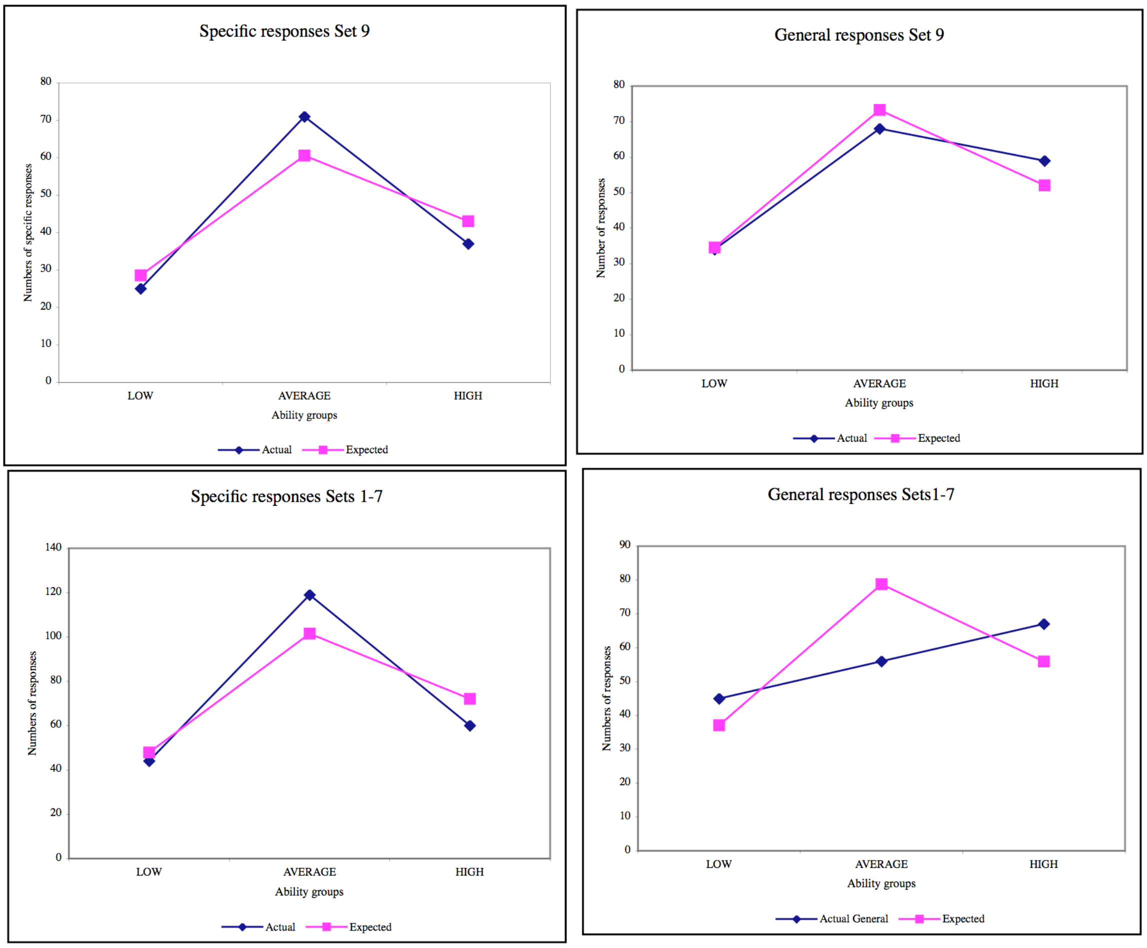


*Figure 6.15: Actual and expected numbers of responses to items in Set 9 and Sets 1-7 by each of the ability groups*

The patterns of actual and expected specific responses to the sets of items appear to be little changed as students move from answering algebraic items to items less focused on mathematical ideas. Low-ability students gave slightly fewer than expected specific responses, Average-ability students more specific responses than expected, and High-ability students fewer than expected.

When general responses are considered, the Average-ability group gives notably fewer general responses than expected when discussing Sets 1 to 7. For Set 9, however, although the number of general responses is still fewer than expected, the difference is less. Low-ability students provide an expected number of general responses, while the High-ability students continue to give more than expected general responses.

## Discussion

Analysis of students' conversations revealed different patterns of response types for each of the ability groups. Low- and Average-ability groups tended to make affective statements, usually expressing their dislike of the type of items ["I hate fractions"], or the fact that carrying out the particular mathematics was "simple" or "easy" (Table 6.9). This last comment was also often used as a justification for a procedure used by a student to manipulate an expression. It appeared from the interviews that students confused the ideas of *simplification* of expressions, with that of an action being *easy*, or a statement *simple* (see Chapter 7 on cancelling).

Many affective statements were made by students when discussing the first set of items. Students' tendency to make such comments decreased as the sets of items became more difficult. These more difficult sets of items were presented towards the middle or end of the interview (except Set 3). Thus, the decrease in statements implying uncertain attitudes, might be contributed to students becoming more confident with the situation, rather than with the particular mathematics. On the other hand, students tended to base interview responses on single items in the more difficult sets; they might therefore not have felt any need to express concern. Hence, no clear association of frequency of affective statements with set difficulty could be clearly established.

The affective responses do suggest that less-able students react emotionally to the appearance of items rather than to the mathematical ideas. If items looked familiar or like ones that students felt they could answer successfully, then they attempted to respond to the set of items.

The strategy of selecting items to which they might respond successfully underlies the observation that students in the Low- and Average-ability groups tended to respond to specific items without articulating any mathematical connections between items. Students would focus on what they could or would do, not on the mathematics. Hence, the Average-ability group tended not to make general statements about the sets of items, focusing only on answering each individual item in a set. Nor did this group tend to suggest any general "rules". Whilst the Low-ability group did make some general statements, these were often descriptions of surface features of items, rather than statements about the mathematics.

On the other hand, both the High-ability group and the Low-ability group quoted general "rules". The "rules" suggested by the Low-ability group tended to be couched in vague informal terms, rather than as propositional statements, and appeared to be based on a single example from the sets of items.

The rules suggested by students in the High-ability group were usually precise and expressed in formal mathematical language. Students in the High-ability group offered most of the observed explanations, but these tended to be elaborations of procedures, rather than conceptual justifications. The inference is that although students in the High-ability group apprehended mathematical concepts that underlie procedures, they were unable to make these understandings explicit.

When students are asked to explain their thinking about various examples of algebra, the type of responses implies a security of knowledge about mathematical concepts in the examples. Affective responses usually express attitudes unambiguously. However, attitudes towards aspects of mathematics, exemplified by the items, did not appear to be indicative of success or failure.

Specific responses to one or more items reveal that students can obtain correct, or incorrect, answers, but also suggest that they are unaware of mathematical connections existing between examples that share mathematical characteristics. Depth of understanding suggested by general statements might be gauged from whether the statements are descriptive of the appearance of items, rules that are stated in vague or precise terms, or explanations that are procedural or conceptual.

## MODALITY

Modality of speech indicates the confidence in, or insecurity of, the speaker's knowledge (discussed in detail in Chapter 2). Low modality is indicated by the occurrence of *hedges* (words or phrases that imply caution or hesitancy), the use of *well,* and hesitations in responding to questions. The first sub-section examines occurrences of a particular type of hedging statement - the *attribution shield* (where the student indicates lack of responsibility for the truth, or otherwise, of the statement uttered). This is followed by the examination of the use of other modal forms. These data are analysed in terms of ability groups and item-set difficulty. Finally, the overall results are discussed.

### Attribution shields

Attribution shields are those words or phrases used by students to excuse their ignorance, or their errors. Students used phrases such as: "That is what I was taught", "My teacher told me" or "It's just a rule". The use of attribution shields by students indicated their lack of confidence in dealing with the algebra items, or in their ability to explain reasons for the procedure they carried out, or in the correctness of the answer they obtained.

The numbers of these statements were small compared to the overall numbers of words in the interview conversations. Response patterns were analysed in two ways: firstly, as the average number of attribution shields per student varied by ability group and set difficulty; secondly, as a comparison of the average actual attribution shields per student in each ability group with the expected average.

#### *Attribution shields and ability groups*

Students in the Average-ability group tended to use more attribution shields per student than students in the Low-ability group. These students tended to use more attribution shields than students in the High-ability group. Students in both groups used, on average, the same number of attribution shields when talking about Set 5 and Set 8.

The average number of attribution shields per student in each ability group spikes at Set 5. This is probably due to the fact that Set 5 consisted of more items to be discussed than other sets, and that students tended to speak about each item individually. Set 1 and Set 9 also show higher numbers of attribution shields. Set 1 was the first set in the interview, and the number of attribution shields, like the number of affective statements (see previous section) might be indicative of lack of confidence engendered by the situation.

Figure 6.16 illustrates the average number of attribution shields per student in each of the ability groups for each set of items. Set 9 (background questions), unsuitable for Rasch modelling, is not included in the order of difficulty.

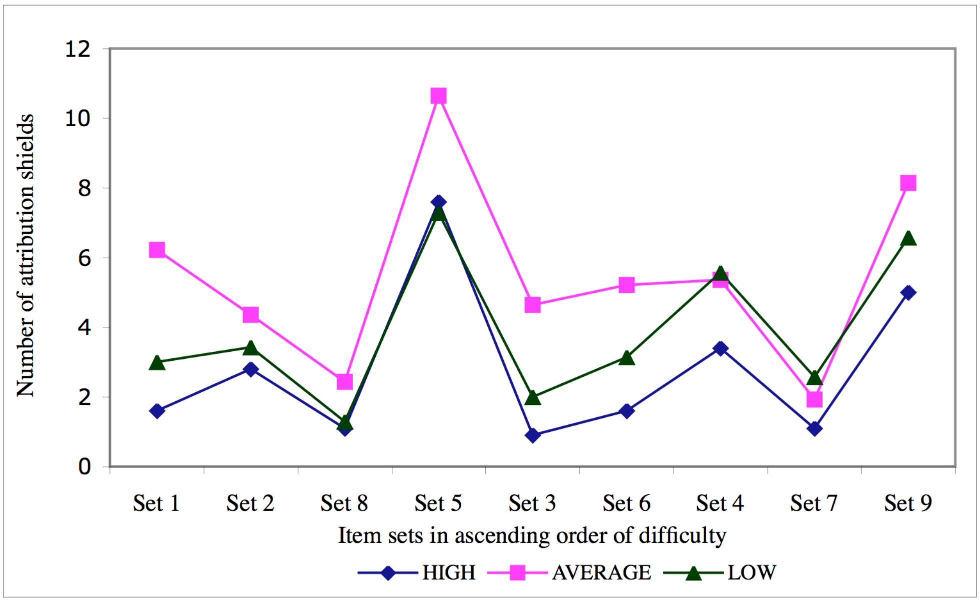


*Figure 6.16: Average numbers of attribution shields per student in each ability group for each set ordered by increasing difficulty (Set 9 excepted)*

Set 9 provoked considerable discussion, and questions were often answered in more personal terms than the items of a more obvious mathematical nature. Hence, there was a tendency for students to hesitate, or offer personal comments that were taken as attribution shields. (E.g., when asked to elaborate a procedure they had described for adding two two-digit numbers together, or connecting to the internet, students often said, “that’s how I was taught”, or “that’s just how I do it”.)

Set 8, although more difficult than Set 5, attracted fewer attribution shields for two reasons: the nature of the questions (Küchemann items, 1981) and the consequent responses; and, the confidence of the students in stating their suggestions for possible re-arrangements of the items. Despite their confidence, many students did not provide correct answers in the survey.

Numbers of attribution shields per student tended to increase as sets after Set 8 (difficulty estimate: -0.09 logits) became more difficult. The numbers then declined in responses to Set 7. This set had the greatest average difficulty (1.92 logits) because of the inclusion of two items, Item 17, $(6xy)^2$, and Item 10, $(x + y)^2$, which were answered incorrectly on the survey by most students. However, most students appeared to be confident when addressing these items with brackets; probably because of their reliance on a simple rule, and so they felt no need to further explain their thinking.

***Attribution shields and sets of items***

In this part, the frequency of use of attribution shields by each of the ability groups is examined with respect to the sets of items. Actual numbers of attribution shields uttered by students as they addressed each set of items are compared with expected numbers of attribution shields. Table 6.17 records the actual numbers of attribution shields uttered by each group for each set of items. The sets are arranged in order of difficulty. Set 9 is included as a reference point for the background use of attribution shields.

*Table 6.17: Numbers of attribution shields by each ability group for each item set, arranged in order of difficulty*

| | Sets of items in order of Rasch difficulty | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Set 1 | Set 2 | Set 8 | Set 5 | Set 3 | Set 6 | Set 4 | Set 7 | Set 9 |
| | -1.7 | -1.56 | -0.09 | -0.08 | 0.107 | 0.377 | 0.92 | 1.92 | |
| Ability Groups | Numbers of attribution shields | | | | | | | | |
| HIGH | 16 | 28 | 11 | 76 | 9 | 16 | 34 | 11 | 50 |
| AVERAGE | 87 | 61 | 34 | 149 | 65 | 73 | 75 | 27 | 114 |
| LOW | 21 | 24 | 9 | 51 | 14 | 22 | 39 | 18 | 46 |

Differences between actual and expected numbers of attribution shields are significant ($p$=0.005) for Sets 1 to 8. For each of the ability groups, the differences between the actual and expected numbers were not significant. However, some patterns do emerge. These are illustrated in Figure 6.17, which shows the comparison of the actual number of attribution shields as an average number per ability group in each set, compared with expected values (see Appendix D9).

The average number of attribution shields in each set reflects the number of words uttered by students in responding to each set of items (see Figure 6.2). Students tended to be more communicative when responding to Set 5 and Set 9. These were also the sets with most items. For each ability group, the average number of attribution shields spiked with Set 5, dropped when students responded to items in Set 8 (because of the way in which the questions were asked) and increased with set difficulty after Set 3. Students were not so verbose when responding to Set 6 and Set 4 (Figure 6.2), but the average numbers of attribution shields has increased with these difficult sets.

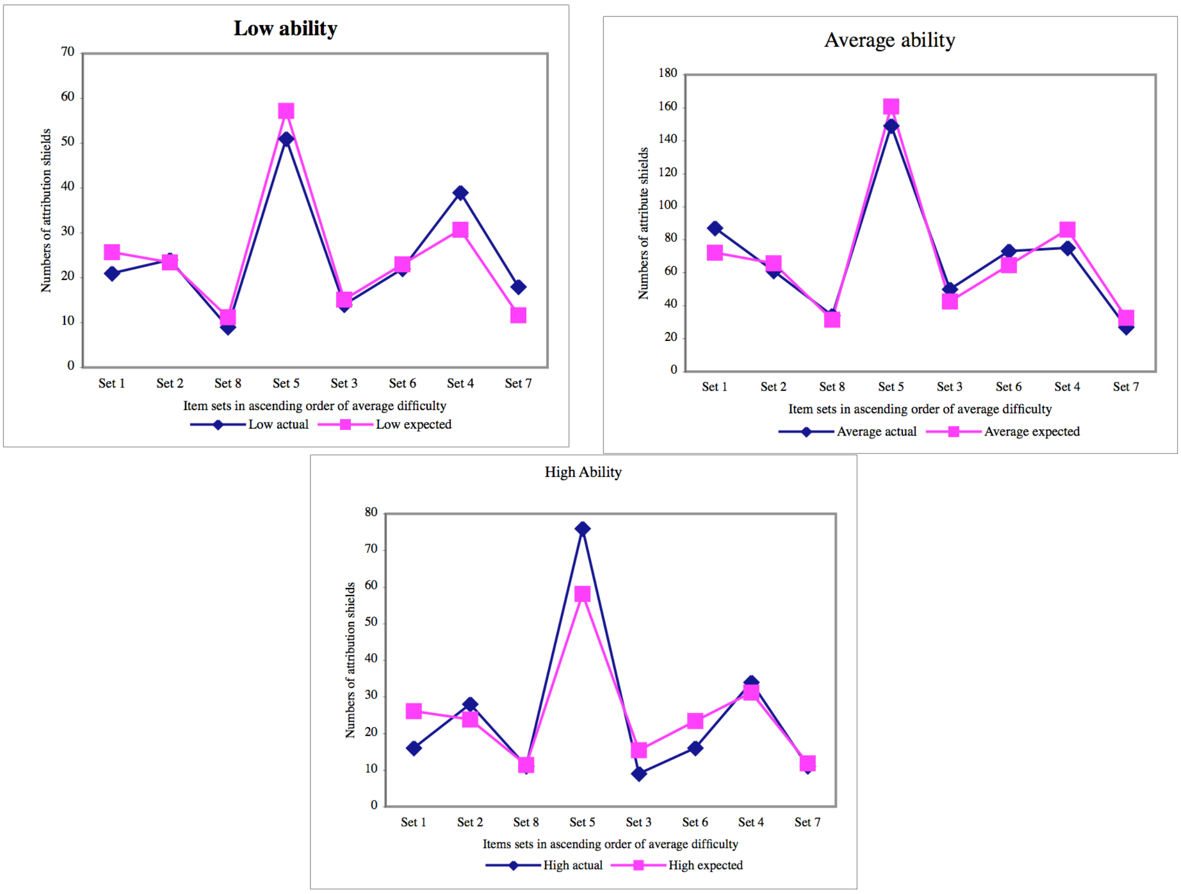


*Figure 6.17: Actual and expected proportions [as words uttered] of attribution shields made by each ability group by sets in order of set difficulty estimates.*

Students in the Low-ability group tended to use attribution shields when responding to the most difficult sets of items, Set 4 and Set 7. At times, these students could not find a single item to which they could respond confidently, as in the following example:

> ... It's stuff that I can't do. When I look at this, it's just stuff that I haven't learnt how to do really well, like, haven't learnt how to do them, like, much. Because I didn't listen when we were doing them [...] Ummm ...It's just sort of mixed up stuff, I didn't pay attention.

Another student explained her thinking as:

> I kind of just remember what I'm told and use that… It's just a rule. I dunno, it's what I've learnt.

Average-ability students tended to preface their interviews with attribution shields such as this student's response to items in Set 1:

> Oh, because, um, 2*ab*, I think it's... like, *a* times *b*, while 3*b* is three times *b* and *a* times… You just plus the 3*b* with the *ab*... I'm not sure. It's hard to explain how you do it 'cause, I'm just going on what, um, the teacher's been telling us… how to do it.

High-ability students did not, as a general practice, use attribution shields, except when prompted to explain their procedures, or results for items in Set 5 (equations). Often, these students stated that they carried out a procedure because it was "easier" or it was "just how I have always done it". Such expressions attribute reasons for a procedure to some vague authoritative source – relieving the student of any responsibility for the mathematics.

## Other modal forms

In this sub-section, analysis focuses on the use of hedges, and the use of "prosodic indicators" (voice changes and pauses) *well, um* and *yeah*. Hedges and prosodic indicators serve as "linguistic pointers to uncertainty and attendant cognitive vulnerability" (Chapter 2 for details). The extent of their use indicates the degree of uncertainty on the part of the student.

Hedges can be approximator shields (e.g., *just, like)*, plausibility shields (e.g., *I think/don't think* or *I don't know*) or prosodic hedges (*um* or *yeah*). Prosodic hedges allow the speaker a pause for thinking (*um*), or indicate insecurity on the part of the speaker as to what has already been said. For example, the response of a student to being asked to explain her thinking about the items in Set 1:

> OK. So look for like terms to, um, like, simplify it. And, um, yeah. Like 3*m* plus 2*m*, and then eight minus five. You can simplify it, um, yeah.

A lack of commitment or clarity might also be indicated by the use of a dispreferred second part (DSP in Figure 6.18), such as the use of *well* introducing a statement.

The most-frequently-used hedge words or phrases were common to all ability groups, and across all sets. The data are analysed firstly, with respect to the ability groups and secondly, with respect to the set difficulty.

### *Modality and ability groups*

The numbers and variety of the hedges and shields used by the students interviewed are illustrated in Figure 6.18. The actual number of hedge words or phrases as a proportion of the utterances recorded by the students in the study was relatively small, in the order of five-to-eight percent. If the actual number of hedges per ability group is counted and compared with the number expected, then the High-ability group used fewer hedges than expected, the Average-ability group more hedges than expected and the Low-ability group, slightly fewer hedges than expected.

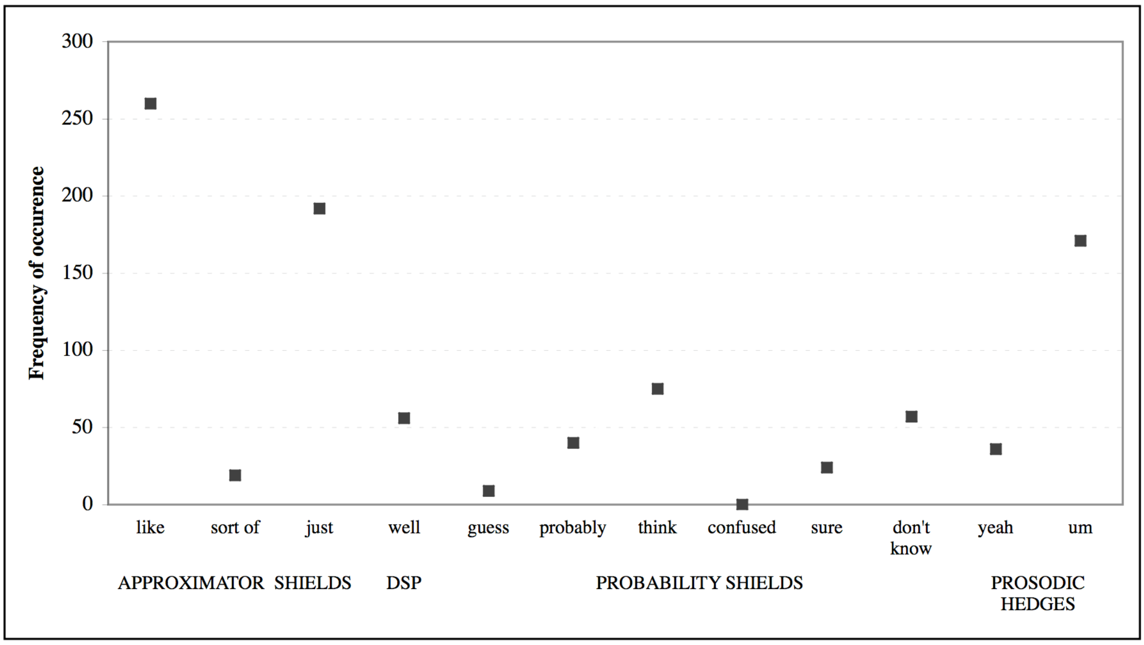


*Figure 6.18: Frequency of occurrence of hedges in Sets 1 to 8 by all students*

The differences between the actual and expected numbers of hedges used by the ability groups are significant (*p*<0.001, Appendix D8). The data are given in Table 6.18, which lists the actual number of hedges for each of the ability groups, the average numbers of hedges per student and the average numbers of hedges per responses made by the group.

*Table 6.18: Actual and Expected numbers of hedge words/phrases uttered by ability groups when talking about items in Sets 1 to 8 (Algebra) and Set 9 (background)*

| | Sets 1-8 | | | | Set 9 | | | |
|---|---|---|---|---|---|---|---|---|
| Ability groups | Responses | Hedges | Hedge/ student | Hedge/ response | Responses | Hedges | Hedge/ student | Hedge/ response |
| HIGH | 688 | 658 | 66 | 0.96 | 222 | 230 | 23 | 1.04 |
| AVERAGE | 787 | 2040 | 146 | 2.59 | 291 | 525 | 38 | 1.80 |
| LOW | 457 | 564 | 81 | 1.23 | 183 | 184 | 26 | 1.01 |
| Totals | 1932 | 3262 | | | 696 | 939 | | |
| Average | 62 | 105 | | | 22 | 30 | | |

Table 6.18 also shows the ability groups, the total numbers of responses by each of the groups to Sets 1 to 8, and Set 9, together with the total number of hedges recorded for those sets. Figure 6.19 illustrates the data for actual and expected numbers of hedges in Sets 1 to 8 by ability groups. There are significant differences between the expected and actual frequencies of hedges occurring within each group when the students discuss the algebra sets. The Low-ability group and High-ability groups used fewer hedges than expected, whilst the Average-ability group used hedges more frequently than expected.

*Figure 6.19: Actual and expected numbers of hedges uttered by ability groups in Sets 1 to 8*

From the table it can also be seen that the number of hedges/response is greater for Sets 1 to 8 than for Set 9, notably so for the Average-ability group. The Average-ability group shows a marked drop in the average number of hedges per response in Set 9, compared to the average number of hedges per response to Sets 1 to 8. The frequency of hedges by the Low-ability group also dropped slightly from Sets 1 to 8 to Set 9, but rose slightly for the High-ability group. These trends are illustrated in Figure 6.20.

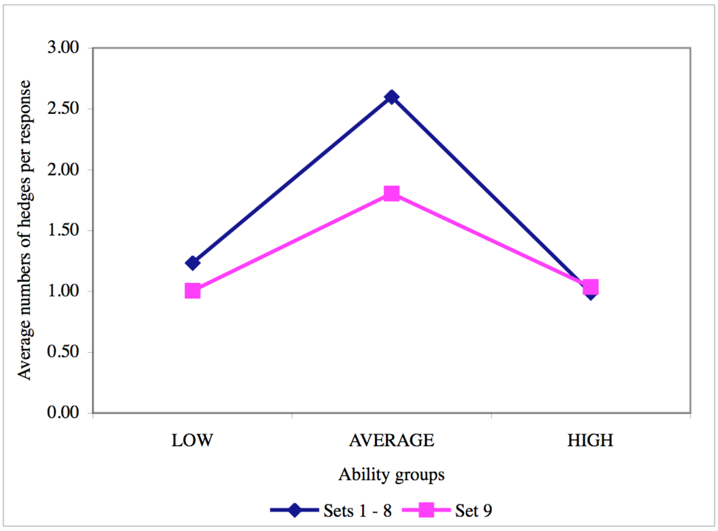


*Figure 6.20: Hedges/response by ability groups for Sets 1 to 8 and Set 9*

The decrease in the number of hedges as students responded to questions in Set 9 might be attributed to their increased confidence in discussing ideas that were not algebraic (although some questions were also explicitly mathematical). It might also be that students had become more comfortable with the interview situation. However, this does not appear to be the case with students in the High-ability group, who showed an increase in the average number of hedges in Set 9.

To further illustrate patterns of occurrence of hedges in each of the ability groups, the numbers of hedges by each student were standardised. Figure 6.21 shows the standardised number of hedges per response in Sets 1 to 8 for individual students ordered according to increasing ability. The numbers of hedges/response by students in the Average-ability group show a greater scatter (more variation) than is the case in the other two groups (Figure 6.21). This group also used the greatest number of words to express their ideas.

Students in the High- and Low-ability groups tended to make fewer hedges, than students in the Average-ability group. All standardised numbers of hedges for these two ability groups fell within one standard deviation from the mean, or lower, as is the case for two students from the High-ability group. The figure also shows clearly the gap between the three most able students in the High-ability group (High-3, see earlier) and the other seven students (High-7) in that group.

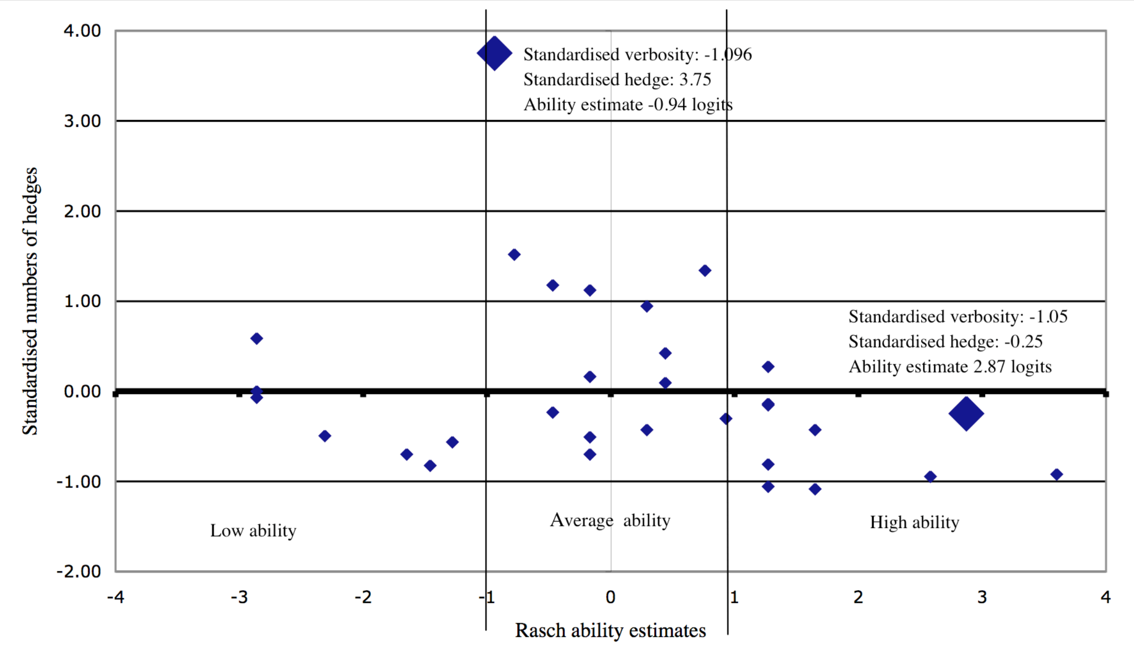


*Figure 6.21: Scatter of the number of hedges per response by individual students in increasing order of ability estimates.*

However, it was not always the numbers of hedges that contributed to relatively large numbers of words used by students. Two interesting examples should be remarked, each of students with low standardised verbosity scores, but high standardised hedge scores. These are indicated by large markers in Figure 6.21.

The first is that of the student (ability estimate -0.94 logits) on the cusp between Average-ability and Low-ability at the -1.0 logit line. The student had a standardised verbosity score of -1.06 (Figure 6.1), but has a standardised hedge score of 3.75. The score indicates the student's uncertainty of knowledge, and how this uncertainty is expressed in utterances of low modality.

> I: Tell me what goes on in your head when you see questions like those and those in Set 1.
>
> S: Um, just trying to answer it. Just, like, mainly the numbers and stuff, and not the letters much.
>
> I: Why? What's wrong with the letters?
>
> S: I don't know. Well, I just don't use the letters that much, and then just…[short pause]. Yeah, I don't work with the letters that much.
>
> I: Why is that?
>
> S: I don't know. Because I know more about the numbers. It's easier.
>
> I: It's easier. So, what you do in the case of Question 1 [Item 1]?

> S: Just, like… Oh well, because it's 3*m* plus 2*m* that's 5*m* and then eight minus five that's three; so that's 5*m* plus three, I think. Something like that.
>
> I: So you are happy to work with those letters?
>
> S: Yeah. I suppose.
>
> […]
>
> I: What do the letters stand for? Or mean?
>
> S: I dunno. There's… I dunno.

The student need not have struggled with many of the items in Set 1 as their difficulty estimates were less than, or near his ability estimate. His conversation demonstrates, however, that he found dealing with the notation of algebra challenging, and difficult to explain coherently.

The other example is that of the student with the second highest ability. She also had a low standardised verbosity score (-1.05) but a standardised hedge score of -0.25. Although this is within the average range, it suggests that the student's utterances, few as they were, consisted of many hedges. Although very successful in responding to the survey, the student was uncomfortable in the situation of having to explain her mathematical understanding. This is indicated by the relatively low modality of her conversation at the beginning of the interview.

> I: What goes on in your head when you see question like those, and those [Set 1]?
>
> S: OK so… Look for like terms, to, um, like simplify it, and, um, yeah… Like 3*m* plus 2*m* and then eight minus five; you can simplify it … um, yeah…
>
> I: When you get expressions like these in group 2 [Set 2]?
>
> S: Just go… I don't know. Just follow the rules we've been taught like… Like in the first one, like, four times five times b; and the second one is *a* times *a* is *a* squared and then, then …

This particular student should have had little difficulty in answering these items correctly; her ability estimate was 2.87 logits, and no items in the first two sets had difficulty estimates greater than 0.33 logits (Item 6, $5a - 2b + 3a + 3b$). However, her lack of confidence in explaining her reasoning is apparent.

The frequency of the use of hedges as students talk about their algebraic thinking indicates their ability to deal with the items, and/or their ability to articulate their mathematical understanding. The more frequent the hedges, the lower the modality of the utterances, the greater the cognitive struggle indicated.

This is most marked with students in the Average-ability group. These students were willing to respond to all items, but struggled to do so successfully in many cases, hence the increased frequency of the occurrence of hedges. The drop in the number of hedges as students in the Average-group responded to Set 9 questions indicates their increased confidence when not having to deal with algebraic ideas.

Students in the Low-ability group could also be expected to struggle with many items, but as has been noted in previous sections, they tended to select items to which they *wanted* to respond, and so avoid feelings of uncertainty that would be expressed in the modality of their talk. Frequency of hedges in the High-ability group, for whom most items would have posed little difficulty, was low overall. One interesting exception has been identified and discussed.

### *Modality and Sets of Items*

When the numbers of hedges and shields uttered by all students to each of the sets of items are analysed, the frequency of occurrence is not so clearly aligned to increasing set difficulty. These data are given in Table 6.19, which lists the sets of items in order of their Rasch difficulty estimates, the total numbers of responses, and the numbers of hedges used in each set, and the expected hedges for each set.

The proportion of hedges and shields in the first two sets is high (1.83 in Set 1 and 1.95 in Set 2). The least numbers of hedges occur in Sets 6 and 3, whilst the proportion of hedges to responses drops as students respond to Set 9. The low frequencies of hedges correspond to a drop in the number of responses to Sets 3 and 6. However, as has been discussed in the previous part, individual students might tend to use a greater number of hedges than others.

Trends in numbers of responses and numbers of hedges are illustrated in Figure 6.22. The peaks of responses, and hedges, indicate points where students in different ability groups began to be challenged. The first drop in the numbers of responses and the number of hedges occurs at -0.09 logits (Set 8), a difficulty level near the average difficulty of all items (0.0 logits), and near the average ability level of all students (-0.64 logits).

*Table 6.19: Comparison of the numbers of hedges for each set of items*

| Set | Difficulty estimate | Responses | Actual hedges | Expected hedges |
|---|---|---|---|---|
| 1 | -1.66 | 237 | 473 | 379 |
| 2 | -1.56 | 181 | 385 | 289 |
| 8 | -0.09 | 113 | 188 | 181 |
| 5 | -0.08 | 562 | 976 | 898 |
| 6 | 0.34 | 205 | 272 | 328 |
| 3 | 0.425 | 245 | 313 | 392 |
| 4 | 1.046 | 259 | 433 | 414 |
| 7 | 1.92 | 130 | 222 | 208 |
| 9 | n/a | 696 | 939 | 1113 |
| Totals |  | 2628 | 4201 | 4201 |

Although the first two sets were relatively easy (Set 1: -1.66 logits; Set 2: -1.56 logits), and within the ability range of the students interviewed from the Low-ability group (-2.86 to -1.28 logits) the number of hedges is high. This might be attributable to the students' insecurity and their settling to the interview situation. It might also be due to the fact that Low-ability students were attempting items that, in some instances, were challenging. The high number of hedges for these sets can therefore be interpreted as indicating cognitive struggle for some students.

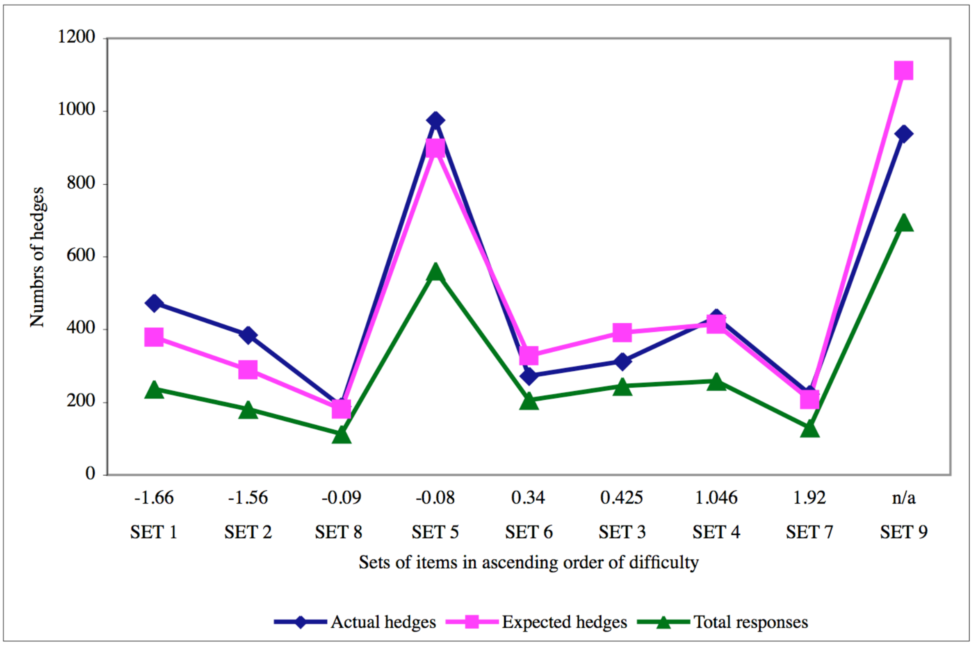


*Figure 6.22: Numbers of responses and numbers of actual and expected hedges by all students to Sets 1 to 9*

Overall, differences between actual and expected numbers of hedges by each set are significant, ($p$<0.001, Appendix D10). The differences between expected and actual numbers of hedges for Sets 1, 2, 6, 3, and 9 would appear to be the main contributing factors to this. The lower numbers of responses, and hedges, to Set 8 is, as has been previously noted, due to the nature of the questions asked for this set. The rise in response numbers, and hedges for Set 5, reflects the participation of most students in the Average-ability group, and their subsequent struggle to deal with all the items.

The rise in numbers of responses and hedges to Set 4 reflects the struggle by the students in the High-ability group to deal with all the items in this set. The High-ability group became more verbose than usual when talking about these items (see also Figure 6.2). When the Average- and Low-ability students were presented with Set 4, most students tended to answer specific items only, often avoiding harder items involving addition and subtraction of fractions. Hence, the number of responses, and consequently the number of hedges, dropped overall.

Responses to Set 7 varied. This set consisted of items, all with brackets (see Table 6.2, and Appendix B2). Students often appeared to be confident when asked to explain their thinking about these items, citing a rule that brackets meant "to do it first". Most students were also confident to state that the result for Item 10 [$(x + y)^2$] was $x^2 + y^2$ (see Chapter 5, Error Analysis). It was this incorrect response, rather than students avoiding this item, that resulted in the high difficulty estimate for Item 10 (3.56 logits), and consequently, the average difficulty estimate for Set 7.

In general, modality of students' utterances declined as sets of items became more difficult. Changes in modality indicated points where the different ability groups of students began to find the mathematical ideas difficult to make explicit. The cognitive struggle by students was made apparent through the change in modality (i.e., an increase in the numbers of hedges) of their conversation.

## Discussion

Modality of students' responses reflects their cognitive security. Low modality, the presence of hedges and shields, and of dispreferred second parts, indicated that students had to struggle with articulating algebraic ideas presented to them. This might be because they have no knowledge of the concepts, or because they could not explain the mathematical reasoning behind the procedures they used.

The frequency of use of attribution shields by ability groups varied according to the difficulty of the set of items, and with the nature of the items. Students in the Average-ability group tended to make more statements containing attribution shields than students in either the Low- or High-ability groups. Students in the Average-ability group also tended to use more hedges than the students in the other two ability groups, across all item-sets. The High-ability group tended to use fewer than expected hedges or shields. The types of hedges and shields used by the students were similar in all ability groups.

The conversations of both Average- and Low-ability groups tended to be of a lower modality (more hedges and shields) than those of the High-ability group when discussing the sets of algebra items (Sets 1 to 8). However, the modality of the conversations of the students in the Low- and Average-ability groups increased when the students responded to items in Set 9 (background). The number of hedges and shields was greatest, and more varied, in the group of Average-ability students.

Students might be willing to respond to items with which they feel confident, based on previous experience, and familiarity of form. However, having to answer items, such as those used in the study, as part of textbook-based practice is different to having to explain one's thinking about the mathematics. Whilst students might be able to execute a procedure by following a rehearsed set of steps, articulating their thinking provides a challenge on a different level. The low modality exhibited by many students at the beginning of the interview might be attributed, at least in part, to this unfamiliar situation. Regardless of the situation, an inability on the part of a student to offer mathematical explanations with a degree of confidence indicates insecurity of knowledge.

The data also suggest that the Low-ability students' conversations were, from the outset, of a lower modality than those of students in the Average- and High-ability groups. The modality of the discourse of students in the Average-ability group became lower as they encountered items that were more difficult than, or approximately the same as, their ability estimates. The modality of the utterances of students in the High-ability group became lower only as they addressed items with high difficulty estimates.

Where modality of a response is low, the conclusion can be drawn that the algebraic ideas are too difficult for the student. Where modality is high, the inference made is that the students understand the ideas about which they speak.

## TENSE

One further aspect of language in use is that of the tense in which ideas are expressed. Use of the past tense (recount genre) describes or explains individual, specific action. On the other hand, general principles are stated in the simple present tense.

Six of the sets of interview questions were examined for the type of tense predominantly used by each student. Almost exclusively, students replied using the simple present tense. There was one instance of a student using the past tense, and 23 instances of students using a conditional tense, either “I would” or “I’d”. The data are summarised in Table 6.20.

*Table 6.20: Numbers of responses using present, past or conditional tenses*

| | NUMBERS OF RESPONSES USING EACH TENSE | | |
|---|---|---|---|
| | Present | Past | Conditional |
| Set 1 | 27 | 0 | 5 |
| Set 2 | 26 | 0 | 4 |
| Set 3 | 27 | 0 | 2 |
| Set 4a | 27 | 0 | 3 |
| Set 4b | 26 | 0 | 2 |
| Set 5 | 29 | 1 | 5 |
| Set 6 | 25 | 0 | 7 |

Set 4, was presented to students in two parts, and the responses recorded for Set 4a, the addition or subtraction of fractions, and Set 4b, simplification by division of fractions. Although students predominantly used the simple present tense, specific instances of the use of the conditional tense were noted. Two students, one in the High-ability group, and one at the lower end of the Average-ability group tended to use the conditional tense in all their responses.

The conditional tense was more likely to be used by students in the High-ability or Average-ability groups: High-ability students used the conditional tense in ten instances; Average-ability students used the conditional tense in 13 instances; and, Low-ability students used the conditional tense in 5 instances. The High-ability students tended to use the conditional tense most frequently when discussing items in Set 6 (four instances out of seven).

### Discussion

Students in the early years of secondary school seem to couch their responses to questions about their understanding of algebra in the simple present tense as a matter of course. This could be the result of their classroom experiences where the teacher structures many statements as generalisations of principles or procedures. Hence, students learn that this is the conventional linguistic form appropriate to classroom use. Many statements made by students in the interviews were, however, about specific examples, not general principles.

The small number of cases where the conditional tense was used seemed to be largely due to an idiosyncratic speech form used by two students when addressing the interviewer. These particular students also displayed an unwillingness to do other than describe the procedures they used, or elaborate on more general statements. They appeared to focus on doing the mathematics, but were insecure when asked to explain their thinking further. Two of these students were in the High-ability group.

Use of the simple present tense by students in the Low-ability group might be attributed to their tendency to make vague, rather than general, statements. The data, therefore, suggest that tense cannot serve as a linguistic indicator of confidence of knowledge. However, the use of the conditional tense does suggest to the listener that the speaker is not necessarily secure in the truth of the statements made.

## CONCLUSION

The pragmatics of language used by students during interviews focused on linguistic elements: verbosity; use of pronouns (*I*, *you* and *it*); types of responses; and modality of the conversations. Linguistic characteristics of students' discourse varied according to student ability and difficulty of the algebra discussed, and indicated changes in the confidence of students. Changes in confidence were interpreted as indicating conceptual security, or conceptual challenge.

Students' uncertainty of knowledge was inferred from: the use of overlong and convoluted explanations, of a low modality; an absence of generality; and a tendency to personal observations, or vagueness. Certainty was inferred from: the use of succinct and precise explanations, often general, and of a high modality; and the use of a literate (mathematical) register.

The greater verbosity of students' responses could be attributed to several factors: the tendency to respond to each individual item in a set of items, rather than making any general observations that included all items; low modality occasioned by frequent hedges and shields; and numerous false starts. Thus students' prolix explanations indicated an insecurity of knowledge.

An absence of ability or inclination to generalise was indicated by less frequent use of the pronoun *you* and the more frequent use of the pronouns *I* or *it* (or plural equivalents). The use of *it* was an indicator of vagueness.

Vagueness might often sound like a generality. Vague statements were expressed in the simple present tense, which tended to imply that an utterance was about a general principle. However, other linguistic features and close analysis of the truth-value of these statements belied this impression. Vagueness became more frequent when students' understanding of ideas was greatly challenged. Not only did the use of *it* become more frequent, but also general statements tended to be couched in informal language, to be descriptive of appearances of items, or to state rules of limited mathematical truth or use.

Where less-confident students resorted to the stating of rules, the rules were often justified by being attributed to an authority rather than by any mathematical rationale. More confident students found little need to justify the rules they stated. However, when students do state rules, the bases on which they make those statements need to be explicitly identified, and articulated. Less-confident students tended to base their rules on a sole example in a set of items; more confident students on mathematical attributes of all items in a set.

Less-verbose responses from students might be the result of their responding to single, selected items, as was often the case for Low-ability group students. On the other hand, the equally laconic responses from capable students were a result of their being able to generalise across all items, seeing little need to say anything further. When more able students did speak about individual items, their discourse was of the appropriate mathematical register.

It is not necessarily the occurrence of a single linguistic feature that indicates confidence, or lack of confidence, in one's utterances. Taken together, however, several linguistic features that characterise the discourse of students can serve as indicators to the depth, or quality, of their understanding. Many examples from interviews have been alluded to in this chapter. To demonstrate more clearly the connections between students' mathematical success in particular contexts, and the associated pragmatics of their language, five examples are elaborated in the following chapter.

# CHAPTER 7: CONNECTING LANGUAGE USE AND ALGEBRA UNDERSTANDING

Listening to students explain their understanding of introductory symbolic algebra, one is struck by how the quality of the discourse appears to reflect the depth and confidence of their mathematical understanding. The luxury of being able to replay tape recordings of interviews, and to read over and over again transcripts of those interviews enables greater attention to be paid to both aspects of classroom conversations – the content, and the use of language – than is afforded teachers during lessons. One is able to hear, and read, subtle linguistic changes that indicate students' secure, confident understanding, or which signal their cognitive struggles.

Rasch-modelling of students' responses to a survey of introductory algebra provided a framework of objectively measured concept development. Students' explanations of their thinking about items in the survey provided evidence for the types of conceptual changes necessary for successful engagement with items of increasing mathematical complexity. Students' use of language also changed with their ability to meet conceptual challenges of the algebra.

In this chapter, the association found between the development of algebraic concepts and that of students' mathematical register, identified by their use of language, is described. Five examples of ways in which students' explanations were structured, together with their algebraic understanding, are discussed. These examples serve to illustrate effects of the language used by students, and what that language both reveals, and conceals, of the students' conceptual development. The four examples discussed in the first sections are aspects of algebra selected because all students across the range of ability addressed them in both the survey and the interviews. The examples are: the role of brackets; the notion of *cancelling*; the concepts of *like terms* and *unlike terms*; and the conjoining of terms. A fifth section examines the generalisations made by students in the High- and Low-ability groups.

In discussions of the examples, use is made of the results of analyses of the survey responses, and the linguistic analyses. Reference is also made to the notion put forward by Pimm (1987) of the use of metaphor in mathematics. The reflective analysis that comprises this discussion highlights possible important influences that understanding metaphor in terms of algebra might have on student's understanding, and on their own use of language.

## BRACKETS

The NSW Mathematics Syllabus first makes mention of brackets in Stage 4 statements. At this time the mathematical meaning of brackets is introduced "as an operator" (Board of Studies NSW, 2002b, p.59). Hence, students first encounter explicit teaching of the formal use of brackets in Stage 4 (first years of secondary school), although interview responses suggest that some students might be introduced to the use of brackets in primary school. (This can be inferred from notation used in the Stage 2 Content statements of the Syllabus, pp.54-55.) Most students explained that brackets meant "you do what is inside first" (25 out of the 31 students interviewed). Five students (out of 31 interviewed) implied a vague understanding of the grouping role of brackets, by stating that brackets "kept it separate".

The idea of brackets grouping terms to make a mathematical entity was expressed by seven students, all of whom were in the High-ability group, or near the upper bound of the Average-ability group. Three of these students demonstrated a flexible understanding of brackets based on the algebraic context, suggesting that they thought of brackets as grouping terms together, but also as conveying an instruction to act. The action was to be carried out *on* the entity enclosed by the brackets by the mathematical objects outside the brackets, rather than *within* the entity itself.

> They [brackets][*] can either mean do this first or they can mean times it by what ever is outside those brackets, or squaring, or square root…

The explanations from the other four students seemed to imply that although they expressed the role of brackets in terms of grouping they saw the grouping as an instruction to operate on the terms within the brackets.

> To tell us the order to do things, do brackets, to work out brackets. It's that part of the question that you can work out first, or just [?] an entire group instead of… um … It's about an entire group of numbers rather than just one or two.

Of the 19 students who described the purpose of brackets as indicating that one should "do" the operation inside the brackets first, four successfully answered Item 20 [Multiply $x + 5$ by 4]. Most of the others gave the answer as $20x$. Two students from

---

[*] Square brackets inserted in a verbatim excerpt from an interview indicate either parts have been left out […], or contain the interviewer's comment for clarification.

the Average-ability group and one from the High-ability group said they "did not know" what the role of brackets might be.

One student who understood the grouping role of brackets gave partly correct responses to Item 17 [$(6xy)^2$] as $(6xy)\times(6xy)$ and Item 10 [$(x + y)^2$] as $(x + y)(x + y)$, but could go no further. It is as if this student remembered that to square a number, it was to be "multiplied by itself", and that the student also saw the brackets as grouping the terms into one mathematical entity. However this remained a surface response and there was no indication that the student saw any relationships between the terms, nor that the factored form could be further manipulated.

Although students' understanding of brackets might have been established in earlier years, the limitations and insecurity of their understanding was evident, as demonstrated by the above examples, and also by their language. To illustrate how one student's lack of commitment to the ideas he explains might be inferred by the listener, the following examines one explanation given by a student in the Average-ability group, in response to the question "What is the role of brackets?"

> To do them first. I don't …Why the... There was a little… Oh no! I remember in Year 6 I did something, it was brackets first, then I think you multiplied then divide then plus and minus. That was how I was taught to do the brackets first, that was the main thing you had to do.

The student's explanation comprises elements of vagueness, low modality, and an appeal to authority, all of which indicate a lack of commitment to the ideas expressed. Vagueness is apparent in the student's use of *them*, ("To do them first"), *something* ("I did something…"), *it* ("…it was brackets first"), and the repetition of the order of operations rule "you multiplied, then divide, then plus and minus". The objects of the operations remain unclear, and vague.

The modality is low: lack of confidence is apparent in false starts as the student tries to elaborate in response to the interviewer's waiting, "I don't… Why the… There was a little…Oh no!" The appeal to authority in "That was how I was taught …" is also evidence that the student cannot make explicit mathematical reasons for the use of brackets.

The student's explanation is, perhaps, acceptable in a classroom context where a teacher tries to understand, and encourage some conversation. It demonstrates a cognitive struggle to articulate ideas that are understood only implicitly. The understanding that is explicit, however, does not allow the student to deal with later more formal use of bracket notation in contexts dissociated from an immediate arithmetic context.

Most responses in the interviews were simple statements of the initially-learnt rule for the use of brackets, “you do it first”, which suggests that that few students’ understandings have been modified in the light of later mathematical experiences. This “static view” of brackets (Warren, 2003), first developed in middle or late primary school seems to persist, and to dominate any later thinking.

Students’ lack of confidence in providing explanations, illustrated by the example, might indicate that these first ideas are being challenged by later experiences in algebra, but that these are as yet imperfectly assimilated. On the other hand, because of their success in handling items such as Item 20 and other similar items, it would appear that the students of greater ability have richer, more flexible understandings of brackets.

High-ability students tended to make the same simple statements about the role of brackets, but did so more confidently than students in other ability groups. Their statements sounded like generalisations that, for the individual student, embodied a greater complexity of meaning, but which were unable to be made explicit, except in the words of the initially encountered (and perhaps rote-learned) rule.

## CANCELLING

*Cancelling* is a term often used in the classroom, in textbooks (e.g., Lynch & Parr, 1982; Fitzpatrick, 1992) and tutorial systems (e.g., www.quickmath.com). Most often the term is used to describe the elimination of terms in an expression, either multiplicatively or additively. Often such elimination is tracked by terms being “crossed off”. However, flexible use of the term by experts in various contexts appears to be confusing for novices, not least, because metaphors of “crossing out” and terms being “the same” are often used in close linguistic association. The following discussion illustrates these confusions, and how students’ use of language indicates their limited understandings.

*Cancelling* was a term used by students in the Average- and High-ability groups when speaking about simplifying expressions of Item 8 [$4ab/4b$] and Item 16 [$2a^2$

× $5a/4$]. No student in the Low-ability group used the term. When asked by the interviewer to explain what was meant by *cancelling,* many students in the Average-ability group offered explanations similar to the following:

> S1: It means there's one $b$ there [indicating the numerator] and one $b$ there [indicating the denominator], so you can scratch them both out. It's like dividing them. And there's two fours so you can scratch them both out because they're the same…

And from another student:

> S2: …you are dividing. $b$ divided by $b$ is zero, and four divided by four is zero. Or one. So, you can take both fours out and both $b$'s out and just have $a$.

This student then responded to the interviewer's question, 'So is $b$ divided by $b$, and four divided by four zero or one?'

> S2: One, I think. Yes it's one. But one, really for a fraction, is the same as zero.

One other student, confident to cancel terms in Item 8 because they were "the same", found Item 16 a greater challenge because,

> S3: …there aren't any like terms, because, like a squared. Oh, you could. You could put 10 over 4a. Can you? I think… I don't know.

The range of ability estimates of students in the Average-ability group was 1.87 logits (-0.94 to 0.93 logits), consequently including students of ability either side of the overall average ability estimate of -0.64 logits. These are the students who have been shown to have the most to say, and whose struggle to understand was often made evident by the way in which they framed their explanations. Students in the Average-ability group felt able to address all or most of the items in the survey, and to respond in the interview to most items, although rarely were they able to make explicit any connections they might perceive between the items.

The examples quoted above contain linguistic features that serve to indicate the uncertainty of the students as they tried to clarify their thinking. The explanations are, at a level of procedural simplicity, adequate, but they are also mathematically impoverished, despite the use of mathematical terminology such as *numerator* and *denominator*.

In the first example, the student (S1) used words that imply his pointing to terms in the expression under discussion ("It means there's one $b$ there…"). The phrases describe a mathematical procedure in terms of a physical act ("so you can scratch them both out") that is *like* dividing. This not understood by the student *as* dividing. The distinction is important: the student apprehends the mathematical idea, at an intuitive level. His conceptual understanding remains tied to the literal interpretation of the metaphor.

In the second example, the student (S2) acknowledges that cancelling is about division, but the phrase "taking both fours out" is a metaphor for removing terms, and hence leaving no fours. The absence of that which can be counted is denoted by zero, but the student also knows that in fractions the removal of terms is denoted by one. Hence the confusion: the language reflects a cognitive struggle to reconcile conflicting ideas as to whether the result of cancelling is to be represented by zero or one. Had the student used words describing a division process, the focus would have been on the mathematics, not on a physical act resulting in the absence of terms.

The description of terms in the expression as being the *same* is interpreted literally by students as meaning the terms have to *look* the same, instead of *being* the same (equivalent) mathematically. The first interpretation requires only the perception of a superficial appearance of sameness, the second requires abstraction about the mathematical relationship between terms that might not look the same but are mathematically identical (e.g., *ab* or *ba*).

This important step in understanding was clearly expressed by the student (S3) stating "…there aren't any like terms, because, like *a* squared". Although the student recognised that $a^2$ was divisible by $a$, she could not see any mathematical relationship (that of the common factors, that do, indeed look the same) between four and ten. The hedging by this same student was also indicative of her cognitive confusion and doubt.

The removal of terms by cancelling was also described by students as "getting rid of" unwanted terms. Two examples follow: the first from a student (S4) at the top of the Average-ability group (ability estimate 0.93 logits); the other from the second most successful student (S5, ability estimate of 3.6 logits).

> S4: Cancelling can be… You can get rid of … You can get rid of something because you don't need it, to make the expression, the answer easier, so because

> cancelling is like… If there are two similar things you can cancel them both out, but if they are different you can't cancel them, but if they are the same you can.
>
> S5: You look for similarities… You can, in question eight [Item 8, $4ab/4a$]… that you can cancel out. So it would be four. You can get rid of the four and the $b$.

The first example (S4) contains elements of vagueness: "You can get rid of something […]". The statement begins as an apparent generalisation, but is, instead, a vague statement, the vagueness betrayed by the use of *similar things*. The statement subsequently becomes tautological: "…cancelling is like… You can cancel them both out …" This example also illustrates that, like many other students in the interviews, the student interprets the instruction *to simplify* an expression as to make *it easier* (or, *simpler*, as another student expressed it). This is an example of the confusion that arises when students do not make the semantic transition from the meaning of words in everyday language to their constrained meaning in a literate register.

The second example (S5) illustrates the struggle of this most able student, who could cope successfully with Item 8, but could not do so with Item 16 [$2a^2 \times 5a/4$]. He tried to generalise, which, up to this point, he had been able to do, but could not extend the concept of similarity beyond terms that looked alike to terms (as factors, embedded in non-equivalent terms) that were of identical mathematical value.

In Item 16, students focused on cancelling the letters that looked the same, but often did not deal with the possible division by two of both the four and the two in the expression. Other students however, focused on the numbers, before multiplying the terms of the expression, but missed the connection between $a$ and $a^2$. These responses were typical of those in the interviews, and indicate the reasons for the high difficulty estimate for Item 16 (3.17 logits), but not for Item 8 (-0.64 logits). In dealing with expressions such as these, students operated only on the visual cues, and appeared not to have a firm idea of the mathematical reasons for the procedure.

It must be conceded that, because the items were written, students resorted to using iconic support by pointing to items, and relying on appearance of the printed expressions in front of them. By pointing to items to support their explanations, students demonstrated their inability, or saw little need, to articulate any mathematical understanding beyond learnt procedures, and these procedures were expressed in a non-mathematical register, or as an appeal to authority. One student (ability estimate 0.29 logits), when asked why her procedure worked, replied, interrogatively, "Because, um, [long pause]…it just does?"

*Cancelling* is a term that is an acceptable part of the mathematical lexicon, but the meaning is often conveyed to students by metaphor or analogy. The evidence suggests that only a limited set of examples is used to illustrate the metaphor, and that continued use of the associated informal language fails to engage many students with the underlying mathematics.

## LIKE AND UNLIKE TERMS

One of the first concepts that students meet explicitly when beginning formal algebra, in the first year of secondary schooling, is that of *like* and *unlike* terms (PAS 4.3, Board of Studies NSW, 2003 p.85). The concept has been implicit in much of primary school mathematics, but it is not until students encounter formal algebra that this idea is addressed explicitly. The generalised concept – of mathematical operations only able to be carried out with like mathematical objects (like quantities, of like value, like size, etc.) – is best illustrated by the use of non-specific symbols, usually letters. At this point in their algebra development, the use of letters to represent some generalised arithmetic concept is also new, and difficult for students. Many students, particularly those in the Low-ability group, could not articulate the point of using letters, stating variously that a letter was a "thingummy", "it puts a complication in", that "letters stand for something", and that letters stand for a "number".

However, students from all three ability groups used correctly and meaningfully the phrase *like terms* when dealing with items in interview Set 1 (Items 1, 2, 5 and 6), although students in the Average-ability group often seemed uncertain about the addition or subtraction of the terms.

> S6: I know that you can only add and subtract like terms… First thing I see is, I get the like terms together so 3*m* plus 2*m* is 5*m*. So then I use the eight minus five, which is obviously three… But sometimes I get confused if I put a plus three or a minus three. But I'd put a plus three. Yeah, I'd put a plus three, so 5*m* plus three …

Uncertainty is conveyed by the frequent use of *I*, the frank admission of confusion over signs, and the hedged general rule expressed as, "*I* know that *you* can only add and subtract like terms…" [my italics]. The use of the personal *I know* indicates that the student is not as committed to the succeeding generalisation as might be expected. Nor is the student confident in his final answer, correct though it is, as indicated by the use of *but,* and *yeah*.

One student (S7, High-ability) thought that *like terms* are indicated by a numeral in front of a letter, not by the letters themselves. Her uncertainty is clear in the opening phrases.

> S7: […] Question 5 [$2ab + 3b + ab$] would be a bit, yeah, harder, because I'd be unsure whether *ab* and 2*ab* are like terms or not.
>
> I: Why would you be unsure?
>
> S7: Well, *ab* is just… it doesn't have a number next to it, so it wouldn't be… It's the same with two [Item 2: $5p - p + 1$]. I wouldn't be sure whether *p* and 5*p* are like terms because they don't [both] have numbers next to them.

Student ability was determined from modelling survey responses on an equal interval scale of the targeted latent trait. Thus, able students would be those who gave more correct responses to more difficult items than did students of Average- or Low-ability. An inference to be made, when students express confusion over the mathematical meanings of relatively easy items (as compared to their ability estimates from the model), is that they are practised in certain procedures, but without being able to attribute any real meaning to them (see, e.g., Frid, 1993).

One student (S8) interviewed, found the use of a term such as *ab* confusing. The student's response to Item 5 in the survey was correct [$2ab + 3b + ab = 3ab + 3b$] but in the interview, he stated vaguely (indicated by the use of *it*) that:

> S8: So it'd be 2*ab* plus 3*b* would be 5*ab* squared plus *ab* is 5*a* squared, *b* cubed.

Other responses to this item (e.g., $6a^2b^3$, $2a^2b + 3b$, $2ab^2 + 3b$, $6ab^2$) indicated that it caused confusions amongst several participants (see Error Analysis, Chapter 5, Table 5.4).

Some students over-generalised rules for adding and subtracting terms:

> S9: Um … only like terms can be added or subtracted or timesed [*sic*] or divided by each other.

This misconception became most apparent when students, particularly those in the Low- and Average-ability groups, who had dealt successfully with Item 4 [*4* × *5b*] and Item 9 [*2ab* × *a*], found Item 13 [4*r* × 5*t* × 3] difficult. Fifteen students (out of the 25 interviewed in these two groups) were uncertain as to what to do with the different letters in Item 13; and, possibly, the fact that three terms were to be multiplied rather than two.

> S10: […] I don't think I'd be quite sure what to do with *r* and *t,* whether it'd be 4 times 5? So that's it's like 20*rt* times 3, which would end up being 60*r* to the power of 3*t*. To the power of 3. I wouldn't be sure if that was right or not.
>
> I: Why not? […]
>
> S10: Oh, I guess it's because they don't go to the third thing.

The referent for the vague *they* was unclear. The pronoun might be supposed to refer to textbook examples that the student had encountered.

One other student (S11) did correct her initial response, although she remained uncertain, as demonstrated by the frequent use of *I*, pauses ("um"), the hedge, "I think", and, after a pause giving her more time to think, the disclamatory "I dunno":

> S11: Um, that I multiply them all by three and then I add 12*r* and 15*t*. Just put 12*r* plus 15*t*
>
> I: Why are you adding?
>
> S11: Because they are not like terms again. Oh, no, no it's not. Um, OK… That's, that'd be 20*rt*, um, equals 60*rt*, I think. … I dunno.

These confusions were not so explicitly stated by students from the Low-ability group, who avoided responding to Item 13, or who tended to put all the Item 13 letters together, regardless of whether they were adding or subtracting or multiplying terms. The simple strategy of *putting all the letters together* resulted in a correct answer when multiplication of terms was required, but when members of this ability group were to add or subtract terms they conjoined the terms. The strategy could also be a reason for the survey items involving simple multiplication comprising the least difficult category of items.

The expression *putting like terms together* is a metaphor that conveys different, often inappropriate, concepts to students when interpreted literally. The metaphor can be understood literally, in arithmetic terms, such as *putting* two (counters) and three (counters) *together* (in one group) to give the unique result of five (counters). When students generalise the arithmetic idea to an algebraic expression, for example, $x + 2$, the result is a closed response of $2x$. The two terms have been *put together* in a literal interpretation of the metaphor that harks back to very early childhood experiences of literally putting one set of counters together with another, and counting the resultant members of a new set.

Students who are unfamiliar with open arithmetic expressions (e.g., $8 + 5 = 14 - 1$) as being valid equivalents are not helped by the metaphor, which does not make explicit the relational role of the arithmetic operators. Students who understand that terms in algebraic expressions can be related by symbols hitherto understood solely as instructions to act, accept the metaphor. However, the interview data suggest that the High-ability students also used terminology (e.g., *add, subtract*) that made explicit the mathematics involved in relating terms in an expression. Low-ability students seemed unable to rephrase informal language using a mathematical (literate) register.

## CONJOINING OF TERMS

The tendency of students to conjoin terms appears to be influenced by two limited conceptions: that of the role of brackets, and that of the addition or subtraction of (like) terms. Both of these concepts have been discussed in preceding sub-sections. In many cases, when explaining how they dealt with examples such as those in Set 1 and Set 3, students spoke about *putting together* like terms. However, students in the Low-ability group tended to put terms together by conjoining all the terms. Having identified and isolated like terms in Item 1 [$3m + 8 + 2m – 5$], Item 2 [$5p – p + 1$], Item 5 [$2ab + 3b + ab$], and Item 6 [$5a – 2b + 3a + 3b$], by circling them, or by rearranging the expression, or simply acting sequentially on each, students in the Low-ability group *put them together* in a different way (by conjoining them) to students in the Average- and High-ability groups.

Students in the two higher ability groups did not tend to conjoin terms in these items, with the exception of Item 5. The presence of two multiplied pronumerals ($ab$) in the terms confused some students in these groups, and seemed to prompt some to multiply the terms. The response quoted below indicates that the student (S8, quoted earlier) has only a vague idea of what the expression means. Vagueness is indicated by the use of the pronoun *it*. Without being acquainted with the item being discussed, one would have little idea of the intent of the student.

> S8: …Because it's $ab$ and just a normal $b$. It's the same letter but it's got another $a$ in it, so it'd be $2ab$ plus $3b$ would be 5 $ab$ squared plus $ab$ is $5a$ squared $b$ cubed.

Additional uncertainty is indicated by the conditional construction of *it'd*. Such a construction indicates conjecture on the part of the student such as, "…if I am right in my thinking, the answer would/could be…"

Students of Average-ability tended to conjoin terms more frequently when dealing with items in Set 3 (those with brackets, Item 18 $[2(x + 4) + 3(x - 1)]$, Item 19 $[2(x + 5) - 8]$, Item 7 $[(a - b) + b]$ and Item 11 $[8p - 2(p + 5)]$); and particularly those in Set 8, the Semi-literal items (Item 20 [Multiply $x + 5$ by 4], Item 21 [Add 4 on to $n + 5$], Item 22 [Add 3 on to $4n$], Item 25 [Take $n$ away from $3n + 1$], Item 26 [If $p + q = 5$, then $p + q + r =$…?], and Item 40 [If $e + f = 8$, then $e + f - g =$ …?]. Table 7.1 below records the numbers of students in each of the ability groups who conjoined terms in their responses to the items listed. The percentage for each group is also recorded.

*Table 7.1: Numbers of students who incorrectly conjoined terms in responses to items, by ability group and item number. The items are arranged in sets as presented in interviews*

| Item | Interview set | Rasch difficulty | Numbers and percentage of students conjoining terms in each ability group | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Low | | Average | | High | |
| | | | Numbers | Percentage | Numbers | Percentage | Numbers | Percentage |
| 2 | Set 1 | -2.6 | 29 | 28 | 2 | 3 | 0 | 0 |
| 1 | | -2.53 | 24 | 1 | 1 | 0 | 0 | 0 |
| 5 | | -1.98 | 24 | 24 | 4 | 6 | 3 | 6 |
| 6 | | 0.33 | 19 | 19 | 0 | 0 | 0 | 0 |
| 19 | Set 7 | -0.27 | 20 | 20 | 1 | 1 | 0 | 0 |
| 18 | | 0.13 | 16 | 16 | 2 | 3 | 0 | 0 |
| 7 | | 1.38 | 11 | 11 | 1 | 1 | 1 | 2 |
| 11 | | 2.28 | 19 | 19 | 11 | 16 | 8 | 15 |
| 21 | Set 8 | -0.58 | 47 | 46 | 15 | 22 | 3 | 6 |
| 22 | | -0.34 | 69 | 68 | 20 | 29 | 1 | 2 |
| 26 | | 0.07 | 23 | 23 | 5 | 7 | 1 | 2 |
| 25 | | 0.2 | 57 | 56 | 20 | 29 | 4 | 8 |
| 40 | | 0.2 | 4 | 4 | 1 | 1 | 0 | 0 |
| 20 | | 0.46 | 42 | 41 | 20 | 29 | 3 | 6 |
| Numbers of students in each ability group | | | $n = 102$ | | $n = 69$ | | $n = 52$ | |

Item 11 [$8p - 2(p + 50]$, in particular, prompted some students in the Average- and High-ability groups to conjoin terms. This may be because they failed to take account of the fact that the item indicated a difference between $8p$ and $2(p + 5)$ rather than a multiplicative relationship between the terms, and so they multiplied throughout – a case of a stimulus causing an automatic response: that when there are brackets in an expression the procedure is to multiply what is "inside" by what is "outside" (discussed in an earlier part of this section).

This short, procedural, informally-phrased, and well-remembered mantra also caused problems with Item 7 [$(a - b) + b$], where some students simply multiplied $(a - b)$ by $b$, because the $b$ was "outside" the brackets. This procedure might explain errors such as $ab - b^2$, or $ab^2$ that were common responses to Item 7. The latter response represents a conjoining of terms, the students paying no attention to the structure of the expression, but instead tending to focus on the single feature of brackets. Hence, a conjunction of two misconceptions – the need to *do something* with the terms inside the brackets, before multiplying by $b$, because it is outside the brackets – that leads to the incorrect response.

> S12: You just put the *b* on the other side and add it to the other side.
>
> I: What do you mean put the *b* on the other side?
>
> S12: Like, um, you times it by the *b* and you times it by *a,* instead of doing it the other way around.
>
> I: Why am I multiplying?
>
> S12: Because… It's confusing there, so you plus *b*… So there's always a times there [in front] where the bracket is.

In this example, the student (S12) is confident in her initial description of what she would do, as a general procedure. This is indicated by her use of the pronoun *you*. However, the hedge *just*, implies that some doubt is there. When the student is pushed by further questions, the hesitations (*um*) and silent pauses become more frequent, as does the use of *it,* and indicative words such as *there*. The student apparently can read the expression, but what is seen appears to conflict with what she expects from previous experiences with examples using brackets.

The conjoining of terms by students in the Average-ability group became much more frequent when they were required to answer items in Set 8 (Semi-literal items: Items 20, 21, 22, 25, 26 and 40; Table 7.1). These items required students to translate from words to mathematical symbols, showing an awareness of appropriate syntax and requiring an ability to deal with deliberate ambiguity in some of the items. Students of High-ability did not tend to make errors by conjoining terms.

There was a marked increase in the numbers of conjoined-term errors as students in the Low- and Average-ability groups responded to the items in Set 8, when compared with the numbers of similar errors made with items in Sets 1 and 3. One possible explanation for this is that items in Sets 1 and 3 were typical textbook examples and students could respond to them by carrying out well-rehearsed procedures, procedures that had trained them not to put together terms by conjoining them. Faced with an unfamiliar context, students with little understanding of the mathematical relationships conveyed by arithmetic operators in an algebraic context provided incorrect responses. (A considerable number of Low-ability students did not attempt Item 40, hence the low frequencies of responses, when frequencies similar to that for Item 26 could have been expected.)

One of the main differences between the ability groups, when they are compared on items to which all could be expected to respond successfully, was that students in the Low-ability group tended to conjoin terms that contained the operators plus and minus. Students in the Average-ability group seem to have learnt not to do so when simplifying expressions presented as a string of symbols (as in Sets 1 and 7), although they need attach no mathematical meaning to the expressions and the manipulation of those expressions. The increased frequency of conjoined terms when students in the Average-ability group were asked questions that had a similar mathematical meaning to those items in Sets 1 and 7, but which were presented in an apparently unfamiliar interpretive context (Set 8), indicated that this was most likely the case.

Students in the High-ability group rarely conjoined terms inappropriately, although a slight increase in numbers did occur for items in Set 8. This suggests that although students in all groups spoke about “putting terms together” their statements carried different meanings.

To *put terms together* is an example of metaphor, in colloquial language, used to assist novices to understand the mathematics to which they have been introduced, and which becomes well-understood shorthand for experts. In natural language, *putting things together* carries implications of physical *joining*. When one puts things together in a non-mathematical context, such as putting an outfit together, one adds several clothing items to a collection that is then viewed as a whole – an outfit. Attention is not paid to the manner of joining, but to the net effect.

In contrast, when mathematical terms are put together, it is the manner in which they are connected that becomes the focus of attention, and which is also necessary in establishing meaning for the new expression. To add to the confusion of metaphor and reality, mathematical putting together might involve some physical act – moving terms around in an expression being manipulated (the use of a physical metaphor again).

The inappropriate conjoining of terms inhibits mathematical understanding that goes beyond students' failure to deal with simple expressions involving the addition and subtraction of terms. The manipulation of expressions with brackets also depends on students being able to accept unclosed expressions as mathematical entities.

If the informal language of the metaphor is used persistently and does not evolve into appropriate mathematical terminology (e.g., *add the [like] terms*, *subtract the [like] terms*) that carries the meaning along with the procedural instructions, students are left to operate only with a literal interpretation of the metaphor, and consequently, conjoin terms.

The use of the phrase *put terms together* appeared to be common in the sample of students interviewed, regardless of ability. The successful responses to the survey items by students in the High-ability group suggest that these students do not attach a literal meaning to the phrase. The lack of success by students in the Low-ability group suggests that they are more likely to do so. High- ability students also have a grasp of the syntax of natural language, and of the ways in which this is translated into the correct mathematical syntax.

## GENERALISATIONS OF LOW- AND HIGH-ABILITY STUDENTS

From the preceding examples, expressions used by students in each of the ability groups often sounded quite similar when taken at face value, without attention being paid to linguistic elements. In the analysis of linguistic features (Chapter 6), statements made by students interviewed were considered as *general* if there was no reference made to specific items in the set being discussed, except, perhaps, as an illustration.

Results indicated that both the Low- and High-ability groups tended to make general statements with approximately the same frequency: 55% of all statements made by students in the High-ability group were general; 51% of all statements made by students in the Low-ability group were general. Students in the Average-ability group tended not to make general statements: only 32% of all statements made by students in this group were general (see Chapter 6, Figure 6.8, Tables 6.11, 6.12).

Central to this study is the notion that an ability to generalise is linked to success (ability) in algebra. The reason, therefore, for the differences in success on the algebra survey by Low- and High-ability students lies in differences in the types of generalisation made by the students, and linguistic structures of these statements. Two types of general statements are discussed, using examples from the interviews, descriptive statements, and rule-like statements. When the linguistic features are examined, some statements, although apparently general, might better be understood as vague utterances rather than as concise statements of principle.

### Descriptive statements

Descriptive generalisations focused on the appearance of items in a set. The Low-ability group made more descriptive general statements than the students in the High-ability group. In response to the question directing a student (S13) in the Low-ability group to all items in Set 4 (fractions):

> I: Can you tell me anything common to all those expressions, and anything different you might see?
>
> S13: They're all fractions [pause]. Yeah, they're all fractions.
>
> I: Do you notice anything different about them?

> S13: That one's addition, subtraction, division and two times …[The student had no more to offer.]

The generalisation merely describes the appearance of the items, and names the arithmetic operators that appear in the items. The use of indicative words (*that one's…*) is vague as the student lists attributes of an individual item as if they applied to all items in the set. The student appears to have focused on common features of how the items *look*, rather than the mathematical *meanings*.

In terms of the linguistic model being proposed, such a response need not be unexpected. The set of items with an average difficulty estimate of 1.046 logits was more difficult than the ability estimate for the student of -1.46 logits. Therefore, the student would most likely struggle to make mathematical sense of the items, and thus would articulate their insecure understanding in vague terms.

Students in the Low-ability group were not the only ones to make such superficial descriptive generalisations, as the following exchange with a High-ability group student (S14, High-7) concerning the items in Set 2 (Multiplication of terms in expressions) demonstrates:

> I: Could you tell me anything common about all those questions?
>
> S14: They're all positive numbers. They're not too hard.
>
> [The student's voice indicated that he had nothing further to add]

What is notable is the student's lack of attention to the mathematics common to the items, despite his overall success in the survey. This was the least difficult set of items (average difficulty -1.56 logits), and well below the student's ability estimate (1.65 logits). The student should have been able to see, and to make explicit, the mathematical similarities. Despite this student's success on the survey items, he found communicating his mathematical ideas difficult, often resorting to reasoning such as items were "easy" because "I can do them"; or, in moments of struggle, admitting that "I forget most of this stuff".

Descriptive generalisations that focus on the appearance of items, allow the student to avoid engagement with the mathematics. These types of responses might be considered to be less conceptually developed than responses from students who provided answers to each individual item in a set of items, but without attempting to articulate some generalisation. These latter types of responses were typical of students in the Average-ability group.

### Rule-like generalisations

Rule-like generalisations were those statements made by students articulating a principle or procedure, although these statements need not necessarily have been correct. For example, statements such as brackets mean to “do it first”, “you collect like terms”, “you cancel” are rule-like. The understanding by students of these concepts has been discussed in detail in preceding sub-sections.

Students in the Low- and High-ability groups tended to make these statements more frequently than students in the Average-ability group, but differed greatly in their success on the survey. It is useful to compare the statements, and the use to which they were put, by students in the High- and Low-ability groups. Students often asserted these rule-like statements with confidence, but when they were probed for greater understanding, it became clear that students in the Low-ability group were using the mathematical terminology without deep understanding. Their elaborating statements were of a low modality, vague, and lacking in a mathematical register.

To investigate the general rules stated by students in the Low- and High-ability groups, Sets 1, 3, and 4 were examined in detail. These were chosen because Set 1 (addition and subtraction of terms) was the least difficult set (average difficulty estimate -1.66 logits), and hence accessible to most of the students interviewed. Set 3 (expressions with brackets) with an average difficulty of 0.425 logits was responded to with a degree of confidence by most of the students interviewed. Set 4 was one of the more difficult sets, consisting of expressions involving fractions and of average difficulty 1.046 logits.

#### *Set 1*

The initial responses to Set 1 by five out of the seven Low-ability students in response to the question, “What goes on in your head when you see items like these?” were to state that they “thought of how [they] got the answer”. One student said that she “thought of what not to do”. Two students mentioned like terms; only one student defined a like term by suggesting that *m* was one (Item 1), when asked what was understood by the phrase *a like term*. Two students used the term *pronumeral*; most referred to *letters*. Four students said they had no idea of the meaning of the letters; one only suggested that they stood for numbers. One student (ability -2.86 logits) stated that the letters were a "thingummy" and elaborated by stating:

> S15: I think when it's sort of getting a letter [unclear] it stands for something else. Like you are adding different things up together, puts a complication in.

The struggle by the student (S15) to make her understanding explicit is apparent when one considers the low modality in the frequent use of phrases such as "I think…", "sort of…" and "like", and the absence of a mathematical register in the use of the vague "something else" and "different things" alluding to the meaning of the letters. Vagueness is also indicated by the use of the third person pronoun *it*.

When describing what they did to manipulate the expressions in Set 1 (addition and subtraction), two students in the Low-ability group, from the same school, stated that they would "put a circle around" like terms, because that was how they were taught. Others stated that they had to "fit together" the like terms (see discussion on the conjoining of terms in a previous sub-section).

Four students in the High-ability group commented that the items in Set 1 and Set 2 were "easy" or "simple" – because they could "do them". Four students used the phrase *like terms* and could identify like terms correctly in all items and explain, by illustration, what they meant. All students in this ability group could explain that the letters (some used the term "pronumeral") stood for an unknown number.

### *Set 3*

Set 3 consisted of three expressions with brackets, to be manipulated. The average difficulty estimate for this set (0.11 logits) was higher than the ability of the most able student in the Low-ability group (ability estimate -1.28 logits). The responses of the students in this group tended to be vague, or expressed in very informal language. Two stated that brackets "told you what to do"; three stated that brackets indicated, "to do it first". Others expressed the idea of brackets grouping numbers as "to make them stand out", "to separate like numbers", to "single out [numbers]" or they were for "grouping numbers inside instead of writing them out".

The idea that one had to "do first" what was in the brackets may be the root cause of the students in this group tending to conjoin terms inappropriately (see preceding sub-section).

One student (S16) began to describe, confidently, rule-like instructions for dealing with items in Set 3. The statement is general in type: he has used the impersonal *you*; he does not complete manipulations of specific items, but indicates what should be done; the instructions are of a reasonably high modality, with few pauses for thinking, and only two uses of the hedge *just*; he uses some logical connectives in the statement "and if you can group the like terms…"; and, he sees similarities of procedure for all items in the set, "… you just do the same as you would for the first one…"

> S16: You just… Back with [item] 18 [$2(x + 4) + 3(x -1)$] … you just do the same as you would for the first one, and you put a plus, and you do the same with the second set, and if you can you can group the like terms, and put them together, shortening it. And with [item] 19 [$2(x + 5) - 8$] … you just write down the answers that you get with the two times $x$ and the two times five, and plus the eight because it's outside the brackets and then you… Ten plus the eight together.

However, the student is at a loss to explain the meaning of the brackets in one particular item. The modality of his statement drops, he uses vague third person pronouns and he cannot find useful mathematical terminology to explain the ideas.

> I: What is the meaning of two, then brackets, $x$ plus five [$2(x + 5)$]?
>
> S16: Um, sort of like… I don't know. Um… just separate like numbers. They're not part of that sum there. That's just a different area. You have to work out that.

The change in the modality of the response suggests that the student's ability not to close on each item as he describes the procedure in the first statement might be because of his lack of confidence in being able to apply the procedure he remembers and recites. The similarities of his instructions as to what might be done when manipulating an expression to that of a generalised procedure becomes, instead, a rather vague statement, when read, or heard, in conjunction with the subsequent statement.

A listener would infer from the tenor of the second statement that the student was not secure in his understanding, and would be likely to reassess the quality of understanding suggested by his previous statement. This is a context which requires a difficult and subtle pragmatic judgement to be made about the quality of the student's understanding.

Most students in the High-ability group explained that brackets indicated a procedure. Nine of the ten students used the terms *multiply* or *expand* suitable to a mathematical register, unlike the low-level terminology in the example above. Four also suggested that simplification followed the expansion; "shortening it" was the phrase used by one student. Three students stated that brackets meant to "do it first". Two students used the phrase that one "took the brackets away" or "got rid of" the brackets when they expanded the expression.

Instructions, such as "do it first" are too simplistic to be mathematically useful; the metaphors of "taking brackets way", or "removing brackets" confuse novice students' mathematical ideas with the visual result of carrying out particular mathematical process. Successful students seem to be able to interpret the metaphor in the appropriate mathematical way; unsuccessful students seem to rely on a literal interpretation.

### *Set 4*

Set 4 was more difficult [average difficulty estimate 1.046 logits] and contained items that were challenging to the students in the High-ability group. The students in the Low-ability group struggled with the mathematics in all items, except Item 8 [$4ab/4b$]. Only one of the seven students used the term *cancelling*, instead describing their simplification of the expression as "crossing out the pairs', "getting rid of" or "cutting out" the same letters.

One student (S17) explained her thinking about items in Set 4 in the following generalisation, which, on close inspection, is of a low modality and is expressed in the language of an informal register:

> S17: When it's over something means divided by, like. So you'd have to divide them, it's a bit harder than the others. Yeah, just divide it as much as I can.
>
> I: Can you tell me more about those?
>
> [Student looks for a question which she thinks she can do]

> S17: Um … there's division ones like top with the bottom and do the same thing to the top or the bottom of the other one. Like, if you do something to the top you have to do the same thing to the bottom.
>
> I: Can you give me some details about that?
>
> S17: Like, with [item ] 12 [$2/a \times 3/b$] you sort of times two and *b* and three and *a* and you have to do the same thing to the top or the bottom as you did to the other side.

There are scattered, confused, generalised half-concepts, such as "when it's over something, means divided by…", and "what you do to the top or bottom, you do to the other side". What is to be done is not stated explicitly. Nor is it clear what is to result from "divide it as much as I can".

No students in the Low-ability group used the term *denominator* or *numerator*, preferring to use *top* and *bottom*, terms that describe the visual arrangement of terms in a fraction, and which are not part of a mathematical register.

Students in the High-ability group also used informal language when responding to this set of items, particularly so with the students in the ability range of 1.27 to 1.65 (High-7). Five students used the term *denominator,* one referred to *equivalent fractions*, without further elaboration. One student avoided answering questions about the two items requiring addition or subtraction of fractions.

Three High-ability students used the term *cancelling*. The meaning of this term was usually explained informally (see previous sub-sections). Four students explained that cancelling meant division. Two suggested that one number divided by another was zero, and two that the result was one.

The students seem to have confused two concepts. The first is that of division involving exponents – subtracting powers. The second confounds the absence of a coefficient in front of a letter (the letter itself signifying one of) with the fact that the letters are crossed out (cancelled), and hence are no longer present in the expression, signified by zero (see the earlier discussion).

Generality and vagueness exist as two extremes of a spectrum of principled statements about mathematical concepts. Both response types lack reference to specific examples; both might use the impersonal *you*; and both might describe in commonly acceptable terminology mathematical procedures or rules. Thus, the students' interviews were classified as *generalisations*. These were distinguished as being either descriptive, or rule-based. On this basis, both the High-ability and Low-ability groups of students made approximately the same proportion of general statements.

The feature that discriminated, and so explained the success of one group and the lack of success by the other was the quality of the statements. Low-ability students often made vague statements, statements of low modality and lacking in the mathematical register, whist using common phrases that were also used by the students in the High-ability group.

General statements by students in the High-ability group tended to be of a higher modality, and to adopt, in part, the mathematical register. However, when challenged by mathematical concepts, students in the High-ability group also made vague statements. Most general statements made by these students were procedural explanations, rather than conceptual.

Generality expressed in statements of high modality and in the mathematical register are more likely to be made by students confident in their mathematical understanding. Vague statements are likely to be made when mathematical ideas are poorly understood.

## CONCLUSION

Explanations by students in interviews suggested that there were several aspects of algebra that appeared to be commonly understood. These concepts were expressed in similar language. Responses of students from different ability groups to survey items indicated that the groups held different mathematical understandings that influenced their success. Students from across the spectrum of abilities used a mix of informal and formal language to describe many aspects of algebra, often using closely similar phrases.

Research question 5 asked whether linguistic features of students' explanations could be associated with the development of their algebra understanding. Demonstrated by the examples in this chapter, it is evident that students' understandings are quite different, and conceptually richer, as ability levels increase (as shown by their success in dealing with the increased complexity of more difficult items on the survey). The students' conceptual development was reflected in their explanations when they were asked to elaborate on common terminology used by them*: brackets, like terms, cancelling*. Further evidence of students' confidence in their understanding was provided by analysing the intention behind the phrase *put like terms together* as demonstrated by survey responses, and expressed in interviews by students from each ability group.

Of particular note was the similarity of expression used by the High- and Low-ability groups when they made generalisations. However, the underlying concepts appeared to be quite different, and indicate possible reasons for the difference in success between these two groups.

Although students were confident to explain their mathematics in terms of stock classroom terminology (which was often colloquial and metaphoric), their elaborations of the understandings embodied in these simple epithets were less confident. This can be a challenge for the most able student, who has not had cause to reflect on and articulate the meanings behind phrases commonly used. However, linguistic structures of the explanations given by students from the three ability groups do indicate the extent of the students' cognitive struggles.

The less well-understood the ideas encapsulated in procedural instructions or informally-phrased rules, the more informal the students' explanations. Discourse tends to be of low modality, lacking the mathematical register. Insecurity of understanding is also marked by literal interpretations of metaphor and analogy. It is also apparent that the strength of metaphor, or analogy, lies in descriptions of mathematical (cognitive) acts on mathematical objects (concepts) as physical acts on visible, and actual symbols. This assists students to obtain correct solutions to familiar examples, but at a cost of explicating the mathematics.

Less successful students struggled to rephrase informal, metaphoric terminology to make explicit underlying mathematical concepts, even when able to respond successfully to the algebra items. The language used by more able students also decreased in modality and register, as they struggled to explain ideas that challenged them, and for which they had a restricted vocabulary.

From student responses to items in Set 8 (Semi-literal items that asked students to rewrite word statements in algebraic symbols in the survey), it was evident that students of Average-ability were experienced in a formalist approach to algebraic manipulations. Yet they had little idea of the meaning behind the signs and syntax. They had had little exposure to the language that would convey appropriate meaning.

High-ability students tended neither to conjoin terms inappropriately, nor to interpret the metaphors of *cancelling, crossing out, doing first* and *putting terms together* literally, thus indicating that they had, in most of the contexts used in the study, a richer understanding of the mathematics. However, they too found articulating their understanding difficult, offering only procedural explanations and generalisations, and no conceptual explanations.

The evidence of the examples discussed suggests that the ways in which students express their mathematical understanding are closely linked to the quality of that understanding. How well students understood the algebraic ideas was indicated by the modality, generality and register of their explanations. These data provide the basis for answering the final research question (6) as to how students' linguistic development might be aligned to stages of mathematical growth, generating a conceptual framework that shows development of students' thinking from the specifics of arithmetic to that of the general relationships of algebra. This is the subject of the following chapter.

# CHAPTER 8: A LANGUAGE-CONCEPTUAL MODEL OF ALGEBRA UNDERSTANDING

> *…Cognitive and linguistic aspects intervene simultaneously in the comprehension and in the use of different means of expression (Sfard, 2001, p.65)*

The preceding chapters provide evidence that the ways in which students use language indicates the quality of their understanding. This chapter generalises from this evidence to address the final research question by proposing a model that aligns algebraic understanding and linguistic behaviour to a conceptual framework.

Language that is hesitant, informal and unnecessarily prolix implies speakers' lack of confidence in the truth-value of their utterances. Alternatively, language that is of a literate register appropriate to the contest conveys the speaker's confidence in the subject matter. Hence, a developmental model of mathematical understanding and language can be formulated. In general, a developmental continuum connecting language and mathematical insight might be represented as:

| | | | |
|---|---|---|---|
| *Experience* | Novice | → | Expert |
| *Understanding* | Procedural | → | Conceptual |
| | Tacit Implicit | → | Explicit |
| *Mathematics* | Real world | → | Dissociated |
| *Language* | Vague | → | General |
| | Low modality | → | High modality |
| | Natural register | → | Mathematical register |
| *Interpretation of metaphor* | Literal | → | Mathematical |

This continuum is elaborated in the developmental model described in this chapter. The model maps the occurrence of particular linguistic features in the explanations of students of different abilities against particular features of mathematical understanding of symbolic algebra. The linguistic features become pointers to how well students know particular aspects of algebra. The model provides an objective basis for judgements made about students' understanding, based on their discursive practice.

The first two sections outline mathematical approaches of students operating at different conceptual levels and the features of the language used. A third section describes the organisation of the language-conceptual model; the mapping of the mathematical and language descriptors to the SOLO model.

## MATHEMATICAL APPROACHES

Mathematical approaches of students, revealed through interview conversations have been analysed and discussed in terms of ability groups indicated from the Rasch model. The approaches that typify the groups are elaborated in this section.

In the first years of secondary school, introductory topics in algebra include the use of letters to represent numbers, the translation of word sentences to algebraic symbols, the generalisation of number patterns, the manipulation of letters according to rules first encountered in arithmetic contexts, and the solution of linear equations. Students' success in algebra depends, therefore, on their making the conceptual shift from the use of particular numbers to a generalised symbol, and in perceiving relationships between numbers in terms of general principles rather than as particular instances.

This means that students need to understand the arithmetic symbols (or words) for addition, subtraction, multiplication and division not only as instructions to operate on a set of numbers, but also as describing relationships between numbers. The intent of the arithmetic symbol must be elicited from the context and so students are required to move between two different, but related, concepts. Gray and Tall (1991) referred to this conceptual flexibility as a *procept*.

More developed algebraic thinking involves the ability to perceive relationships between numbers and to generalise those relationships. This generalisation results in the use of symbols (conventionally, letters) other than particular numbers, and the manipulation of relational expressions as mathematical. Therefore, in terms of the SOLO model, the thinking required to carry out formal symbolic algebraic processes successfully occurs in the formal mode.

However, many introductory algebra examples can be dealt with arithmetically, particularly the solution of simple one- or two-step linear equations which have small positive integer solutions. Arithmetic thinking, dealing as it does with symbols representing abstractions from the real world, occurs in the concrete symbolic mode of the SOLO model. It is not until students meet examples that demand their attention to number relationships and they conceive of the idea that the obtaining of a unique, closed answer is not always the object of mathematical manipulation that they develop an orientation that might be described as algebraic thinking.

Although the target mode for algebra is the formal mode, earlier modes of thinking, such as the ikonic mode, might also be used by students to support their mathematical reasoning. This was most evident in responses from the Low-ability students, who seemed to use predominantly the ikonic mode to support their thinking. The items on which the students in the Low-ability group were most successful were the least complex mathematically, requiring a single step to simplify an expression, or a single positive whole number answer to an equation.

These students also appeared to use a literal interpretation of the metaphors (e.g., *remove/get rid of brackets*, i.e., rewrite an expression without brackets), often used by teachers to explain mathematical procedures. Many of these metaphors also described a mathematical act in terms of a physical act (e.g., *cancelling* as crossing out terms that were/looked the same). These students could carry out simple procedures relying only on the appearance of items, ignoring any mathematical meaning.

Students in the Average-ability group tended to respond to every item in each of the sets – providing simplifications of all expressions, and solutions to each equation, provided the equation could be solved where each step in the solution could be closed, or confirmed by a numerical answer. It would appear, therefore, that they had begun to connect arithmetic ideas with the more abstract algebra, but still relied on the real-world referents of numbers to confirm their algebraic thinking. This suggests that the type of thinking for the Average-ability group remains in the concrete symbolic mode. Some students in this group struggled to deal with arithmetic that involved numbers other than positive whole numbers (e.g., $68/x = 180$; $10y = 5$), and could not deal with the items that required dissociation from all numeric referents (e.g., $(x + y)^2$ , or $ax = 5$ ).

Students in the High-ability group struggled with the more abstract (literal) items in the sets of items, tending to employ appropriate procedures, but without considering the structure of the items. Hence, when attempting Item 16 [$2/a^2 \times 5a/4$]

they crossed out terms that looked alike, but not before they multiplied each term in the original expression to obtain $10a/4a^2$. No attempt was made to factorise, or explain their procedures in terms of factors. Many of these students needed prompts to conclude that $x = 5/a$ in Item 39 [$ax$=5]). The three most able students interviewed found the expansion of Item 10 $(x + y)^2$ difficult.

Many of the algebra examples used as items in this study could have been answered successfully by students operating in the concrete symbolic mode. It appears from the interview data that students regard numbers, particularly positive whole numbers, as concrete ideas, that are sufficient basis on which to verify their abstractions of algebraic manipulations. This was particularly apparent with students in the Average-ability group. Only those students who had internalised general relationships between numbers and arithmetic operations could deal successfully with the more difficult, literal items. These students were considered to be operating at early levels in the formal mode.

## LANGUAGE FEATURES

When understanding of algebra is considered in terms of linguistic features, arithmetic thinkers, operating in the concrete symbolic mode, are most likely to describe specific instances, often using the first person. Algebraic thinkers would tend to explain general principles in an impersonal, procedural genre, using the simple present tense. Students whose level of cognitive development is transitional between the arithmetic and algebraic, could be expected to use a mix of the personal and impersonal, and to use several specific examples as typifying all members of a set without articulating a general principle.

In the following, descriptors of language that characterise levels of understanding described in the model are discussed. The first sub-section considers the descriptors in the specific context of the study – developing algebra understanding. The second reframes the descriptors in general terms that might be used in a broadly-based model of mathematical understanding.

## Language descriptors in the context of developing algebra understanding

This study has found that students whose understanding of algebra is different to that of others, adopt different registers when speaking about their understanding. These registers are distinguished by the presence of particular linguistic features that together provide a set of language descriptors. In the following parts, these descriptors are outlined for each of the ability groups of students identified in the study.

### *Low-ability students*

The Low-ability group of students had little success in dealing with examples of algebraic expressions. When they did try to address algebra items, they did so by operating in early levels of the concrete symbolic mode, with considerable support from the ikonic mode. They seemed not to engage with the mathematics. With little information or insight to share: they therefore had little to say. Their conversations were brief, even laconic, using an everyday, natural register.

These students, whose algebra understanding was fragile, tended to focus on a single item in a set, and to explain in short statements what they, personally, would do. They tended to recite simple taught rules, or informally phrased procedures, as justification for their responses. They had difficulty in elaborating these procedural rules, or interpreting the underlying mathematics of the metaphors used to formulate the rules.

Many explanations given by these students consisted of vague statements of low modality, lacking in conventional mathematical terminology, and dominated by language of the natural register. When these students did attempt to generalise, they did so either with statements about their attitude towards the mathematics, or with vague statements such as “I think of what I would not do”. When general statements did focus on the mathematics these tended to be descriptions of the appearance of expressions, or simple procedural rules.

### *Average-ability students*

Students in the Average-ability group were those whose understanding was moving away from a strictly arithmetic focus. These students engaged with the algebra – at some level, had some information with which to work, who saw that connections could be made, who hypothesised. These were the students working in a conjecturing environment.

In order to organise and connect the disparate threads of their understanding these students had much to say as they talked to themselves, and to others. They tended to give very lengthy, indeed verbose, explanations describing step-by-step their (often personal) procedures. Given a number of mathematically similar items, they did not articulate any mathematical connection between the items, but instead gave a sequence of isolated, specific answers.

The register of their discourse varied from the informality of natural language to the formality of the mathematical register, depending on their security of understanding. The predominantly natural, colloquial language of their discourse at times adopted appropriate mathematical terminology, indicating a developing mathematical register. Metaphors used to describe mathematical procedures seemed to be partly understood in mathematical terms, although students were unable to articulate that understanding.

The range of abilities for this group of students resulted in a wide variation along the developmental language continuum described. Students who could deal successfully with items of difficulty less than their ability estimates used language of a higher modality, although it remained predominantly in the natural register, and focussed on specifics. As students attempted items of greater abstraction and complexity, their utterances increasingly consisted of frequent hedges and pauses, vague statements, and informal vocabulary.

### *High-ability students*

Students in the High-ability group were those whose thinking focused on relationships between numbers, and the general behaviour of numbers and number systems. This conceptualisation of algebraic ideas indicates that the High-ability students were operating in the early levels of the formal mode, or relational level of the second cycle of the concrete symbolic mode.

These students used generalisations to encapsulate complex ideas. Hence, they too had little to say, assuming that their concise explanations would be clearly understood, and sufficient. Their discourse assumed a more literate, mathematical register.

Explanations, although tending to be statements of procedural rules rather than of mathematical concepts, were phrased in the general, objective mathematical language of propositions. Metaphors were reinterpreted in terms of underlying mathematics, which students could articulate clearly and distinctly. As with the group of students of Average-ability, when faced with the items that were most difficult – literal and dissociated from any arithmetic referent – the modality of their language dropped, and statements become vague, characterised by the increased use of the third person pronoun *it*.

The broad linguistic characteristics of students' utterances can be summarised as follows: the better a student's understanding, the explanation tends to be more general, clearer and likely to be couched in terminology appropriate to a mathematical register. Given a new context, a set of new concepts with which they must grapple intellectually, students' thinking and their language returns to earlier, less sophisticated, less developed levels.

### Generalised language descriptors of a language-conceptual model

This second sub-section describes linguistic features of the model in terms that characterise levels of understanding in a broader mathematical context, and which reflect the cyclical nature of learning. The generalised descriptors form a developmental continuum, a spectrum of increasing linguistic sophistication and register. In this sub-section the language descriptors outlined above in the specific context of algebra are elaborated sequentially, from levels where students are challenged by mathematical ideas, to levels where they have a sound understanding of the concepts.

*Explanations by students struggling with new ideas* tend to focus on descriptions of specific examples. Use of the pronouns *I*, and *it*, imply a personal, rather than an objective perspective, and understanding that is vague – yet to become fully-formed. A predominant use of informal terminology often reflects a high degree of dependence on ikonic-mode support, and a literal interpretation of metaphors commonly used to ground mathematical concepts and procedures in contexts students already understand.

Sentences may be simple, with no accompanying logical connectives or subordinate clauses, reflecting still-to-be-developed relational thinking. Informal, epithetic rules based on single accessible or familiar examples are often offered as explanations. At the early stages of concept-development, students have little understanding to explicate. Hence, their utterances tend to be brief, and laconic.

*Students for whom some aspects of mathematics* are clear, and can be made explicit, while other aspects remain dimly apprehended at an intuitive level, operate in a conjecturing environment. These students try out newly-formed hypotheses, and reconceptualise previous understanding. They use many examples to elucidate a point, but are unable to make explicit any mathematical connections between the items. Hence, their discourse often consists of hedged, verbose, involved explanations, usually in the simple present tense. Their use of the general *you* indicates some security of understanding. Their reversion to the first person *I*, or the vague *it*, signifies a cognitive struggle with imperfectly understood concepts. They use many hedges and attribution shields as they struggle with the ideas.

*Students who understand and can apply particular mathematical concepts* explain their thinking with general statements that are propositional in form. The statements use the general *you* (instead of the more formal *one*) and are couched in the simple present tense. The statements often make no reference to any specific example; when they do so, examples are cited as representative of a class of concepts. Statements are coherent, of a high modality and use the appropriate mathematical register.

Linguistic features characterise particular registers. Changes in linguistic features of student discourse reflect development of those registers. Development proceeds from the almost exclusive use of natural language to the use of discipline-appropriate language. In parallel to linguistic development, algebraic understanding develops from arithmetic thinking, grounded in numeric (real) experiences and personal perspectives, to the consideration of abstracted and generalised mathematical relationships, dissociated from real referents. How these two aspects are brought together on a theoretical framework is discussed in the following section.

## ORGANISATION OF THE MODEL

The SOLO model has been used as the theoretical framework on which mathematical responses and linguistic features are organised in a developmental hierarchy. The model was chosen for two reasons: Collis' early suggestions of how language development fits the SOLO framework (1982) and the multi-modal nature of the framework.

One of the strengths of the SOLO framework is that it recognises learning occurs in different modes – the sensorimotor, the ikonic, the concrete symbolic and the formal. The sensori-motor and ikonic modes become available from birth and early childhood, but remain accessible. The learning of algebra, for which the target mode of operation is the formal mode, appears to rely on support from the early modes, particularly the ikonic mode, and later, the concrete-symbolic mode. Much of algebra relies on the recognition of visual clues in the notation, and the manipulation of patterns of symbols as mathematical objects. Understanding of the mathematical meaning of the notation, however, proceeds from the concrete-symbolic mode where arithmetic knowledge is drawn on, into the formal mode where abstract relationships are understood.

The language-conceptual model is presented schematically in Figure 7.1. Modes and levels of the SOLO model are summarized in the left-hand column of the diagram. The progression in linguistic register and algebraic conceptualisation is summarised as a descriptive, qualitative spectrum, aligned broadly with SOLO levels and modes. Features of algebraic conceptual growth and associated linguistic development are described as a continuum.

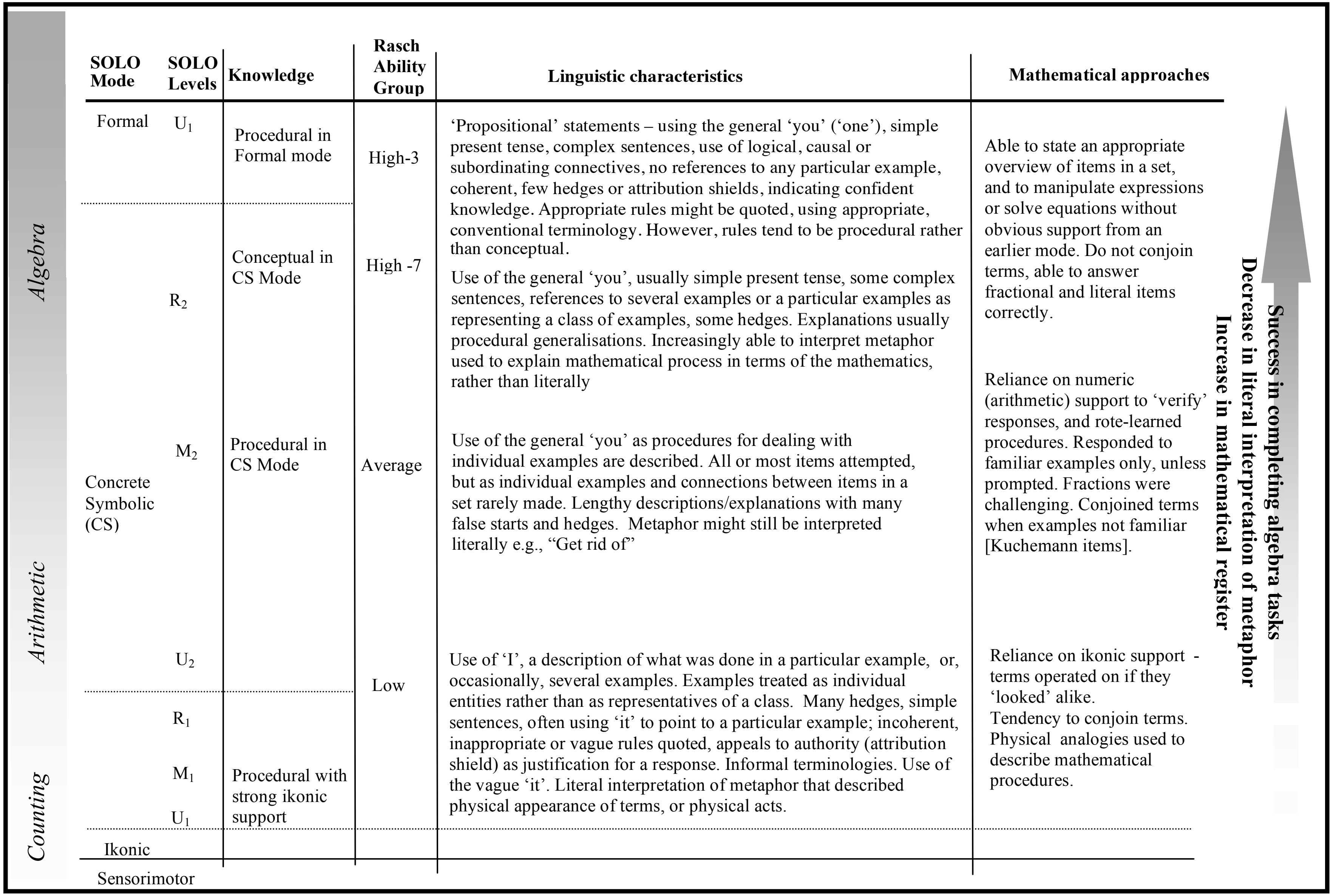
Linguistic characteristics have been approximately aligned with levels in the concrete symbolic (CS) mode of the SOLO model, based on that proposed by Collis (1982, summarised in Chapter 2). The levels in the concrete-symbolic mode form recurring cycles as learners acquire understanding of new concepts. The levels are

Unistructural, Multistructural and Relational (U, M, R) (described in detail, Chapter 2). Much of the algebra taught in the early years of secondary school, as exemplified by the survey items, seems to have been accessed by students thinking in this mode, as evidenced by the interviews.

This suggests that students develop formal algebraic techniques from contexts where the algebra can be supported by a concrete referent, such as numerically verifiable solutions. Few students responded in ways that could be inferred as their having dissociated mathematical ideas from a real (numeric) referent, as would be the case if they were thinking in the formal mode.

At higher levels, however, the language used by students seems to indicate that they are operating in the early formal mode. Collis (1982) suggested that this is the mode of cognitive development where learners are able to make meaningful, propositional statements about mathematical ideas that have no immediate real world referent, such as the items to which these students successfully responded.

The development of algebra from arithmetic understandings to abstract, formal symbol-manipulations is illustrated as a continuum on the far left of the model diagram (Figure 7.1). To the right of the diagram is described a continuum of success in the algebra tasks required of the students in the study, together with a continuum of linguistic changes that appear to accompany increased success in algebra.



| SOLO Mode | SOLO Levels | Knowledge | Rasch Ability Group | Linguistic characteristics | Mathematical approaches |
|---|---|---|---|---|---|
| Formal | $U_1$ | Procedural in Formal mode | High-3 | ‘Propositional’ statements – using the general ‘you’ (‘one’), simple present tense, complex sentences, use of logical, causal or subordinating connectives, no references to any particular example, coherent, few hedges or attribution shields, indicating confident knowledge. Appropriate rules might be quoted, using appropriate, conventional terminology. However, rules tend to be procedural rather than conceptual. | Able to state an appropriate overview of items in a set, and to manipulate expressions or solve equations without obvious support from an earlier mode. Do not conjoin terms, able to answer fractional and literal items correctly. |
|  | $R_2$ | Conceptual in CS Mode | High -7 | Use of the general ‘you’, usually simple present tense, some complex sentences, references to several examples or a particular examples as representing a class of examples, some hedges. Explanations usually procedural generalisations. Increasingly able to interpret metaphor used to explain mathematical process in terms of the mathematics, rather than literally |  |
| Concrete Symbolic (CS) | $M_2$ | Procedural in CS Mode | Average | Use of the general ‘you’ as procedures for dealing with individual examples are described. All or most items attempted, but as individual examples and connections between items in a set rarely made. Lengthy descriptions/explanations with many false starts and hedges. Metaphor might still be interpreted literally e.g., “Get rid of” | Reliance on numeric (arithmetic) support to ‘verify’ responses, and rote-learned procedures. Responded to familiar examples only, unless prompted. Fractions were challenging. Conjoined terms when examples not familiar [Kuchemann items]. |
|  | $U_2$ |  | Low | Use of ‘I’, a description of what was done in a particular example, or, occasionally, several examples. Examples treated as individual entities rather than as representatives of a class. Many hedges, simple sentences, often using ‘it’ to point to a particular example; incoherent, inappropriate or vague rules quoted, appeals to authority (attribution shield) as justification for a response. Informal terminologies. Use of the vague ‘it’. Literal interpretation of metaphor that described physical appearance of terms, or physical acts. | Reliance on ikonic support - terms operated on if they ‘looked’ alike.<br>Tendency to conjoin terms.<br>Physical analogies used to describe mathematical procedures. |
|  | $R_1$<br>$M_1$<br>$U_1$ | Procedural with strong ikonic support |  |  |  |
| Ikonic |  |  |  |  |  |
| Sensorimotor |  |  |  |  |  |

*Figure 7.1: Language-conceptual model showing the association of language development and algebra understanding within the SOLO cognitive framework*

The type of mathematical knowledge displayed by the students is also described, aligned with levels of the SOLO model. Rasch ability groups are listed in descending order in the third column. The schematised Rasch model of ability estimates and item-difficulty estimates (Chapter 4) allows these groupings to be identified by clusters of students, and of items.

Although the linguistic model described here has been developed in the context of students' understanding of early, basic symbolic algebra concepts, it might also be applied in other contexts, and in modes other than the concrete-symbolic and early formal. Thus, the model can be more generally applied: fluent, complex, but concise explanations that are based on the articulation of appropriate, general mathematical principles indicate a secure understanding of concepts relevant to the context, whilst statements that are simple, but vague, often rambling or verbose, and/or expressed as personal perspectives indicate a more fragile understanding.

## CONCLUSION

If "language is the dress of thought", how students express their understanding of mathematical ideas offers a way of seeing into their minds. Correct answers and correct terminology do not guarantee correct, or mathematically useful, thinking. It is often the subtle linguistic clues that should prompt teachers to probe students' responses more carefully, rather than the apparent mathematical correctness of the response.

Hence, in the course of normal classroom interactions, teachers make assessments about the progress of their students based on factors associated with conversational "gambits" – the gestures, voice inflexions and tonality, and use of language. These are signals used in all human communications. However, they are difficult to identify, partly because they are used automatically and intuitively, and partly because it is the nature of speech that it is immediate, and transient. A model that links language use with conceptual development provides a theoretical, substantive basis for such assessment. The language-conceptual model provides a framework based on the pragmatic analysis of language and objectively measured mathematical success; thereby making implicit "professional knowledge" explicit.

To devise a model linking mathematical understanding with language use through a theoretical conceptual framework was the aim of this study, articulated specifically through the sixth research question, and the focus of this chapter. The model (Figure 7.1) aligns student success in algebra, with linguistic features identified from interviews with students. The linguistic features that serve as indicators are: the use of personal pronouns, the occurrence of hedges and attribution shields, the extent of generality or specificity of the response, and the formality of the register of the explanations, and their mathematical appropriateness.

Cooperative principles of language function in classroom discourse and cannot be disregarded in making judgements about students' mathematical understanding. The language-conceptual model provides a rationale, and a focus, for the conduct of productive and developmental classroom discourse in mathematics. In the following chapter, implications of the model for teaching and research are considered, and limitations to the study acknowledged.

# CHAPTER 9: EVALUATION AND IMPLICATIONS OF THE STUDY

> *The extremely subtle pragmatic interpretive judgements regularly made by teachers and students in the course of mathematics teaching and learning will move steadily to the fore as a research topic (Pimm, 1994, p.167).*

Pimm's prediction has been realised by subsequent research with students in primary school (e.g., Bills, 2002; Bills & Gray, 2001). Researchers working in other languages also have begun to explore possible associations between the pragmatics of their language and students' mathematical understanding at tertiary level (e.g. Boero, Douek, & Ferrari, 2002; Ferrari, 2004). This study has examined the pragmatics of students' language in association with their first encounters with symbolic algebra in the early years of secondary school. It has taken a snap-shot, using students in Years 8, 9 and 10 – the years in which formal, symbolic algebra is introduced, as evidenced by syllabus documents. The study goes some way to filling the chronological gap – and a conceptual gap – between the work of Gray and Bills, and that of Boero, Douek, and Ferrari.

This chapter identifies and discusses limitations of the study, and summarises the results in the first two sections. A third section discusses implications of the results, and of adoption of the model, for the teaching of mathematics, and in particular, for the teaching of algebra. Implications for further research are discussed in a fourth section.

## LIMITATIONS OF THE STUDY

This section considers limitations of the study, and consequently some of the aspects of the study that could be considered with caution, and which also suggest avenues for further research. The limitations discussed are the sample size of students interviewed, a necessary result of opportunistic sampling, and possible issues arising from the classroom contexts of the students in the study.

## The sample size of students chosen for interview

In hindsight, it would have been valuable to have interviewed more students at the upper and lower levels of achievement on the survey. Out of the 222 students who responded to the algebra survey, 51 were identified on the basis of raw scores on the survey and approached for interview.

Students who gave few responses to the survey items were not approached for interview. Experience from the pilot study suggested that students with very low scores on the algebra test had little to contribute to an interview. To explore the mathematical thinking of these students would have required a different set of data collecting instruments. This was outside the scope of the study.

Of the 51 students identified, 31 agreed to participate, and all were interviewed. This slightly smaller sample meant that there were fewer students in the High-ability group. This was of some concern as, when in later analysis of linguistic features, the High-ability group appeared to comprise two sub-groups. Some conjectures have been made regarding the linguistic characteristics of the language of the three most able students. The trends in the model are suggestive but, because of the very small sample available, the results may have been influenced by particular speech habits of individual students.

Despite the smaller than anticipated sample of students interviewed, interviews resulted in approximately 65000 words spoken by the students. Of these approximately 7500 words involved countable linguistic features such as pronouns and hedges. Given the exploratory nature of the study, the numbers were adequate to provide a basis for general findings regarding these aspects of language use. Other aspects of the study, such as the determination of verbosity, tense, and the frequency of attribution shields, did not require a finely-tuned examination of numbers of particular words.

However, the focus of the study was not on the determination of a model of normative behaviour. The study was designed to establish broadly-based developmental pathways in the context of students' early algebra learning and their language use.

### Contextual issues

Students' learning experiences must influence their thinking, and their use of language. Thus, a second possible limitation concerns the context of language use in the students' classrooms. There are two issues. The first concerns the language used by students and teachers and the extent to which students' language mirrored remembered or practised comments by their teachers. The second issue is whether teachers encourage and provide opportunities for students to talk about mathematics, and the nature of that talk.

Interviewed students were drawn from 12 classes from three different schools, taught by a total of ten different teachers. This provided a broad, general perspective on the language practices across a variety of classrooms. These practices could be inferred from interview responses. However – and this is taken up and discussed in the section on further research – it would be valuable to explore the extent of opportunities for students to engage in mathematical discussions and the language use of individual students, as well as that of entire classrooms.

The model, however, has been developed by the aggregation of several linguistic features, not by attending to any one particular aspect. The limitations of the study are those imposed by the sample of students interviewed, and the mathematical context chosen. The conclusions reached have been presented cautiously, conscious of these important constraints. The limitations identified suggest areas of research that would elaborate and extend the model, and provide further insight into reasons behind some of the linguistic behaviours noted.

## SUMMARY OF RESULTS OF THE STUDY

The aim of the study was to devise a developmental model that aligned linguistic features of students' explanations with the quality of their algebraic understanding. Data from six research questions were used to inform the model. Three of the questions (1, 2 and 3) sought to explore students' understanding of algebra, two (4 and 5) explored linguistic features of students' utterances. The final research question was designed to bring together the algebraic and linguistic aspects of the study.

Students' algebraic understanding was established by analysing their responses to a survey of algebra items. Rasch-modelling of responses resulted in identification of groups of students clustered around three significantly different ability estimates. Corresponding to these were three clusters of items. Content of statements by students was used to determine possible developmental stages in algebra understanding The ability-groups of students and item-difficulty clusters were used to structure linguistic analyses of student utterances. Mapping both mathematical approaches and language used by students to express their understanding to the SOLO model of cognitive development resulted in the language-conceptual model.

The SOLO model would suggest that the target mode for algebraic thinking is the formal mode, whilst that for arithmetic thinking is the concrete symbolic mode. The students studied represented a period of transition in cognitive development from that within the concrete symbolic mode to that of the formal mode.

Low-achieving students (Low-ability group) tended to answer only those items that were essentially arithmetic. High-achieving students (High-ability group) were able to respond successfully to items that were to some extent, dissociated from any numerical referent. Responses from students in the Average-ability group could be described on a continuum between these two extreme groups.

Low-ability students demonstrated a reliance on ikonic support, by their focus on superficial features of algebraic expressions, and a literal interpretation of procedural metaphors. These students also demonstrated fragile number fact recall, articulated by their counting on or recitation of tables-facts.

Average-ability students, tended to rely on arithmetic results to confirm their algebraic responses. In terms of the SOLO model, these students were operating in the concrete symbolic mode. Their number-sense was often sound. In the lower-achieving levels of this group, this was limited to understanding of relationships between positive whole numbers. The higher-achieving students of this group were more able to articulate and use relationships between arithmetic operations. The procedures used were those that produced intermediate answers, so that multi-step solutions or manipulations were executed by students recording "answers" at each step. Often relationships between each step were not attended to, resulting in many errors.

High-ability students were able to respond successfully to items in terms of remembered procedures, the results of which they did not always need to confirm by arithmetic. This suggests that some of the highest achieving students were operating in the formal mode. However, none of those interviewed could provide explanations that addressed underlying mathematical concepts.

The language used by students as they attempted to explain their thinking changed as they encountered algebraic ideas that they found difficult to understand. Challenging ideas were expressed in informal language, often as metaphors. Ideas that were more confidently understood were expressed in language of a higher register. Students operating in a conjecturing environment used language that tended to be a mixture of the more formal mathematical register and the informal, everyday register.

The informal register tended to use the first person, or the vague third person pronoun, be of low modality, and use terminology lacking mathematical specificity. Confidence in understanding mathematical ideas was conveyed by students' use of an increasingly formal register – use of the general *you*, high modality and appropriate use of mathematical vocabulary.

The development of an explicit, mathematically appropriate register was found to align with students' development of arithmetic relationships and the generalisation of those relationships to formal algebra. Aligning algebraic understanding, and linguistic changes to a conceptual development hierarchy provides objective empirically derived bases for the subtle intuitive judgements made by teachers.

## IMPLICATIONS FOR THE TEACHING OF MATHEMATICS

Teachers, and students, need to be aware of how language functions to communicate understanding. This requires a shift in perspective about the nature of school mathematics, how mathematics might be taught, and how classroom discourse might be managed.

Teaching involves several partnerships: those between the teacher and students; between teachers and parents; between teachers and the school system and a wider community. Hence, teaching necessarily involves teachers in activities that have repercussions beyond the classroom. Possibly the most discussed is that of assessment

– its purposes and consequent implementation. Inescapably associated with assessment is classroom practice. What is to be assessed must first be taught. This section considers implications of the results of the study for pedagogy, with a focus on classroom language practices, the introduction of formal algebra, and formative assessment.

### Classroom language practices

Research over at least two decades has supported a Vygotskian paradigm where classroom conversations are essential for students developing mathematical understanding (e.g., Austin & Howson, 1979; Halliday 1985; Mousley & Marks, 1991; Pimm, 1987; Zack & Graves, 2002). The development of a language-conceptual model suggests that there needs to be greater attention given to ways in which language is used in mathematics teaching and learning. A pedagogical shift in this direction should then influence approaches to algebra away from procedural, symbolic manipulation to an emphasis on meanings and connections, developed through discussion and argument. Classroom discourse with discussion and argument also provides teachers with the opportunity to assess the quality of students' understanding inferred from their use of language.

The language-conceptual model demonstrates that as understanding develops, so too does language change and the appropriate register develop. Results from the study suggest that greater attention needs to be paid to language use in mathematics classrooms, both to the content and the quality of expression. The consequences could be a change in the types of mathematical activities, and the ways in which mathematical understanding could be assessed.

By explicitly developing a mathematical register, teachers provide students with the language with which to dress their thoughts. To do so, teachers model the appropriate language, and provide opportunities for students' language to develop from the informal and colloquial to the more formal, literate register of the subject. The teachers' focus then must shift between an awareness of content, and an awareness of language use. Teachers need to listen not just for content information, but also for linguistic clues that indicate students' confidence in, and commitment to, the truth-value of their utterances (Rowland, 2000).

Teachers also need a consciousness of their own language use, and a clear understanding of the ways in which conceptual metaphors (Nunez, 2000) can assist the development of understanding. The use of metaphor to help students to link abstractions with more readily accessible ideas is important. However, the mathematics behind the metaphor needs to be made explicit and elaborated. Whilst communication of mathematical ideas relies on metaphors using physical and visual analogies in informal language, students' mathematical understanding remains limited.

Bills and Gray (2001) suggested that primary school students who were successful in mathematics were more likely to use the register (words and syntax) of the teacher. However, it would appear that much of the remembered, and imitated, teacher-language is not of a mathematical register, as evidenced by students' use of the various informal phrases identified in the study. Metaphors and simplistic illustrations can help students understand mathematical ideas. When language remains at a non-literate, or even idiomatic level, and when examples that students encounter do not extend ideas embodied in the language, students' mathematical development becomes constrained. Hence, classroom activities need to be framed in ways that move students' thinking, and talking, beyond these limits.

Assessing students' understanding by listening to their conversations will only be useful if students are used to talking mathematics. If talking mathematics is an unfamiliar experience, students' use of language will not necessarily reflect their mathematical understanding through linguistic features: the act of talking will be cognitively demanding, and hence the language used will reflect that cognitive challenge – as well, perhaps, as that of the mathematics.

A language-conceptual model demonstrates the intimate and essential association between language development and development of algebra understanding. Without the language, students are left without the means of framing their understanding. Students can develop the appropriate register when they hear it modelled by their teachers, have the language explicitly taught, and have opportunities for practice. Students are then able to move between colloquial and formal registers, in much the same way as they move between procedural and conceptual aspects in the transition from arithmetic to algebra.

## The introduction of formal algebra

Algebra itself was not the focus of the study. It was used as the means by which linguistic structures associated with conceptual growth could be investigated. Throughout the study, however, several aspects have emerged that have important implications for the teaching of algebra, and which are closely linked to the necessary emphasis on the use and development of language.

The first is that the students in the study, although successful, could express only a procedural understanding of the mathematics. The second is that the examples chosen from the syllabus failed to extend students' thinking beyond the arithmetic in many cases. By the time many students had been studying algebra for four years, many remained tied to arithmetic strategies – trial-and-error, or reliance on simple number facts. A third concern was the limited arithmetic background of many students.

In order for students to move beyond the conceptual limitations of metaphors used to structure their algebraic understanding, they need to encounter examples that challenge the ideas in the metaphors. Examples of algebraic procedures need to be such that move students to the point where they comprehend the necessity to dissociate their mathematical thinking from the real world of arithmetic and deal with generalised abstractions about number behaviour and relationships.

Many, if not most, of the students studied possessed limited arithmetic understanding: most usually understood numbers as consisting of positive integers only. Even the better students needed to be prompted to consider the existence of fractions, or of negative numbers. This limited understanding appears to be reinforced by the lack of textbook examples using a broad range of number types, or addressing number theory in piecemeal fashion. The development of such reasoning, as is here advocated, is facilitated by having students to engage in mathematical discourse, but that is difficult when students' experiences of arithmetic are limited.

The simplicity of many textbook and syllabus examples might be justified on the grounds that introducing students to new ideas without adding conceptual complexity in the form of difficult arithmetic manipulations. However, if algebra is to be approached from the perspectives of generalised arithmetic and generalised patterns, the arithmetic and the patterns have to be complex enough to disallow simple additive strategies. Data from the study suggest that more time needs to be spent on developing flexible, imaginative arithmetic understanding, before students are introduced to the formal aspects of manipulative symbolic algebra. Although such a focus would begin in primary school (e.g., Warren, 2003), it should take on the necessarily greater complexity in secondary school.

## Assessment practices

Assessment is a complex issue for teachers, associated as it is with formal systemic reporting and accountability requirements. Where the assessment is formative, influencing teaching decisions, evidence to support decisions made on-the-run is often tenuous, yet vital if teachers are required to justify decisions. Teachers' attention to classroom discourse more often than not focuses on the content rather than the pragmatics of the conversation, but the "gambits" that humans use to communicate ideas inevitably play a sub-conscious role in the judgement-making of teachers. If teachers can listen with one ear to the content, and one to the manner in which the understanding of content is communicated, they will have a basis on which they can make explicit the conscious rationale for their decisions.

A language-conceptual model provides such a basis, although it is not intended to provide yet another set of "outcomes" to be checked off. Like the SOLO model, which provides the developmental framework for the model, linguistic indicators are task-specific; they apply in a particular instance, at a particular time. Over time, a pattern might become apparent allowing judgements to be made about the conceptual development of individual students in a broader mathematical context. The model serves to illustrate cognitive struggle through the language used by students. It assumes that as understanding develops, the struggle diminishes, and the language used by students increasingly adopts an appropriate register, high modality and a generality of focus.

The model can also serve to alert teachers to the potential of listening to students – not just for what is said, but how it is said. Thus, teachers become aware of the need to use and develop appropriate language in the classroom, making the mathematical meanings explicit.

# IMPLICATIONS FOR RESEARCH

Limitations of the study and the conclusions to be drawn from the evidence anticipate areas for further research. Two particular areas are: fine-tuning and extending the language-conceptual model; and approaches to the teaching of algebra that focus on the connections to arithmetic, and the development of criteria that would inform a partial-credit model of item-difficulty.

## The language-conceptual model

Many of the patterns and trends that have been described and which have informed the language-conceptual model need to be further elaborated by evidence from a larger cohort; in particular, more extensive evidence is needed for the upper and lower levels of the model. Even the most able students were unable to give conceptual explanations of the mathematics involved. Thus, it would be useful to extend the study to include students in the senior years of secondary school, and at the most advanced levels of mathematics in secondary school. Such research would connect with, and support the findings of Boero, Douek, and Ferrari (2002). This would be of interest also if it can be found that the language-conceptual model is applicable in languages other than English, and in other educational cultures.

Other aspects of discourse analysis also are needed to substantiate the model. Listening to the tapes of the interviews, once more, it is apparent that prosodic indicators (voice inflexions, emphases, volume) play an important part in communicating students' attitudes and confidence in what they say. These have not been examined in the study. Such examination would be assisted by video-taping of interviews, with the result that gestures also could be examined. This would contribute to understanding of students' conceptual development, and classroom negotiation of meaning.

Teaching contexts were not considered in the study, but because the voice of the teacher was apparent in many phrases used by the students, this influence needs to be explored further. In particular, level-reduced procedures, metaphors (such as "circling like terms") and colloquially expressed explanations were probably used by teachers, and repeated by their students in response to the interview questions.

The evidence suggests that teachers need explicitly to link informal language and metaphors to the appropriate mathematics so that the mathematical register is developed with all students. Teachers also need to interrogate closely student statements that appear to approximate the teacher's expectations. How this is effected in a classroom, and with what regularity must have direct influence on the language use and mathematical understanding of students. Investigation of this aspect would provide valuable data with which to elaborate the model, and provide evidence for changes in classroom practice.

### Algebra aspects of the model

There is a strong indication that students' arithmetic understanding is fragile, particularly at the lower end of ability. Such fragile understanding would seem to contribute to the struggles many students have in grasping concepts in the formal, generalised system of symbolic algebra. A study, of similar design to this reported study, using a survey that examined late arithmetic/pre-algebra understanding of upper-primary and early secondary students (middle-school), and their explanations would provide data that elaborate the language-conceptual model at the lower end of ability (in algebra terms).

At the same time, it would build on the work of Bills and Gray (2001) and of Bills (2002) linking primary students' linguistic behaviours to that of beginning secondary students. This could highlight important issues affecting students' learning in the transition from primary to secondary education in general, and mathematics in particular. Such a study would also contribute to elucidating the foundations upon which the teaching of formal algebra could best be laid.

Other aspects of algebra understanding also have been suggested by analysis of responses to the survey items. Error analysis suggests that criteria for establishing a partial-credit model to analyse similar items (e.g., Items 16 and 8) could be found. This could then be used as a basis for a developmental teaching sequence. Responses to the survey, and during interviews, also suggest that the form of an expression or equation, and the type of solution to an equation influence a student's success in manipulating the expression or solving the equation.

Algebra dominates the secondary syllabus and anecdotal evidence suggests that success in algebra becomes a benchmark of secondary mathematics prowess. It also appears that many students are excluded from the time of their first encounters with the subject. One of the reasons for that exclusion might be the ways in which language is used, or not used, to foster thinking that can lead students to understand the ideas of algebra.

## CONCLUSION

> *Language is so central to the whole of the educational process that its role was never even talked about, since no-one could conceive of education without it. (Halliday, 1985. p.96)*

Understanding how language works, and ensuring that students have the opportunity to acquire and develop the language in which to dress their mathematical thoughts is integral to their mathematical development. The research reported in this study provokes more questions than it set out to answer. The proposed language-conceptual model needs further investigation over a wider-ranging sample of students, so that constraints identified in this study are avoided, and the model extended and elaborated. In effect, the model provides a starting point for a range of research – into its applicability in other languages, and other areas of mathematics, as well as into the development of algebraic understanding, and associated shifts in pedagogy when language aspects are given greater prominence in classrooms.

A focus on language in a mathematics classroom is not so very new, but it is not yet a commonly understood or accepted aspect of mathematics pedagogy. Institutional requirements go some way to addressing this, but until the communicative human nature of mathematics is appreciated, understood, and accepted, the focus is likely to remain on the procedural and formal aspects. The language-conceptual model provides a rationale, and a focus, for the conduct of productive and developmental classroom discourse in mathematics. Teachers, recognising that students' language both informs and is informed by their understanding, will listen, not only for the mathematical content of students' statements, but also for the ways in which language is used, and thus make qualitatively richer judgements about students' learning.

# REFERENCES

Adams, R.J., & Khoo, S. (1996). *Quest: The interactive test analysis system* (Version 2.1) [Computer software]. Melbourne: Australian Council for Educational Research

Arzarello, F. (1998). The role of natural language in prealgebraic and algebraic thinking. In H. Steinbring, M. Bussi & A. Sierpinksa (Eds.), *Language and Communication in the Mathematics Classroom* (pp. 249-261). Reston, VA: National Council of Teachers of Mathematics.

Austin, J., & Howson, A. (1979). Language and mathematical education. *Educational Studies in Mathematics, 10*(3), 161 - 197.

Bechervaise, N. (1992). Mathematics: a foreign language? *The Australian Mathematics Teacher, 48*(2), 4 - 8.

Bell, G. (1993). Mathematics x Language = Language x Mathematics. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 140 - 150). Melbourne: Australian Council for Educational Research.

Bickmore-Brand, J. (1993). Applying language-learning principles to mathematics teaching. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 79-90). Melbourne: Australian Council for Educational Research.

Bickmore-Brand, J., & Gawned, S. (1990). Scaffolding for improved mathematical understanding. In J. Bickmore-Brand (Ed.), *Language in Mathematics* (pp. 43-58). Melbourne: Australian Reading Association.

Biggs, J., & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.

Biggs, J., & Collis, K. (1983). Matriculation, degree structures, and levels of student thinking. *The Australian Journal of Education, 27*(2), 151-163.

Biggs, J., & Collis, K. (1989). Towards a model of school-based curriculum development and assessment using the SOLO Taxonomy. *The Australian Journal of Education, 33*(2), 149-161.

Biggs, J., & Collis, K. (1991). Multi-modal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), *Intelligence Reconceptualisation and Measurement* (pp. 57-76). Hillsdale, NJ: Lawrence Erlbaum.

Bills, C. (2002). Linguistic pointers in young children's descriptions of mental calculation. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Annual Conference of the International Group for the Psychology of Mathematics Education,* (pp. 97-104). Norwich: PME.

Bills, C., & Gray, E. (2001). The "particular", "generic" and "general" in young children's mental calculations. In M. van Heuval-Panhuizen (Ed.), *Proceedings of the 25th Annual Conference of the International Group for the Psychology of Mathematics Education*, (pp. 153-160). Utrecht: PME.

Bills, L. (1999). Students talking: An analysis of how students convey attitude in maths talk. *Education Review, 51*(2), 161 - 171.

Board of Studies NSW. (2002a). *Mathematics 7 - 10 Syllabus 2002 (Vol. 7 - 10)*. Sydney: Board of Studies NSW.

Board of Studies NSW. (2002b). *Mathematics K - 6 Syllabus 2002 (Vol. K - 6)*. Sydney: Board of Studies NSW.

Board of Studies NSW. (1999). *Mathematics Years 7-8 Syllabus Outcomes*. Sydney: Board of Studies NSW.

Board of Studies NSW. (1998). *Mathematics K - 6 Outcomes & Indicators*. Sydney: Board of Studies NSW.

Board of Studies NSW. (1996a). *Mathematics Years 9-10: Advanced Course Stage 5*. Sydney: Board of Studies NSW.

Board of Studies NSW. (1996b). *Mathematics Years 9-10: Intermediate Course Stage 5*. Sydney: Board of Studies NSW.

Board of Studies NSW. (1996c). *Mathematics Years 9-10: General Course Stage 5*. Sydney: Board of Studies NSW.

Board of Secondary Education NSW. (1990). *Mathematics Syllabus Years 7 & 8*. Sydney: Board of Secondary Education NSW.

Boero, P., Douek, N., & Ferrari, P. (2002). Developing mastery of natural language: Approaches to theoretical aspects of mathematics. In L.D. English (Ed.), *Handbook of International Research in Mathematics Education* (pp. 241-267). Mahwah, NJ: Lawrence Erlbaum.

Bond, T., & Fox, C. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Bond, T., & Fox, C. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Charbonneau, M., & John-Steiner, V. (1988). Patterns of experience and the language of mathematics. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and Cultural Influences on Learning Mathematics* (pp. 91-100). Hillsdale, N.J: Lawrence Erlbaum.

Clarkson, P., & Thomas, J. (1993). Communicating mathematics bilingually. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 263-273). Melbourne: Australian Council for Educational Research.

Cobb, P., Wood, T., & Yackel, E. (1995). Learning through problem-solving. In P. Murphy, M. Selinger, J. Bourne & M. Briggs (Eds.). *Subject Learning in the Primary Curriculum: Issues in English, Science and Mathematics* (pp. 232 - 251). London: Routledge.

Cockcroft W.H. (1982). *Mathematics Counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr W.H. Cockcroft*. London: Her Majesty's Stationery Office, 1982.

Collis, K. (1982). *Language development and intellectual functioning*. Occasional paper No. 10, presented to APEID/UNESCO Forum, July 1982, Bangkok, Thailand.

Collis, K. (1992). Curriculum and assessment: A basic cognitive model. In G. Leder (Ed.), *Assessment and Learning of Mathematics* (pp. 24-45). Hawthorn: Australian Council for Education Research.

Collis, K. (1994). Mathematics and cognition. In H. Chick & J. Watson (Eds.), *Mathematics and Teaching - Topics for the professional development of teachers*. (pp. 337-369). Adelaide: Australian Association of Mathematics Teachers

Collis, K., Watson, J., & Campbell, K. (1993). Cognitive functioning in mathematical problem solving during early adolescence. *Mathematics Education Research Journal, 5*(2) pp.107-123.

Davey, G. (1988). *Primary School Geometry - the van Hiele challenge in mathematical interfaces*. Paper presented at the 12th Biennial Conference of the Australian Association of Mathematics Teachers, Newcastle.

Dawe, L. (1995). Language and culture in the teaching and learning of mathematics. In L. Grimison & J. Pegg (Eds.), *Teaching Secondary School Mathematics* (pp. 230 - 245). Sydney: Harcourt - Brace.

Donaldson, M. (1986). *Children's Explanations: A Psycholinguistic Study*. Cambridge: Cambridge University Press.

Douek, N., (2002). Context, complexity, and argumentation. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2). (pp. 297-304). Norwich: PME.

Ellerton, N., & Clarkson, P. (Eds). (1996). *Language Factors in Mathematics Teaching and Learning* (Vol. 4). Dortrecht: Kluwer.

Ellerton, N., & Clements, M. (1991). *Mathematics in Language: A Review of Language Factors in Mathematics Learning*. Melbourne: Deakin University Press.

Ellerton, N., & Clements, M. (1996). Researching language factors in mathematics education: The Australasian contribution. In W. Atweh, K. Owens & P. Sullivan (Eds.), *Review of Mathematics Education in Australasia 1992-1995* (pp. 191-235). Melbourne: Mathematics Education Research Group of Australasia.

Ericsson, K., & Simon, H. (1980). Verbal Reports as Data. *Psychological Review, 87*(3), 215-224.

Ferrari, P. (2002). Developing language through communication and conversion of semiotic systems. In A. Cockburn & E. Nardi (Eds.), P*roceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2). (pp. 353-360). Norwich: PME.

Fitzpatrick, J.B. (1992). *New Senior Mathematics: Two unit course for Years 11 & 12*. Port Melbourne: Heinemann Education Australia.

Fletcher, J., & Loakes, D. (2006). Patterns of rising and falling in Australian English. In P. Warren & C. Watson (Eds.) *Proceedings of 11th Australian International Conference on Speech Science and Technology*, Auckland, NZ. Accessed 13 July, 2007 from http://www.assta.org/sst/2006/sst2006-32.pdf

Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dortrecht: D. Reidel.

Frid, S. (1993). Communicating mathematics: A social sharing of language and decisions relating to truth and validity. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Research* (pp. 26 - 40). Melbourne: Australian Council for Educational Research.

Gawned, S. (1990). An emerging model of the language of mathematics. In J. Bickmore-Brand (Ed.), *Language in Mathematics* (pp. 27 - 42). Melbourne: Australian Reading Association.

Gooding, A., & Stacey, K. (1993). How children help each other learn in groups. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 41 - 50). Melbourne: Australian Council for Educational Research.

Gray, E, & Tall, D. (1991). Duality, ambiguity & flexibility in successful mathematical thinking. In *Proceedings of the 15th Conference of the Psychology of Mathematics Education* (Vol. 2). (pp. 72– 79). Assisi: PME.

Gray, E., & Tall, D. (1992a). *Success and Failure in Mathematics: Procept and Procedure – A Primary Perspective*. Workshop on Mathematics Education and Computers, Taipei National University, April 1992, pp. 209– 215.

Gray, E., & Tall, D. (1992b). *Success and Failure in Mathematics: Procept and Procedure – Secondary Mathematics*. Workshop on Mathematics Education and Computers, Taipei National University, April 1992, pp. 216– 221.

Gray, E., & Tall, D. (1994) Duality, ambiguity and flexibility: A proceptual view of simple arithmetic. *The Journal for Research in Mathematics Education*, *26* (2), 115– 141.

Halliday, M. (1985). *Spoken and Written Language*. Melbourne: Deakin University Press.

Halliday, M. (1976). *System and Function in Language*. Oxford: Oxford University Press.

Halliday, M., & Hassan, R. (1985). *Language, Context and Text: Aspects of Language in a Social-semiotic Perspective*. Melbourne: Deakin University Press.

Howard, P. (1999). "*Just the words that she used I can't understand, I don't know what they mean*". Paper presented at the 1999 Annual Conference of the Mathematics Association of NSW, Port Macquarie, NSW.

Jacobsen, E., & Dawe, L. (1984). *Topic area: Language and Mathematics*. Paper presented at the Fifth International Congress on Mathematical Education, Adelaide.

Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws (Ed.) *Handbook of Research on Mathematics Teaching and Learning* (pp. 390-419). New York: Macmillan.

Kress, G. R., & Hodge, R. (1979). *Language and Ideology*. London: Routledge & Kegan Paul.

Stiff, L.V. & Curcio, F.R. (Eds.). *Developing Mathematical Reasoning in Grades K-12: NCTM 1999 Yearbook*. Reston, VA: National Council of Teachers of Mathematics.

Laborde, C. (1990). Language and mathematics. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education* (pp. 53 -69). Cambridge: Cambridge University Press .

Leder, G. E. (1992). *Assessment and Learning of Mathematics*. Hawthorn: Australian Council for Education Research.

Leedy, P. (1993). *Practical Research: Planning and Design*. London: Macmillan.

Linacre, J.M. (2002). What do infit and outfit, mean-square and standardized mean? In *Rasch Measurement Transactions*, *16*(2). Retrieved on 28 August, 2008 from http://www.rasch.org/rmt162f.htm

Love, E., & Mason, J. (1995). Telling and asking. In P. Murphy, M. Selinger, J. Bourne & M. Briggs (Eds.), *Subject Learning in the Primary Curriculum: Issues in English, Science and Mathematics* (pp. 254-266). London: Routledge.

Lowry, R. *t-Test for the Significance of the Difference between the Means of Two Independent Samples*. Retrieved on 13 July, 2006 from http://faculty.vassar.edu/lowry/ch11pt1.html

Lynch, B, & Parr, R. (1982). *Maths 10*. Melbourne: Longman Sorrett.

MacGregor, M. (1990). Reading and writing in mathematics. In J. Bickmore-Brand (Ed.), *Language in Mathematics* (pp. 100 - 108). Melbourne: Australian Reading Association.

MacGregor, M. (1993). Interaction of language competence and mathematics learning. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Research* (pp. 51 - 59). Melbourne: Australian Council for Education Research.

MacGregor, M. (2002). Using words to explain mathematical ideas. *Australian Journal of Language and Literacy, 25*(1), 78 - 90.

MacGregor, M., & Price, E. (1999). An exploration of aspects of language proficiency and algebra learning. *Journal for Research in Mathematics Education, 30*(4), 449 - 467.

MacGregor, M., & Stacey, K. (1994). Metalinguistic awareness and algebra learning. In J. da Ponte & J. Matos (Eds.), *Proceedings of the 18th International Conference for the Psychology of Mathematics Education* (Vol. 3, pp. 200-207). Lisbon: PME.

MacGregor, M., & Stacey, K. (1996). Origins of students' interpretations of algebraic notation. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 297-304). Valencia: PME.

Malone, J., & Miller, D. (1993). Communicating mathematical terms in writing: Some influential variables. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 177 - 190). Melbourne: Australian Council for Education Research.

Mason, J. (1996). Expressing generality and roots of algebra. In N. Bednarz, C. Kieran & L. Lee (Eds.), *Approaches to Algebra: Perspectives for Research and Teaching* (pp. 65-86). Dortrecht: Kluwer.

Mousley, J., & Marks, G. (1991). *Discourses in Mathematics*. Melbourne: Deakin University Press.

National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Nunes, T. (1997). Systems of signs and mathematical reasoning. In T. Nunes & P. Bryant (Eds.), *Learning and Teaching Mathematics: An International Perspective* (pp. 29 - 44). Hove: Psychology Press.

Nunez, R. (2000). Mathematical idea analysis: What embodied cognitive science can say about the human nature of mathematics. In T. Nakahara & M. Koyama (Eds.) *Proceedings of 24th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 3 –38). Hiroshima University: PME.

Padula, J., Lam, S., & Schmidtke, M. (2001). Syntax and word order: Important aspects of mathematical English. *Australian Mathematics Teacher, 57*(4), 31 - 35.

Pegg, J. (1992). Students' understanding of geometry: Theoretical perspectives. In B. Southwell, B. Perry & K. Owens (Eds.), *Space - The First and Final Frontier* (pp. 18-36). Sydney: Mathematics Education Research Group of Australasia.

Pegg, J. (2003). Assessment in mathematics: A developmental approach. In J. Royer (Ed.) *Mathematical Cognition* (pp.227-259). Charlotte: Information Age Publishing.

Pegg, J., & Tall, D. (2002). Fundamental cycles of cognitive growth. In A.D. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*, (pp.41-48). Norwich: PME.

Pegg, J., & Tall, D. (2005). The fundamental cycle of concept construction underlying various theoretical frameworks. *International Reviews on Mathematical Education*, 37(6), 468-475.

Pengelly, H. (1990). Acquiring the language of mathematics. In J. Bickmore-Brand (Ed.), *Language in Mathematics* (pp. 10 - 26). Melbourne: Australian Reading Association.

Perry, J. (2001). Fifth-graders' mathematical communication: Lessons from the field. *The Educational Forum, 66*(1), 71-79.

Pimm, D. (1987). *Speaking Mathematically: Communication in Mathematics Classrooms*. London: Routledge & Kegan Paul.

Pimm, D. (1991). Communicating mathematically. In K. Durkin & B. Shire (Eds.), *Language in Mathematical Education: Research and practice*. London: Open University Press.

Pimm, D. (1994a). Another psychology of mathematics education. In P. Ernest (Ed.), *Constructing Mathematical Knowledge: Epistemology and Mathematics Education* (pp. 111- 124). London: The Falmer Press.

Pimm, D. (1994b). Mathematics classroom language: form, function and force. In R. Beihler, R. W. Scholz, R. Strasser & B. Winkelmann (Eds.), *Didactics of Mathematics as a Scientific Discipline* (pp. 159 - 169). Dortrecht: Kluwer.

Pirie, S. (1998). Crossing the gulf between thought and symbol: Language as (slippery) stepping-stones. In H. Steinbring, M. Bussi & A. Sierpinksa (Eds.), *Language and Communication in the Mathematics Classroom, 1998 Yearbook*. (pp. 7-29). Reston, VA: National Council of Teachers of Mathematics.

Quesada, J. (1996). *Mathematics and Languages: Report from working group 10*. Paper presented at the 8th International Congress on Mathematical Education, Seville.

Reeves, N. (1990). The mathematics-language connection. In J. Bickmore-Brand (Ed.), *Language in Mathematics* (pp. 90 - 99). Melbourne: Australian Reading Association.

Rowland, T. (2000). *The Pragmatics of Mathematics Education: Vagueness in Mathematical Discourse*. London: Falmer Press.

Schmidt, S. (1996). Semantic structures of word problems - mediators between mathematical structures and cognitive structures? In C. Alsina, J. Alvares, B. Hodgson, C. Laborde & A. Perez (Eds.), *8th International Congress on Mathematical Education. Selected Lectures*. Sevilla: SAEM 'Thales'.

Schweiger, F. (1994). Mathematics is language. In D. Robitaille, D. Wheeler & C. Keiran (Eds.), *Selected Lectures from the 7th International Congress on Mathematical Education* (pp. 297 - 309). Quebec: Les Presses del'Universite Laval.

Sfard, A. (2001). There is more to discourse than meets the ears: looking at thinking as communicating to learn more about mathematical learning. *Educational Studies in Mathematics, 46*, 13 -57.

Sfard, A. (2002a). Mathematics as a form of communication. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 145-149). Norwich: PME.

Sfard, A. (2002b). Thinking in metaphors and metaphors for thinking. In D. Tall & M. Thomas (Eds.), *Intelligence, Learning and Understanding in Mathematics*. (pp. 79-96). Flaxton: Post Pressed.

Sibley, R. (1999). Math out loud! *Instructor (1999), 112*(7), 24 - 26.

Sierpinska, A., & Lerman, S. (1996). Epistemologies of mathematics and of mathematics education. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International Handbook of Mathematics Education* (Vol. 4, pp. 827-876). Dortrecht: Kluwer.

Southwell, B. (1993). Development and assessment of student's writing. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 223 - 236). Melbourne: Australian Council for Education Research.

Stacey, K., & Steinle, V. (2006). A case of the inapplicability of the Rasch Model to mapping conceptual learning. *Mathematics Education Research Journal*, *18*(2), 77-92.

Stacey, K., Helme, S., & Steinle, V. (2001) Confusions between decimals, fractions and negative numbers: A consequence of the mirror as a conceptual metaphor in three different ways. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 217 - 224). Utrecht: PME.

Steinbring, H. (1992). *Language and Communication in the Mathematics Classroom.* Paper presented at the 7th International Congress on Mathematical Education, Quebec.

Thomas, M., & Tall, D. (1991). Encouraging versatile thinking in algebra using the computer,. *Educational Studies in Mathematics*, *22*(2), 125– 147.

Usiskin, Z. (1996). Mathematics as a language. In P.C. Elliot & M.J. Kenney (Eds.), *Communication in Mathematics – K-12 and Beyond, 1996 Yearbook*. (pp. 231-243). Reston, VA: National Council of Teachers of Mathematics.

Vergnaud, G. (1997). The nature of mathematical concepts. In T. Nunes & P. Bryant (Eds.), *Learning and Teaching Mathematics: An International Perspective* (pp. 5-28). Hove: Psychology Press.

Warren, E. (2003). The role of arithmetic structure in the transition from arithmetic to algebra. *Mathematics Education Research Journal*, *15*(2), 122-137.

Waywood, A. (1993). A phenomenology of report writing: from 'I am' to 'I think' through writing. In M. Stephens, A. Waywood, D. Clarke & J. Izard (Eds.), *Communicating Mathematics: Perspectives from Classroom Practice and Current Research* (pp. 153-163). Melbourne: Australian Council for Education Research.

White, R., & Gunstone, R. (1992). *Probing Understanding*. London: The Falmer Press.

Wiersma, W. (2002). *Research Methods in Education*. Boston: Allyn and Bacon.

Wright, B., & Masters, G. (1982). Rating Scale Analysis. *Rasch Measurement*. Chicago: MESA Press.

Wright, B., & Stone, M. (1979). Best Test Design. *Rasch Measurement*. Chicago: MESA Press.

Zack, V., & Graves, B. (2002). Making mathematical meaning through dialogue: "Once you think of it, the z minus three seems pretty weird". *Educational Studies in Mathematics, 46*, 229 - 271.

# APPENDICES

## APPENDIX A: PILOT STUDY

The pilot study was conducted at a small rural central school. The aim of the study was originally to provide a colleague with information about the algebra performance of his students in the school. To that end, the survey was developed, using syllabus documents (Board of Studies, NSW, 1998) and various school texts as sources of items. The survey was to be given to all students Years 7 – 10 in the second term of the school year. Consequently, the items predominantly covered topics from Stage 4 algebra. Some items were selected from Stage 5 (Board of Studies, NSW, 2002)[*]. The survey was given to the students as a class test, administered by their teachers, but marked by the researcher. The survey was followed up by interviews with some students.

### Participants in the survey

The student cohort consisted of 22 Stage 4 students and 18 Stage 5 students. These students were taught in combined classes for some of the time 9 as years 7/8, and 9/10) and in separate, graded classes for part of the time. This was important for the Stage 5 students, some of whom were attempting the Advanced (as it was then) course for Years 9 and 10. The Year 7/8 class was split for approximately one third of the time into Year groups.

### The interviews

Initially, interviews had not been considered, but the test results indicated that further probing was necessary. Several students were asked to take part in the interviews, which were audio-taped. Interview participants were five Year 7/8 students, two students from Year 8, and three students from Year 7; and, four Year 9/10 students, three of whom were from Year 9, and one student from Year 10 student.

---

[*] The content of the 2002 syllabus, changed little from that in the previous versions. Current understanding of school year levels is best expressed in terms of the new, 2002, syllabus structure.

Students were selected on the basis of representing a range of test success: the Year 7/8 cohort scores ranged from 86/120 to 12/120, and the scores of those interviewed from 86/120 to 56/120; the Year 9/10 scores ranged from 82/120 to 3/120, and the scores of those students interviewed from 82/120 to 15/120. Students with very low scores, did not answer many questions at all, and did not wish to be interviewed. The interview protocol consisted of an informal discussion of individual items with the students. Items were selected either because the students gave correct responses, and the researcher wished to investigate the procedures used; or because the responses were incorrect, or not left as a blank on the survey paper. Questioning about the incorrect responses was designed to elicit student's understanding of the item itself, and their approach to it.

## Results

Survey responses are summarised for each student in Table A.1 below. Results were scored as correct (3), partially correct (2), attempted (1) and no attempt (0). Hence the highest possible score was 120.

*Table A1: Test scores for survey participants*

| Student Year/ID | Test score/120 | Student Year/ID | Test score/120 |
|---|---|---|---|
| 9/10M26 | 82 | 7/8M117 | 27 |
| 9/10M215 | 47 | 7/8M115 | 67 |
| 9/10M134 | 3 | 7/8M118 | 59 |
| 9/10M22 | 31 | 7/8M117 | 12 |
| 9/10M28 | 15 | 7/8M113 | 57 |
| 9/10M214 | 31 | 7/8M15 | 86 |
| 9/10M21 | 12 | 7/8M18 | 48 |
| 9/10M29 | 3 | 7/8M11 | 74 |
| 9/10M23 | 27 | 7/8M119 | 58 |
| 9/10M24 | 60 | 7/8M116 | 56 |
| 9/10M217 | 66 | 7/8M111 | 40 |
| 9/10M212 | 69 | 7/8M112 | 52 |
| 9/10M46 | 4 | 7/8M110 | 61 |
| 9/10M210 | 8 | 7/8M19 | 82 |
| 9/10M27 | 68 | 7/8M121 | 30 |
| 9/10M211 | 55 | 7/8M122 | 53 |
| 9/10M43 | 7 | 7/8M14 | 85 |
| 9/10M4 | 23 | 7/8M12 | 82 |
| | | 7/8M13 | 58 |
| | | 7/8M120 | 55 |
| | | 7/8M116 | 56 |
| | | 7/8M18? | 53 |

## Discussion of results

Some points of interest from the study are outlined briefly in the following.

1. Year 7/8 students were more willing to attempt all items, although not necessarily more successfully than 9/10 students. This has resulted in overall higher scores for students in Years 7/8, despite the fact that they would have had less experience of algebra.

2. Several Year 9/10 students did not participate fully; five out of the 18 students in that year group answered no more than 6 items, none correctly. These students did not agree to be interviewed.

3. The items for which there were most correct responses were:

    a. Item 27: If $x + 5 = 7$, then $x =$…?

    b. Item 28: If $4y = 20$, then $y =$ …?

    c. Item 24: If $a + b = 43$, then $a + b + 2 =$ …?

    d. Item 1: Simplify $3m + 8 + 2m - 5$

    e. Item 2: Simplify $5p - p + 1$

    f. Item 4: Simplify $4 \times 5b$

    g. Item 12: Simplify $2/a \times 3/b$; and,

    h. Item 5: Simplify $2ab + 3b + ab$

4. Many of the misunderstandings/ and understandings common to the students in the large study were apparent from the survey results, and the interviews. These included the manipulation of expressions with brackets, and algebraic fractions.

5. Students also tended to conjoin terms when adding terms, or when dealing with expression within brackets. The following response to Item 10 $(x + y)^2$

describes the confusion demonstrated by many students:

> S: That just goes through my head and goes straight out the other side. ‘Cause its got…No idea. You’re meant to do the brackets first, I think. Go *x* plus *y*, but that’s nearly … I dunno.

6. Students also expressed their confusion with understanding the notation for *a* squared ($a^2$) and 2 times *a* (2*a*).

7. Interviews suggested that students relied on using additive and recursive strategies

8. The students who had greater success on the survey could remember procedures and use them, but could not explain the mathematics. When they did so, it was in informal language, particularly when trying to explain the mathematics behind a procedure, such as “cancelling” in Item 8 [4*ab*/4*b*].

> S: I think I just cancelled out, cancelled out the fours, and cancelled out the b’s and come up with a.
>
> I: What does cancelling mean?
>
> S: When you have two of the same units, and um… Like that for example [pointing to Item 8] four over four. It’s not leaving anything. You are just taking … I dunno… I can’t explain that either.

9. When supplementary questions were asked of some students it became apparent that the structure of expressions such as $2(a + 4) + 3$ was more easily approached than if the expression had been written as $3 + 2\ (a + 4)$. These types of responses were not examined in the main study, but have been alluded to when discussing implications for further research (Chapter 7).

10. The students with low scores who were interviewed, provided relatively little data that focused on the mathematics. They often declined to be drawn on items they found difficult – or simply stated that they did not know. One student (52/120) demonstrated the common perception amongst the students of the rather random nature of mathematics, and algebra in particular:

11. The student was discussing Item 7 [(a - b) + b] with the interviewer. His response had been to give the answer as 6.

I: What about question [item] 7?

S: I'd say I had a pretty good stab at that

I: [Waiting for a further response which did not come]… and you were left with a 6…?

S: 6…I don't know…. Like…

I: how do you think you got 6?

S: I probably just guessed that.

I: What would have prompted you to think of 6? Any idea?

S: That's where I play football; number 6. So… that's probably what I done there.

## Modifications for the main study

Modifications were made both to the survey and the interview protocol.

The pilot study was designed to test the appropriateness of the various algebra items: accessibility to all, or most, students; the ability of the items to elicit a range of responses that provided clues to the thinking of the students; and, whether the selection of items reflected a range of difficulty, and algebra experiences. The interviews were used chiefly to support information from the survey responses, and to help confirm the retention or rejection, or rewriting of items. The changes made are described in the following.

### *Survey*

In the final survey, all items from the original were retained, except for Item 16 (which duplicated Item 12, $2/a \times 3/b$)

This item was replaced with $2/a^2 \times 5a/4$, in the main study. This was one of the most difficult items, and which provided interesting data about student understanding of "cancelling" procedures.

Item 13 was changed from $4p \times 5q \times 3$ to $4r \times 5t \times 3$ because many students seemed to have confused $p$ and $q$, with each other, or as read one or the other or both as "9".

Other items remained unchanged – they elicited a range of responses, and the students interviewed appeared to be willing to address them.

### *Interview protocol*

The interview protocol was changed so that items were grouped according to type – as described in the syllabus. This was to give a more rigorous structure to the interview and to make the time frame for conducting each interview shorter. By asking students about each item most interviews ran to about one hour time. It was felt that this was trying on many students, and that a shorter time-frame needed to be developed. It was anticipated that most interviews in the main study would take approximately half-an-hour. When students addressed one item at a time, they had no opportunity to examine groups of items for common mathematical elements, and hence make generalised statements. As this was to be an important part of the main study, the interview had to be restructured to allow this to occur.

Linguistic features of interviews in the pilot study were not analysed.

# APPENDIX B: DATA COLLECTION INSTRUMENTS

## Appendix B1: Algebra Survey

# Algebra Survey

**Instructions to teachers**

Allocate an ID number to each student. Begin at 101… For different classes use 101…, 201…, 301… etc. Record the student number in your mark book next to the name of the relevant student.

Ensure that students mark their Class/Grade and Stage and their gender.

Allow students about 45 to 60 minutes to complete this survey.

Students need to attempt as many questions as they can to the best of their understanding. (There may be some types which they have not encountered in class, or some which cover work they have not yet met.)

Students must not use a calculator – there is no need to.

Students are to write their answers in the spaces provided. Any working they need to do can be done on the back of each question/answer sheet.

# Algebra Survey

**Student ID** ☐ **Year** 8 | 9 **Stage** 4 | 5 **Male** ☐ **Female** ☐

| Item | **Write these expressions as simply as possible** | **Answers** |
|---|---|---|
| 1 | $3m + 8 + 2m - 5$ | |
| 2 | $5p - p + 1$ | |
| 3 | $\frac{a}{5} + \frac{a}{10}$ | |
| 4 | $4 \times 5b$ | |
| 5 | $2ab + 3b + ab$ | |
| 6 | $5a - 2b + 3a + 3b$ | |
| 7 | $(a - b) + b$ | |
| 8 | $\frac{4ab}{4b}$ | |
| 9 | $2ab \times a$ | |
| 10 | $(x + y)^2$ | |
| 11 | $8p - 2(p + 5)$ | |
| 12 | $\frac{2}{a} \times \frac{3}{b}$ | |
| 13 | $4r \times 5t \times 3$ | |
| 14 | $\frac{x}{3} \div \frac{y}{4}$ | |

| Item | Write these expressions as simply as possible | Answers |
|---|---|---|
| 15 | $\frac{3p}{4} - \frac{p}{8}$ | |
| 16 | $\frac{2}{a^2} \times \frac{5a}{4}$ | |
| 17 | $(6xy)^2$ | |
| 18 | $2(x + 4) + 3(x - 1)$ | |
| 19 | $2(x + 5) + 8$ | |
| **Item** | **Question** | **Answer** |
| 20 | Multiply $x + 5$ by 4 | |
| 21 | Add 4 on to n + 5 | |
| 22 | Add 3 onto 4n | |
| 23 | What can you say about m if m = 3n + 1 and n = 4 ? | |
| 24 | If a + b = 43<br>a + b + 2 = …? | |
| 25 | Take n away from 3n + 1 | |
| 26 | If p + q = 5, then<br>p + q + r = …? | |
| 27 | If $x + 5 = 7$, then $x$ = …? | |
| 28 | If 4y =20 , then y = ….? | |

| Item | Question | Answer |
|---|---|---|
| 29 | What is $t$ if $2t - 23 = 49$? | |
| 30 | If $\frac{x}{4} = 12$, what is $x$? | |
| 31 | Solve $4(p + 3) = 32$ | |
| 32 | If $10y = 5$, $y = \ldots$? | |
| 33 | $\frac{x + 3}{2} = 7$<br>What is $x$ equal to? | |
| 34 | $x + (x + 2) = (x - 1) + 8$<br>What is $x$ equal to? | |
| 35 | $x + \frac{x}{3} = 4$<br>What is $x$ equal to? | |
| 36 | If $\frac{63}{x} = 180$<br>What is x equal to? | |
| 37 | Solve: $5a - 4 = 2a + 8$ | |
| 38 | If $r - 82 = 7$, then<br>$r - 83 = \ldots$..? | |
| 39 | If $ax = 5$, then $x = \ldots$? | |
| 40 | If $e + f = 8$, then<br>$e + f - g = \ldots$? | |

**Appendix B2: Interview Protocol**

# Final Semi – Structured Interview Questions

## Part A: Survey - Items

1. What is going on in your head when you see questions like these?

1. 3m + 8 +2m –5 #
2. 5p – p + 1 #

and these

5. 2ab + 3b + ab
6. 5a – 2b + 3a + 3b

Can you tell me more about these four expressions?

2. What is going on in your head when you see questions like:
4. 4 x 5b #
9. 2ab x a…(and)… #
13. 4r x 5t x 3

Can you tell me more about these expressions? Have you seen an expression like this before (13)?

3. What is going on in your head when you see expressions like this:
20. 2(x + 5)
18. $2(x + 4) + 3(x - 1)$
19. $2(x + 5) + 8$

…and…

7. (a – b) + b *

Can you tell me more? What can you tell me about these and the first expression you were shown?

4. What goes on in your head when you see expressions like:
3. $\frac{a}{5} + \frac{a}{10}$ *
15. $\frac{3p}{4} - \frac{p}{8}$

…and these…

8. $\frac{4ab}{4b}$

12. $\frac{2}{a} \times \frac{3}{b}$

16. $\frac{2}{a^2} \times \frac{5a}{4}$ *

What do these expressions mean to you?

5. What goes on in your head when you answer questions like:

27. If $x + 5 = 7$, then $x = \ldots$? #

$t - 48.4 = 201.9$

29. What is $t$ if $2t - 23 = 49$?

37. Solve: $5a - 4 = 2a + 8$

34. $x + (x + 2) = (x - 1) + 8$
What is $x$ equal to?

31. Solve $4(p + 3) = 32$

In what ways are these equations similar? different?

**How do you know that x = ... is the answer?**

(wait for responses) …and these…

28. If $4y = 20$, then $y = \ldots$? #

32. If $10y = 5$, $y = \ldots$? All or none

What if I gave you this?

39. If $ax = 5$, then $x = \ldots$? *

What do these equations mean?

6. What is going on in your head when you see questions like this?

30. If $\frac{x}{4} = 12$, what is $x$? All or none

33. $\frac{x + 3}{2} = 7$ What is $x$ equal to?

36. If $\frac{63}{x} = 180$ What is x equal to? *

7. What can you tell me about these expressions?

17. $(6xy)^2$ *

10. $(x + y)^2$ *

7. $(a - b) + b$ *

11. 8p – 2(p + 5) *

8. Read these statements out aloud and then tell me how you could rewrite them.

21. Add 4 on to n + 5

22. Add 3 on to 4n

25. Take n away from 3n + 1

26. If p + q = 5 then p + q + r = ?

40. If e + f = 8 then e + f – g = ? *

## Why did you think that?

# indicates questions for which most students gave correct responses
*indicates questions for which there were few correct responses or attempts.

**Part B: Language background questions**
**Mathematical:**

1m. I want you to tell me what goes on in your head when you add two two-digit numbers?

2m. I want you to tell me what goes on in your head when you add fractions together.

3m. What do we mean when we talk about squaring numbers?

4m. What do brackets ( ) mean when they appear in an expression or equation?

5m. One story for the number 14 is 6 + 8. What can you tell me about other possible stories for 14?

6m What do you think is meant by the statement "the difference between two numbers is 4'?

**General:**

1g. What is the first thing you think of when I say:
one thousand

yellow

ball

tiger

square

three quarters

adjective

2g. Follow up with, “Can you tell me anything more?”

3g. How would you help someone who knew nothing about it, get onto the Internet? What would you say?

4g. If you had a casual job, how would you go about working out what you would earn in a month?

5g. How would you go about finding the next number in a sequence of numbers?

# APPENDIX C: RASCH CASE AND ITEM ESTIMATES

## Appendix C1: Rasch case estimates for all survey participants

| Student ID | No. correct | Total Items | Ability | Error | Infit | Outfit | Infit t | Outfit t |
|---|---|---|---|---|---|---|---|---|
| 390101 | 33 | 40 | 2.32 | 0.5 | 1.07 | 0.72 | 0.33 | 0.03 |
| 390102 | 19 | 40 | -0.17 | 0.39 | 1.05 | 1.34 | 0.33 | 0.86 |
| 390104 | 28 | 40 | 1.27 | 0.43 | 1.16 | 0.89 | 0.74 | 0.01 |
| 390107 | 24 | 40 | 0.6 | 0.4 | 0.59 | 0.44 | -2.47 | -1.4 |
| 390109 | 35 | 40 | 2.87 | 0.56 | 1.32 | 1.39 | 0.97 | 0.7 |
| 390110 | 25 | 40 | 0.76 | 0.4 | 1.06 | 1.13 | 0.37 | 0.42 |
| 390111 | 30 | 40 | 1.65 | 0.45 | 0.82 | 0.57 | -0.65 | -0.47 |
| 390112 | 35 | 40 | 2.87 | 0.56 | 1.1 | 0.98 | 0.41 | 0.43 |
| 390113 | 28 | 40 | 1.27 | 0.43 | 1.19 | 3.51 | 0.86 | 2.8 |
| 390114 | 30 | 40 | 1.65 | 0.45 | 1.14 | 1.04 | 0.6 | 0.29 |
| 390115 | 22 | 40 | 0.29 | 0.39 | 0.93 | 1.17 | -0.38 | 0.52 |
| 390116 | 23 | 40 | 0.44 | 0.39 | 1.35 | 1.75 | 1.78 | 1.55 |
| 390117 | 31 | 40 | 1.86 | 0.46 | 1.09 | 0.83 | 0.43 | 0.05 |
| 390118 | 14 | 40 | -0.94 | 0.41 | 0.68 | 0.47 | -1.83 | -1.08 |
| 390120 | 17 | 40 | -0.47 | 0.39 | 0.93 | 0.75 | -0.37 | -0.47 |
| 390121 | 15 | 40 | -0.78 | 0.4 | 0.9 | 1.11 | -0.52 | 0.37 |
| 390122 | 30 | 40 | 1.65 | 0.45 | 0.92 | 0.79 | -0.23 | -0.07 |
| 390123 | 27 | 40 | 1.1 | 0.42 | 0.97 | 0.88 | -0.07 | -0.04 |
| 390124 | 32 | 40 | 2.08 | 0.48 | 0.9 | 0.54 | -0.29 | -0.33 |
| 390125 | 33 | 40 | 2.32 | 0.5 | 1.05 | 1.15 | 0.27 | 0.48 |
| 390126 | 31 | 40 | 1.86 | 0.46 | 0.92 | 0.79 | -0.21 | -0.02 |
| 390127 | 22 | 40 | 0.29 | 0.39 | 1.13 | 1.3 | 0.76 | 0.79 |
| 390128 | 19 | 40 | -0.17 | 0.39 | 0.98 | 0.91 | -0.04 | -0.09 |
| 390129 | 31 | 40 | 1.86 | 0.46 | 0.94 | 1.67 | -0.16 | 0.97 |
| 180201 | 23 | 40 | 0.44 | 0.39 | 0.74 | 0.86 | -1.47 | -0.19 |
| 181202 | 6 | 40 | -2.57 | 0.52 | 0.76 | 0.48 | -0.73 | -0.16 |
| 180203 | 5 | 40 | -2.86 | 0.55 | 1.12 | 5.14 | 0.46 | 2.05 |
| 181209 | 11 | 40 | -1.46 | 0.43 | 0.72 | 0.59 | -1.23 | -0.49 |
| 180205 | 15 | 40 | -0.78 | 0.4 | 0.74 | 0.51 | -1.5 | -1.03 |
| 180204 | 14 | 40 | -0.94 | 0.41 | 0.85 | 1.08 | -0.76 | 0.32 |
| 181208 | 22 | 40 | 0.29 | 0.39 | 1.25 | 1.44 | 1.35 | 1.04 |
| 181207 | 12 | 40 | -1.28 | 0.42 | 1.38 | 1.94 | 1.61 | 1.41 |
| 181206 | 25 | 40 | 0.76 | 0.4 | 1.09 | 0.92 | 0.52 | -0.01 |
| 181210 | 11 | 40 | -1.46 | 0.43 | 0.76 | 1.46 | -1 | 0.82 |
| 181220 | 7 | 40 | -2.31 | 0.5 | 0.81 | 0.55 | -0.59 | -0.18 |
| 181219 | 9 | 40 | -1.86 | 0.46 | 1.16 | 0.82 | 0.64 | 0.03 |
| 181221 | 5 | 40 | -2.86 | 0.55 | 0.99 | 0.6 | 0.08 | 0.11 |
| 181222 | 6 | 40 | -2.57 | 0.52 | 1.17 | 0.83 | 0.62 | 0.24 |
| 181224 | 10 | 40 | -1.65 | 0.44 | 0.77 | 0.52 | -0.9 | -0.55 |
| 181223 | 5 | 40 | -2.86 | 0.55 | 0.79 | 0.44 | -0.56 | -0.07 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 181217 | 13 | 40 | -1.11 | 0.41 | 0.65 | 0.45 | -1.84 | -1.04 |
| 181218 | 10 | 40 | -1.65 | 0.44 | 1.06 | 0.85 | 0.32 | 0.03 |
| 181215 | 10 | 40 | -1.65 | 0.44 | 0.8 | 0.59 | -0.76 | -0.4 |
| 181216 | 8 | 40 | -2.07 | 0.47 | 0.82 | 0.64 | -0.56 | -0.15 |
| 181225 | 17 | 40 | -0.47 | 0.39 | 1.22 | 1.48 | 1.27 | 1.09 |
| 180211 | 9 | 40 | -1.86 | 0.46 | 0.58 | 0.4 | -1.77 | -0.68 |
| 180213 | 4 | 40 | -3.18 | 0.6 | 0.64 | 0.23 | -1.02 | -0.18 |
| 180214 | 3 | 40 | -3.58 | 0.66 | 1.2 | 1.27 | 0.59 | 0.78 |
| 180227 | 12 | 40 | -1.28 | 0.42 | 1.14 | 0.9 | 0.69 | 0.05 |
| 180226 | 11 | 40 | -1.46 | 0.43 | 1.29 | 1.04 | 1.22 | 0.28 |
| 180212 | 11 | 40 | -1.46 | 0.43 | 0.9 | 0.72 | -0.36 | -0.23 |
| 180129 | 18 | 40 | -0.32 | 0.39 | 0.79 | 0.61 | -1.34 | -0.91 |
| 181112 | 11 | 40 | -1.46 | 0.43 | 0.78 | 0.71 | -0.91 | -0.25 |
| 180114 | 22 | 40 | 0.29 | 0.39 | 1.29 | 1.16 | 1.57 | 0.49 |
| 181105 | 19 | 40 | -0.17 | 0.39 | 1.16 | 1.08 | 0.96 | 0.33 |
| 180126 | 20 | 40 | -0.01 | 0.39 | 1.01 | 0.8 | 0.1 | -0.37 |
| 180119 | 23 | 40 | 0.44 | 0.39 | 0.97 | 1.16 | -0.11 | 0.49 |
| 180128 | 15 | 40 | -0.78 | 0.4 | 1.5 | 2.67 | 2.44 | 2.5 |
| 180116 | 29 | 40 | 1.46 | 0.44 | 0.88 | 0.6 | -0.46 | -0.49 |
| 180125 | 28 | 40 | 1.27 | 0.43 | 0.97 | 0.78 | -0.04 | -0.19 |
| 180106 | 5 | 40 | -2.86 | 0.55 | 1.03 | 0.58 | 0.19 | 0.09 |
| 181103 | 18 | 40 | -0.32 | 0.39 | 0.91 | 0.71 | -0.5 | -0.59 |
| 180109 | 17 | 40 | -0.47 | 0.39 | 1.03 | 0.94 | 0.21 | 0.02 |
| 180110 | 12 | 40 | -1.28 | 0.42 | 1.46 | 1.53 | 1.9 | 0.94 |
| 181101 | 25 | 40 | 0.76 | 0.4 | 0.74 | 0.55 | -1.34 | -0.96 |
| 181102 | 21 | 40 | 0.14 | 0.39 | 0.92 | 0.69 | -0.44 | -0.7 |
| 180118 | 18 | 40 | -0.32 | 0.39 | 0.86 | 0.67 | -0.8 | -0.72 |
| 181123 | 31 | 40 | 1.86 | 0.46 | 0.71 | 0.52 | -1.13 | -0.47 |
| 180108 | 16 | 40 | -0.62 | 0.4 | 1.19 | 1.28 | 1.09 | 0.71 |
| 181113 | 17 | 40 | -0.47 | 0.39 | 1.26 | 1.36 | 1.47 | 0.87 |
| 181104 | 6 | 40 | -2.57 | 0.52 | 0.77 | 0.47 | -0.69 | -0.18 |
| 180115 | 14 | 40 | -0.94 | 0.41 | 1.35 | 1.52 | 1.67 | 1.02 |
| 180127 | 12 | 40 | -1.28 | 0.42 | 0.93 | 0.68 | -0.27 | -0.38 |
| 180117 | 13 | 40 | -1.11 | 0.41 | 0.96 | 0.71 | -0.13 | -0.38 |
| 180107 | 17 | 40 | -0.47 | 0.39 | 1.47 | 2.31 | 2.47 | 2.3 |
| 180122 | 32 | 40 | 2.08 | 0.48 | 1.35 | 3.98 | 1.23 | 2.27 |
| 180302 | 6 | 40 | -2.57 | 0.52 | 0.47 | 0.22 | -2.01 | -0.63 |
| 181304 | 5 | 40 | -2.86 | 0.55 | 1.2 | 2.4 | 0.67 | 1.2 |
| 181305 | 8 | 40 | -2.07 | 0.47 | 1 | 2.76 | 0.09 | 1.62 |
| 181316 | 8 | 40 | -2.07 | 0.47 | 0.99 | 0.79 | 0.06 | 0.05 |
| 181310 | 10 | 40 | -1.65 | 0.44 | 0.92 | 0.75 | -0.23 | -0.12 |
| 181314 | 8 | 40 | -2.07 | 0.47 | 0.8 | 0.69 | -0.66 | -0.08 |
| 181301 | 3 | 40 | -3.58 | 0.66 | 0.8 | 0.26 | -0.36 | 0.11 |
| 180308 | 5 | 40 | -2.86 | 0.55 | 0.6 | 0.24 | -1.29 | -0.38 |
| 180309 | 5 | 40 | -2.86 | 0.55 | 0.86 | 0.49 | -0.32 | -0.01 |
| 180312 | 7 | 40 | -2.31 | 0.5 | 0.58 | 0.35 | -1.58 | -0.52 |
| 180307 | 3 | 40 | -3.58 | 0.66 | 1.03 | 1.49 | 0.21 | 0.87 |
| 180306 | 11 | 40 | -1.46 | 0.43 | 0.69 | 0.54 | -1.38 | -0.61 |
| 180318 | 2 | 40 | -4.09 | 0.78 | 1.05 | 6.59 | 0.28 | 1.88 |
| 181311 | 7 | 40 | -2.31 | 0.5 | 0.77 | 0.44 | -0.75 | -0.36 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 181315 | 5 | 40 | -2.86 | 0.55 | 0.63 | 0.25 | -1.15 | -0.36 |
| 181303 | 6 | 40 | -2.57 | 0.52 | 0.42 | 0.2 | -2.27 | -0.68 |
| 191304 | 5 | 40 | -2.86 | 0.55 | 1.03 | 0.63 | 0.2 | 0.14 |
| 190312 | 8 | 40 | -2.07 | 0.47 | 1.34 | 0.86 | 1.17 | 0.15 |
| 191303 | 8 | 40 | -2.07 | 0.47 | 0.65 | 0.39 | -1.33 | -0.57 |
| 190307 | 8 | 40 | -2.07 | 0.47 | 1.04 | 0.86 | 0.23 | 0.14 |
| 191302 | 7 | 40 | -2.31 | 0.5 | 0.96 | 0.67 | -0.02 | -0.02 |
| 191302 | 8 | 40 | -2.07 | 0.47 | 0.98 | 0.91 | 0.04 | 0.2 |
| 190311 | 4 | 40 | -3.18 | 0.6 | 1.37 | 1.27 | 1.02 | 0.69 |
| 190316 | 12 | 40 | -1.28 | 0.42 | 1.03 | 2.11 | 0.23 | 1.58 |
| 191301 | 7 | 40 | -2.31 | 0.5 | 0.92 | 0.74 | -0.16 | 0.06 |
| 190310 | 7 | 40 | -2.31 | 0.5 | 0.99 | 0.69 | 0.08 | 0.01 |
| 191317 | 6 | 40 | -2.57 | 0.52 | 0.86 | 0.61 | -0.36 | 0 |
| 190107 | 39 | 40 | 4.93 | 1.06 | 0.9 | 0.17 | 0.17 | 0.9 |
| 190106 | 36 | 40 | 3.2 | 0.6 | 0.97 | 0.73 | 0.05 | 0.34 |
| 190115 | 24 | 40 | 0.6 | 0.4 | 0.64 | 0.49 | -2.05 | -1.22 |
| 191119 | 26 | 40 | 0.93 | 0.41 | 1.08 | 0.92 | 0.45 | 0.01 |
| 190111 | 34 | 40 | 2.58 | 0.52 | 1.06 | 1.35 | 0.29 | 0.66 |
| 191104 | 29 | 40 | 1.46 | 0.44 | 1.12 | 0.92 | 0.57 | 0.1 |
| 191122 | 23 | 40 | 0.44 | 0.39 | 0.94 | 0.73 | -0.28 | -0.53 |
| 190116 | 37 | 40 | 3.6 | 0.67 | 0.8 | 0.33 | -0.37 | 0.17 |
| 191103 | 22 | 40 | 0.29 | 0.39 | 1.06 | 0.93 | 0.39 | -0.03 |
| 191121 | 21 | 40 | 0.14 | 0.39 | 1.06 | 0.95 | 0.39 | 0.01 |
| 190110 | 21 | 40 | 0.14 | 0.39 | 1.04 | 0.92 | 0.31 | -0.07 |
| 190123 | 19 | 40 | -0.17 | 0.39 | 1.33 | 1.25 | 1.84 | 0.69 |
| 191108 | 29 | 40 | 1.46 | 0.44 | 1.13 | 0.88 | 0.58 | 0.02 |
| 190101 | 30 | 40 | 1.65 | 0.45 | 1.53 | 1.32 | 1.88 | 0.64 |
| 191105 | 30 | 40 | 1.65 | 0.45 | 0.94 | 0.88 | -0.18 | 0.06 |
| 191109 | 35 | 40 | 2.87 | 0.56 | 1 | 0.7 | 0.1 | 0.2 |
| 191120 | 28 | 40 | 1.27 | 0.43 | 1.25 | 1.14 | 1.09 | 0.42 |
| 191112 | 22 | 40 | 0.29 | 0.39 | 0.81 | 0.99 | -1.06 | 0.11 |
| 190118 | 38 | 40 | 4.12 | 0.78 | 0.98 | 0.28 | 0.14 | 0.42 |
| 191117 | 29 | 40 | 1.46 | 0.44 | 0.72 | 0.53 | -1.23 | -0.66 |
| 190102 | 28 | 40 | 1.27 | 0.43 | 1.22 | 1.13 | 0.95 | 0.41 |
| 191113 | 31 | 40 | 1.86 | 0.46 | 0.83 | 0.67 | -0.61 | -0.21 |
| 190114 | 32 | 40 | 2.08 | 0.48 | 1.38 | 2.33 | 1.3 | 1.39 |
| 191204 | 3 | 40 | -3.58 | 0.66 | 0.94 | 0.64 | -0.01 | 0.44 |
| 190212 | 5 | 40 | -2.86 | 0.55 | 0.83 | 0.46 | -0.43 | -0.05 |
| 190209 | 9 | 40 | -1.86 | 0.46 | 1.15 | 0.91 | 0.63 | 0.16 |
| 190216 | 6 | 40 | -2.57 | 0.52 | 0.5 | 0.23 | -1.86 | -0.6 |
| 190217 | 19 | 40 | -0.17 | 0.39 | 0.86 | 0.71 | -0.86 | -0.63 |
| 190206 | 15 | 40 | -0.78 | 0.4 | 1.11 | 0.89 | 0.64 | -0.06 |
| 190203 | 10 | 40 | -1.65 | 0.44 | 1.3 | 1.47 | 1.16 | 0.81 |
| 190205 | 6 | 40 | -2.57 | 0.52 | 1.4 | 1.42 | 1.24 | 0.71 |
| 191214 | 4 | 40 | -3.18 | 0.6 | 1.04 | 0.82 | 0.24 | 0.42 |
| 190215 | 8 | 40 | -2.07 | 0.47 | 1.01 | 0.86 | 0.15 | 0.15 |
| 191208 | 16 | 40 | -0.62 | 0.4 | 0.99 | 0.73 | -0.01 | -0.47 |
| 191201 | 7 | 40 | -2.31 | 0.5 | 1.01 | 0.91 | 0.13 | 0.25 |
| 191210 | 8 | 40 | -2.07 | 0.47 | 0.79 | 0.5 | -0.69 | -0.37 |
| 191218 | 3 | 40 | -3.58 | 0.66 | 1.07 | 0.52 | 0.3 | 0.35 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 191202 | 6 | 40 | -2.57 | 0.52 | 1.04 | 0.7 | 0.24 | 0.11 |
| 191220 | 12 | 40 | -1.28 | 0.42 | 0.63 | 0.43 | -1.85 | -0.98 |
| 191219 | 17 | 40 | -0.47 | 0.39 | 0.96 | 0.74 | -0.17 | -0.49 |
| 191211 | 11 | 40 | -1.46 | 0.43 | 0.8 | 0.6 | -0.83 | -0.47 |
| 191207 | 8 | 40 | -2.07 | 0.47 | 0.89 | 0.7 | -0.29 | -0.07 |
| 191213 | 3 | 40 | -3.58 | 0.66 | 0.6 | 0.17 | -0.98 | -0.03 |
| 281217 | 12 | 40 | -1.28 | 0.42 | 1 | 0.76 | 0.07 | -0.22 |
| 281201 | 5 | 40 | -2.86 | 0.55 | 1.17 | 0.89 | 0.59 | 0.37 |
| 281208 | 4 | 39 | -3.14 | 0.6 | 0.94 | 0.51 | -0.04 | 0.17 |
| 281218 | 8 | 40 | -2.07 | 0.47 | 1.33 | 1.07 | 1.15 | 0.38 |
| 281209 | 6 | 40 | -2.57 | 0.52 | 1.14 | 0.75 | 0.54 | 0.16 |
| 281205 | 5 | 40 | -2.86 | 0.55 | 0.88 | 0.47 | -0.28 | -0.04 |
| 281210 | 3 | 40 | -3.58 | 0.66 | 1.03 | 0.34 | 0.22 | 0.19 |
| 281221 | 9 | 40 | -1.86 | 0.46 | 0.65 | 0.43 | -1.4 | -0.63 |
| 281202 | 4 | 40 | -3.18 | 0.6 | 0.94 | 0.56 | -0.05 | 0.21 |
| 281215 | 4 | 40 | -3.18 | 0.6 | 0.94 | 0.56 | -0.05 | 0.21 |
| 281214 | 4 | 40 | -3.18 | 0.6 | 0.82 | 0.29 | -0.4 | -0.1 |
| 281212 | 5 | 40 | -2.86 | 0.55 | 1.31 | 1.26 | 0.96 | 0.63 |
| 281203 | 5 | 40 | -2.86 | 0.55 | 1.53 | 1.67 | 1.48 | 0.87 |
| 281220 | 1 | 40 | -4.89 | 1.05 | 0.84 | 0.14 | 0.09 | 0.9 |
| 281213 | 12 | 40 | -1.28 | 0.42 | 1.03 | 1.31 | 0.23 | 0.66 |
| 281222 | 9 | 40 | -1.86 | 0.46 | 0.76 | 0.51 | -0.88 | -0.46 |
| 281219 | 9 | 40 | -1.86 | 0.46 | 0.76 | 0.51 | -0.88 | -0.46 |
| 281223 | 10 | 40 | -1.65 | 0.44 | 1.4 | 1.22 | 1.48 | 0.53 |
| 281216 | 5 | 40 | -2.86 | 0.55 | 1.36 | 4.12 | 1.08 | 1.78 |
| 281204 | 5 | 40 | -2.86 | 0.55 | 1.18 | 1.11 | 0.63 | 0.53 |
| 281101 | 15 | 40 | -0.78 | 0.4 | 1.05 | 1.03 | 0.33 | 0.23 |
| 281102 | 26 | 40 | 0.93 | 0.41 | 1.21 | 1.05 | 1.01 | 0.27 |
| 281103 | 28 | 40 | 1.27 | 0.43 | 1.18 | 0.94 | 0.8 | 0.1 |
| 281104 | 14 | 40 | -0.94 | 0.41 | 0.85 | 0.96 | -0.77 | 0.1 |
| 281105 | 9 | 40 | -1.86 | 0.46 | 1.25 | 0.89 | 0.96 | 0.14 |
| 281106 | 28 | 40 | 1.27 | 0.43 | 1.19 | 1.6 | 0.84 | 1.05 |
| 281107 | 12 | 40 | -1.28 | 0.42 | 1.11 | 1.17 | 0.55 | 0.47 |
| 281108 | 27 | 40 | 1.1 | 0.42 | 0.67 | 0.47 | -1.62 | -1.01 |
| 281109 | 23 | 40 | 0.44 | 0.39 | 0.82 | 0.63 | -0.98 | -0.83 |
| 281110 | 28 | 40 | 1.27 | 0.43 | 1.14 | 1.54 | 0.68 | 0.97 |
| 281111 | 11 | 40 | -1.46 | 0.43 | 0.99 | 1.68 | 0.03 | 1.06 |
| 281112 | 15 | 40 | -0.78 | 0.4 | 0.93 | 0.68 | -0.33 | -0.55 |
| 281113 | 18 | 40 | -0.32 | 0.39 | 0.94 | 0.77 | -0.33 | -0.44 |
| 281114 | 13 | 40 | -1.11 | 0.41 | 0.7 | 0.52 | -1.55 | -0.83 |
| 281115 | 18 | 40 | -0.32 | 0.39 | 1.19 | 2.64 | 1.15 | 2.78 |
| 281116 | 27 | 40 | 1.1 | 0.42 | 0.94 | 0.68 | -0.18 | -0.48 |
| 281117 | 27 | 40 | 1.1 | 0.42 | 0.97 | 1.25 | -0.05 | 0.61 |
| 291205 | 16 | 40 | -0.62 | 0.4 | 0.87 | 0.65 | -0.69 | -0.69 |
| 291201 | 18 | 40 | -0.32 | 0.39 | 0.89 | 1.07 | -0.61 | 0.29 |
| 291204 | 11 | 40 | -1.46 | 0.43 | 0.94 | 0.8 | -0.16 | -0.1 |
| 291209 | 21 | 40 | 0.14 | 0.39 | 1.23 | 1.38 | 1.3 | 0.95 |
| 291208 | 11 | 40 | -1.46 | 0.43 | 0.8 | 0.57 | -0.81 | -0.53 |
| 291202 | 15 | 40 | -0.78 | 0.4 | 1.14 | 0.9 | 0.79 | -0.04 |
| 291206 | 12 | 40 | -1.28 | 0.42 | 0.86 | 0.64 | -0.62 | -0.47 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 291207 | 16 | 40 | -0.62 | 0.4 | 0.96 | 0.81 | -0.18 | -0.27 |
| 291203 | 16 | 40 | -0.62 | 0.4 | 1.37 | 2.05 | 1.97 | 1.87 |
| 291301 | 19 | 40 | -0.17 | 0.39 | 0.91 | 0.66 | -0.52 | -0.78 |
| 291302 | 8 | 40 | -2.07 | 0.47 | 1.21 | 0.93 | 0.79 | 0.22 |
| 291303 | 4 | 40 | -3.18 | 0.6 | 0.82 | 0.29 | -0.4 | -0.1 |
| 291304 | 10 | 40 | -1.65 | 0.44 | 1 | 0.75 | 0.09 | -0.12 |
| 291305 | 8 | 40 | -2.07 | 0.47 | 1.03 | 0.7 | 0.19 | -0.07 |
| 291306 | 9 | 40 | -1.86 | 0.46 | 0.77 | 0.52 | -0.84 | -0.44 |
| 291307 | 2 | 40 | -4.09 | 0.78 | 0.83 | 0.21 | -0.15 | 0.37 |
| 291308 | 6 | 40 | -2.57 | 0.52 | 1.22 | 1.09 | 0.76 | 0.47 |
| 291101 | 17 | 40 | -0.47 | 0.39 | 0.91 | 0.71 | -0.49 | -0.57 |
| 291102 | 23 | 40 | 0.44 | 0.39 | 1.21 | 1.04 | 1.12 | 0.25 |
| 291103 | 27 | 40 | 1.1 | 0.42 | 1.22 | 1.03 | 1 | 0.24 |
| 291104 | 20 | 40 | -0.01 | 0.39 | 1.09 | 1.13 | 0.55 | 0.43 |
| 291105 | 24 | 40 | 0.6 | 0.4 | 0.73 | 0.57 | -1.48 | -0.97 |
| 291106 | 25 | 40 | 0.76 | 0.4 | 0.77 | 0.56 | -1.2 | -0.93 |
| 291107 | 23 | 40 | 0.44 | 0.39 | 0.85 | 0.65 | -0.82 | -0.77 |
| 291108 | 32 | 40 | 2.08 | 0.48 | 1.07 | 1.08 | 0.32 | 0.38 |
| 291109 | 30 | 40 | 1.65 | 0.45 | 0.9 | 0.87 | -0.34 | 0.05 |
| 291110 | 28 | 40 | 1.27 | 0.43 | 1.16 | 1.45 | 0.72 | 0.87 |
| 291112 | 36 | 40 | 3.2 | 0.6 | 1.44 | 1.78 | 1.17 | 0.93 |
| 291113 | 24 | 40 | 0.6 | 0.4 | 1.11 | 1.5 | 0.61 | 1.1 |
| 291114 | 18 | 40 | -0.32 | 0.39 | 0.82 | 0.59 | -1.13 | -0.98 |
| 291115 | 31 | 40 | 1.86 | 0.46 | 1.18 | 0.88 | 0.75 | 0.11 |
| 291116 | 34 | 40 | 2.58 | 0.52 | 0.66 | 0.29 | -1.15 | -0.5 |
| 291118 | 31 | 40 | 1.86 | 0.46 | 0.95 | 1.73 | -0.1 | 1.02 |
| 291119 | 34 | 40 | 2.58 | 0.52 | 1.04 | 3.08 | 0.24 | 1.57 |
| 291124 | 30 | 40 | 1.65 | 0.45 | 1.06 | 0.95 | 0.33 | 0.17 |
| 291125 | 32 | 40 | 2.08 | 0.48 | 1.64 | 2.03 | 2.03 | 1.19 |
| 291128 | 31 | 40 | 1.86 | 0.46 | 0.98 | 0.82 | 0 | 0.02 |
| 291129 | 38 | 40 | 4.12 | 0.78 | 0.85 | 0.24 | -0.11 | 0.39 |
| 291130 | 16 | 40 | -0.62 | 0.4 | 1.44 | 1.57 | 2.28 | 1.19 |

## Appendix C2: Rasch item estimates for all survey items

| Item Number | Number correct | Total | Thresholds | Error | Infit mnsq | Outfit mnsq | Infit z | Outfit z |
|---|---|---|---|---|---|---|---|---|
| 1 | 170 | 222 | -2.53 | 0.19 | 1.13 | 1.42 | 1.47 | 0.97 |
| 2 | 172 | 222 | -2.6 | 0.19 | 1.17 | 1.12 | 1.75 | 0.4 |
| 3 | 43 | 222 | 1.59 | 0.21 | 1 | 1.05 | 0.03 | 0.25 |
| 4 | 189 | 222 | -3.27 | 0.21 | 1.3 | 1.55 | 2.36 | 0.93 |
| 5 | 153 | 222 | -1.98 | 0.18 | 1.44 | 2.72 | 4.62 | 3.58 |
| 6 | 77 | 222 | 0.33 | 0.18 | 1.25 | 1.66 | 2.4 | 2.47 |
| 7 | 48 | 222 | 1.38 | 0.21 | 0.92 | 1.81 | -0.73 | 1.91 |
| 8 | 110 | 222 | -0.7 | 0.18 | 0.98 | 0.9 | -0.23 | -0.42 |
| 9 | 123 | 221 | -1.08 | 0.17 | 0.97 | 0.91 | -0.33 | -0.3 |
| 10 | 12 | 222 | 3.56 | 0.34 | 0.95 | 0.53 | -0.11 | -0.13 |
| 11 | 29 | 222 | 2.28 | 0.24 | 1.13 | 1.99 | 0.97 | 1.52 |
| 12 | 108 | 222 | -0.64 | 0.18 | 1.35 | 1.51 | 3.52 | 2.23 |
| 13 | 98 | 222 | -0.34 | 0.18 | 1.22 | 1.31 | 2.24 | 1.45 |
| 14 | 48 | 222 | 1.38 | 0.21 | 0.87 | 0.91 | -1.22 | -0.14 |
| 15 | 53 | 222 | 1.18 | 0.2 | 1.06 | 0.99 | 0.57 | 0.09 |
| 16 | 16 | 222 | 3.17 | 0.3 | 1.11 | 0.67 | 0.58 | -0.1 |
| 17 | 73 | 222 | 0.46 | 0.19 | 1.02 | 0.84 | 0.21 | -0.61 |
| 18 | 83 | 222 | 0.13 | 0.18 | 0.95 | 0.79 | -0.52 | -0.97 |
| 19 | 96 | 222 | -0.27 | 0.18 | 0.91 | 0.81 | -0.97 | -0.94 |
| 20 | 73 | 222 | 0.46 | 0.19 | 0.81 | 0.63 | -2.05 | -1.68 |
| 21 | 106 | 222 | -0.58 | 0.18 | 1.03 | 1.12 | 0.37 | 0.62 |
| 22 | 98 | 222 | -0.34 | 0.18 | 0.83 | 0.86 | -1.9 | -0.65 |
| 23 | 126 | 222 | -1.17 | 0.17 | 0.86 | 0.76 | -1.67 | -1.03 |
| 24 | 183 | 222 | -3.02 | 0.2 | 0.88 | 0.55 | -1.16 | -0.73 |
| 25 | 81 | 222 | 0.2 | 0.18 | 0.94 | 0.9 | -0.56 | -0.39 |
| 26 | 85 | 222 | 0.07 | 0.18 | 1.01 | 0.85 | 0.15 | -0.65 |
| 27 | 186 | 222 | -3.14 | 0.21 | 0.9 | 0.69 | -0.91 | -0.36 |
| 28 | 164 | 222 | -2.33 | 0.18 | 0.83 | 0.6 | -2.08 | -1.03 |
| 29 | 104 | 222 | -0.52 | 0.18 | 1.13 | 1.21 | 1.44 | 1.05 |
| 30 | 118 | 222 | -0.93 | 0.17 | 0.7 | 0.58 | -3.72 | -2.2 |
| 31 | 91 | 222 | -0.12 | 0.18 | 0.92 | 0.78 | -0.83 | -1.11 |
| 32 | 98 | 222 | -0.34 | 0.18 | 0.9 | 0.86 | -1.1 | -0.68 |
| 33 | 106 | 222 | -0.58 | 0.18 | 0.8 | 0.66 | -2.29 | -1.83 |
| 34 | 45 | 222 | 1.51 | 0.21 | 0.78 | 0.42 | -2.07 | -1.78 |
| 35 | 31 | 222 | 2.17 | 0.23 | 1.05 | 0.85 | 0.41 | -0.1 |
| 36 | 23 | 222 | 2.64 | 0.26 | 0.95 | 0.8 | -0.23 | -0.08 |
| 37 | 52 | 222 | 1.22 | 0.2 | 0.85 | 0.52 | -1.42 | -1.61 |
| 38 | 106 | 222 | -0.58 | 0.18 | 1.09 | 1.15 | 1.01 | 0.77 |
| 39 | 17 | 222 | 3.09 | 0.29 | 0.83 | 0.3 | -0.86 | -0.82 |
| 40 | 81 | 222 | 0.2 | 0.18 | 1.08 | 0.99 | 0.87 | 0.01 |

# APPENDIX D: CHAPTER 6 SIGNIFICANCE TESTS

## Appendix D1: $\chi^2$ tests for actual and expected numbers of words by each ability group in each item set

$\chi^2$ Test for verbosity by ability groups and item sets

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Observed Responses | | | | | | | | | | |
| | | | Sets of items in order of average difficulty | | | | | | | |
| | SET 1 | SET 2 | SET 8 | SET 5 | SET 6 | SET 3 | SET 4 | SET 7 | SET 9 | Totals |
| Ability group | | | | Average set difficulty | | | | | | |
| | -1.66 | -1.56 | -0.09 | -0.08 | 0.34 | 0.425 | 1.046 | 1.92 | n/a | |
| LOW | 1009 | 869 | 1020 | 2681 | 864 | 856 | 1337 | 1071 | 3509 | 13216 |
| AVERAGE | 3016 | 1978 | 2146 | 7728 | 2792 | 2312 | 2849 | 1742 | 8821 | 33384 |
| HIGH | 1155 | 1139 | 1516 | 3846 | 1295 | 1242 | 2688 | 890 | 4171 | 17942 |
| Totals | 5180 | 3986 | 4682 | 14255 | 4951 | 4410 | 6874 | 3703 | 16501 | 64542 |
| Expected Responses | | | | | | | | | | |
| LOW | 1061 | 816 | 959 | 2919 | 1014 | 903 | 1408 | 758 | 3379 | 13216 |
| AVERAGE | 2679 | 2062 | 2422 | 7373 | 2561 | 2281 | 3556 | 1915 | 8535 | 33384 |
| HIGH | 1440 | 1108 | 1302 | 3963 | 1376 | 1226 | 1911 | 1029 | 4587 | 17942 |
| | | | | | | | | | | |
| *Overall: p=* | 3.68E-191 | | | | | | | | | |
| df= 27-1 | 26 | | | | | | | | | |
| $\chi^2$ by sets: *p*= | 1.05E-22 | 0.02149 | 5E-16 | 2E-09 | 4E-11 | 0.2145 | 1E-100 | 3E-36 | 4E-12 | |

## Appendix D2: $\chi^2$ tests for actual and expected numbers of words by each ability group

CHI Test for total numbers of words used by ability groups

| Ability group | Number of students | Observed | Expected |
|---|---|---|---|
| LOW | 7 | 13216 | 14574 |
| AVERAGE | 14 | 33384 | 29148 |
| HIGH | 10 | 17942 | 20820 |
| Totals | 31 | 64542 | 64542 |

p= 2.865E-248

# Appendix D3: $\chi^2$ tests for actual and expected numbers of pronouns I and you by each ability group aggregated across all item sets

Comparison of frequencies of *I* and *you* by ability groups over all sets of items

| | Actual | | Total | Expected | | $\chi^2$ by |
|---|---|---|---|---|---|---|
| Ability group | Total *I* | Total *you* | pronouns | *I* | *You* | groups, $p=$ |
| LOW | 446 | 458 | 904 | 504 | 400 | 1.02E-04 |
| AVERAGE | 1374 | 918 | 2292 | 1278 | 1014 | 5.34E-05 |
| HIGH | 621 | 561 | 1182 | 659 | 523 | 2.59E-02 |
| Totals | 2441 | 1937 | 4378 | | | |
| $\chi^2$ over all, $p=$ | 3.20E-04 | 3.93E-05 | | | | |

# Appendix D4: $\chi^2$ tests for actual and expected frequencies of pronouns I and you by each ability group for each set of items

Differences between *I* and *You* frequencies by ability groups on each set of items

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| LOW ABILITY - Actual responses | | | | | | | | | | |
| | SET 1 | SET 2 | SET 8 | SET 5 | SET 6 | SET 3 | SET 4 | SET 7 | SET 9 | Totals |
| | -1.66 | -1.56 | -0.09 | -0.08 | 0.34 | 0.425 | 1.046 | 1.92 | n/a | |
| *I* | 41 | 32 | 21 | 92 | 42 | 29 | 48 | 20 | 121 | 446 |
| *You* | 26 | 36 | 15 | 101 | 23 | 44 | 54 | 32 | 127 | 458 |
| Totals | 67 | 68 | 36 | 193 | 65 | 73 | 102 | 52 | 248 | 904 |
| LOW ABILITY - Expected responses | | | | | | | | | | |
| *I* | 33 | 34 | 18 | 95 | 32 | 36 | 50 | 26 | 122 | 446 |
| *You* | 34 | 34 | 18 | 98 | 33 | 37 | 52 | 26 | 126 | 458 |
| Chi test for Sets 1 -8 (algebra) | | | | | | | | | | |
| $\chi^2$ over all, *p*= | | 0.02 | | | | | | | | |
| $\chi^2$ by sets,*p*= | 0.05221 | 0.70718 | 0.28026 | 0.64303 | 0.01374 | 0.10052 | 0.64547 | 0.11676 | 0.86346 | |
| AVERAGE ABILITY - Actual responses | | | | | | | | | | |
| | SET 1 | SET 2 | SET 8 | SET 5 | SET 6 | SET 3 | SET 4 | SET 7 | SET 9 | Totals |
| | -1.66 | -1.56 | -0.09 | -0.08 | 0.34 | 0.425 | 1.046 | 1.92 | n/a | |
| *I* | 135 | 101 | 87 | 336 | 131 | 76 | 130 | 49 | 329 | 1374 |
| *You* | 44 | 49 | 35 | 220 | 34 | 71 | 105 | 55 | 305 | 918 |
| Totals | 179 | 150 | 122 | 556 | 165 | 147 | 235 | 104 | 634 | 2292 |
| AVERAGE ABILITY - Expected responses | | | | | | | | | | |
| *I* | 107 | 90 | 73 | 333 | 99 | 88 | 141 | 62 | 380 | 1374 |
| *You* | 72 | 60 | 49 | 223 | 66 | 59 | 94 | 42 | 254 | 918 |
| Chi test for Sets 1 -8 (algebra) | | | | | | | | | | |
| $\chi^2$ over all, *p*= | | 5.3054E-12 | | | | | | | | |
| $\chi^2$ by sets,*p*= | 0.00002 | 0.06489 | 0.01042 | 0.81583 | 3.4E-07 | 0.04129 | 0.14761 | 0.00757 | 0.00003 | |
| HIGH ABILITY - Actual responses | | | | | | | | | | |
| | SET 1 | SET 2 | SET 8 | SET 5 | SET 6 | SET 3 | SET 4 | SET 7 | SET 9 | Totals |
| | -1.66 | -1.56 | -0.09 | -0.08 | 0.34 | 0.425 | 1.046 | 1.92 | n/a | |
| *I* | 49 | 47 | 28 | 175 | 38 | 32 | 66 | 29 | 157 | 621 |
| *You* | 28 | 40 | 32 | 120 | 38 | 43 | 122 | 21 | 117 | 561 |
| Totals | 77 | 87 | 60 | 295 | 76 | 75 | 188 | 50 | 274 | 1182 |
| HIGH ABILITY - Expected responses | | | | | | | | | | |
| *I* | 40 | 46 | 32 | 155 | 40 | 39 | 99 | 26 | 144 | 621 |
| *You* | 37 | 41 | 28 | 140 | 36 | 36 | 89 | 24 | 130 | 561 |
| Chi test for Sets 1 -8 (algebra) | | | | | | | | | | |
| $\chi^2$ over all, *p*= | | 5.14133E-06 | | | | | | | | |
| $\chi^2$ by sets,*p*= | 0.05115 | 0.78150 | 0.36242 | 0.01963 | 0.65769 | 0.08690 | 1.7E-06 | 0.43927 | 0.11450 | |

## Appendix D5: $\chi^2$ tests for differences between actual and expected response types for each set of items, ability groups aggregated

| | | Aggregated response types | | |
|---|---|---|---|---|
| Actual | | | | |
| Sets | Set difficulty | Specific | General | Total |
| 1 | -1.66 | 18 | 30 | 48 |
| 2 | -1.56 | 28 | 22 | 50 |
| 5 | -0.09 | 50 | 43 | 93 |
| 6 | 0.34 | 35 | 6 | 41 |
| 3 | 0.425 | 17 | 31 | 48 |
| 4 | 1.046 | 49 | 28 | 77 |
| 7 | 1.92 | 26 | 13 | 39 |
| Totals | | 223 | 173 | 396 |
| Expected | | Specific | General | Total |
| 1 | -1.66 | 28 | 21 | 49 |
| 2 | -1.56 | 29 | 22 | 51 |
| 5 | -0.09 | 55 | 40 | 95 |
| 6 | 0.34 | 24 | 18 | 42 |
| 3 | 0.425 | 28 | 21 | 49 |
| 4 | 1.046 | 45 | 33 | 79 |
| 7 | 1.92 | 23 | 17 | 40 |
| $\chi^2 = 1\times 10^{-5}$ | | | | |

## Appendix D6: $\chi^2$ tests for differences between actual and expected response types by each ability group aggregated over all item sets

$\chi^2$ Test for response types by ability groups

| | | | Response type | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Specific | | | General | | | |
| Ability group | Affective | One item | Multiple items | Description | Rule | Explanation | Total responses | Total math responses |
| High | 12 | 25 | 35 | 7 | 26 | 39 | 144 | 132 |
| Average | 55 | 72 | 47 | 18 | 14 | 24 | 230 | 175 |
| Low | 26 | 30 | 14 | 18 | 16 | 11 | 115 | 89 |
| Totals | 93 | 127 | 96 | 43 | 56 | 74 | 489 | 396 |
| Expected (maths only) | | | | | | | | |
| High | | 42 | 32 | 14 | 19 | 25 | 163 | 132 |
| Average | | 56 | 42 | 19 | 25 | 33 | 216 | 175 |
| Low | | 28.54 | 21.58 | 9.66 | 12.59 | 16.63 | 109.90 | 89.00 |
| $\chi^2 : p =$ | | 0.00293 | 0.17953 | 0.00410 | 0.01445 | 0.00188 | | |
| | | | | | | | $\chi^2$ by groups; $p$= | |
| Expected (all response types) | | | | | | | | |
| High | 27 | 37 | 28 | 13 | 16 | 22 | 0.000006 | |
| Average | 44 | 60 | 45 | 20 | 26 | 35 | 0.01 | |
| Low | 22 | 30 | 23 | 10 | 13 | 17 | 0.02 | |
| $\chi^2 : p =$ | 0.00211 | 0.03634 | 0.08476 | 0.01151 | 0.00264 | 0.00006 | | |

## Appendix D7: $\chi 2$ tests for differences between actual and expected specific and general responses to Sets 1 to 7 and Set 9 by ability groups

| Actual | | | Response Type | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Ability group | Set 9 | Specific<br>Sets 1-7 | Total | Set 9 | General<br>Sets 1-7 | Total | Total responses | | |
| LOW | 25 | 44 | 69 | 34 | 45 | 79 | 148 | | |
| AVERAGE | 71 | 119 | 190 | 68 | 56 | 124 | 314 | | |
| HIGH | 37 | 60 | 97 | 59 | 67 | 126 | 223 | | |
| Totals | 133 | 223 | 356 | 161 | 173 | 334 | 690 | | |
| Expected | Set 9 | Specific<br>Sets 1-7 | Total | Set 9 | General<br>Sets 1-7 | Total | Total responses | Specific *p*= | General *p*= |
| LOW | 29 | 48 | 76 | 35 | 37 | 72 | 148 | 0.389 | 0.194 |
| AVERAGE | 61 | 101 | 162 | 73 | 79 | 152 | 314 | 0.028 | 0.008 |
| HIGH | 43 | 72 | 115 | 52 | 56 | 108 | 223 | 0.091 | 0.077 |
| Totals | 132 | 221 | 353 | 159 | 171 | 331 | 685 | 0.01 | 0.003 |
| *p*= | 0.214 | 0.068 | 0.015 | 0.517 | 0.0054 | 0.0115 | | | |

## Appendix D8: $\chi^2$ tests for numbers of actual and expected hedges by each ability group in Sets 1 to 8 and Set 9

$\chi^2$ test for actual and expected hedges by ability groups (excluding attribution shields)

Actual

| | Sets 1-8 | | Set 9 | | |
|---|---|---|---|---|---|
| Ability groups | Responses | Hedges | Responses | Hedges | All responses |
| LOW | 457 | 564 | 183 | 184 | 640 |
| AVERAGE | 787 | 2040 | 291 | 525 | 1078 |
| HIGH | 688 | 658 | 222 | 230 | 910 |
| Totals | 1932 | 3262 | 696 | 939 | 2628 |

Expected

| | Sets 1-8 | Set 9 | By ability groups across sets |
|---|---|---|---|
| Ability groups | Hedges 1-8/all | Hedges 9/9 | *p*= |
| LOW | 794 | 229 | 3.71E-18 |
| AVERAGE | 1338 | 385 | 3.602E-93 |
| HIGH | 1130 | 325 | 7.630E-51 |
| Totals | 3262 | 939 | |
| *By algebra and background:*<br>*p*= | 6.1003E-138 | 1.08793E-19 | |

## Appendix D9: $\chi^2$ tests for actual and expected numbers of attribution shields

| | Numbers of attribution shields for each set, by ability groups | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Actual | Set 1 | Set 2 | Set 8 | Set 5 | Set 3 | Set 6 | Set 4 | Set 7 | Set 9 | Totals |
| | -1.695 | -1.563 | -0.09 | -0.08 | 0.11 | 0.38 | 0.92 | 1.92 | | |
| Ability groups | | | | | | | | | | |
| HIGH | 16 | 28 | 11 | 76 | 9 | 16 | 34 | 11 | 50 | 251 |
| AVERAGE | 87 | 61 | 34 | 149 | 65 | 73 | 75 | 27 | 114 | 685 |
| LOW | 21 | 24 | 9 | 51 | 14 | 22 | 39 | 18 | 46 | 244 |
| Totals | 124 | 113 | 54 | 276 | 88 | 111 | 148 | 56 | 210 | 1180 |
| Expected | | | | | | | | | | |
| HIGH | 26 | 24 | 11 | 59 | 19 | 24 | 31 | 12 | 45 | 251 |
| AVERAGE | 72 | 66 | 31 | 160 | 51 | 64 | 86 | 33 | 122 | 685 |
| LOW | 26 | 23 | 11 | 57 | 18 | 23 | 31 | 12 | 43 | 244 |
| Totals | 124 | 113 | 54 | 276 | 88 | 111 | 148 | 56 | 210 | 1180 |
| $\chi$2 sets: *p*= | 0.018 | 0.609 | 0.717 | 0.038 | 0.007 | 0.163 | 0.143 | 0.102 | 0.522 | |
| $\chi$2 Sets 1-8 : *p*= | 0.005 | | $\chi$2 all sets: *p*= | | 0.0009 | | | | | |

## Appendix D10: $\chi^2$ tests for actual and expected numbers of hedges by sets of items (excluding attribution shields)

| Set | Difficulty estimate | Responses | Actual hedges | Expected hedges | Differences |
|---|---|---|---|---|---|
| SET 1 | -1.66 | 237 | 473 | 379 | 94 |
| SET 2 | -1.56 | 181 | 385 | 289 | 96 |
| SET 8 | -0.09 | 113 | 188 | 181 | 7 |
| SET 5 | -0.08 | 562 | 976 | 898 | 78 |
| SET 6 | 0.34 | 205 | 272 | 328 | -56 |
| SET 3 | 0.425 | 245 | 313 | 392 | -79 |
| SET 4 | 1.046 | 259 | 433 | 414 | 19 |
| SET 7 | 1.92 | 130 | 222 | 208 | 14 |
| SET 9 | n/a | 696 | 939 | 1113 | -174 |
| | Totals | 2628 | 4201 | 4201 | |
| | $\chi^2$: *p*= | 2.00506E-21 | | | |