DOCUMENT RESUME

ED 326 544 TM 015 767

AUTHOR Cheung, K. C.; And Others

TITLE Meaningful Measurement in the Classroom Using the Rasch Model: Some Exemplars.

INSTITUTION Institute of Education (Singapore).

REPORT NO ISBN-9971-953-80-3; ISSN-0218-0839

PUB DATE Oct 90

NOTE 64p.

PUB TYPE Collected Works - General (020) -- Reports - Evaluative/Feasibility (142) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS *Classroom Techniques; *Computer Science; *Educational Assessment; Educational Research; *Evaluation Methods; Foreign Countries; Higher Education; *Item Response Theory; Language Skills; Mathematics Skills; *Measurement Techniques; Two Year College Students

IDENTIFIERS Computer Anxiety; *Rasch Model; Rating Scale Analysis

ABSTRACT

This collection of research papers considers extensions of the Rasch model. In Chapter 1, K. C. Cheung explains how measurement and testing in the classroom can be made more meaningful, particularly through the use of Rasch family item response models. In Chapter 2, K. W. Koon and K. C. Cheung apply the Rasch model in constructing a part-whole concept continuum in mathematics. In Chapter 3, using a different conceptual framework, S. K. Cheung and M. L. Choo use the Rasch model to establish a proposed hierarchy of language skills. In Chapter 4, M. L. Choo and K. C. Cheung use Rating Scale Analysis--an extension of the Rasch model--to check on a newly constructed computer programming anxiety instrument normed on a sample of computer science students in junior colleges in Singapore. These papers demonstrate that it is possible to base qualitative interpretation on quantitative measurement. Chapters 2, 3, and 4 show that meaningful measurement is possible in the classroom. Five figures and two tables present supporting data. A 40-item list of references is attached. A multiple-choice test on fractions, and the computer programming anxiety scale are included. (SLD)

Meaningful Measurement
in the Classroom Using
the Rasch Model:
Some Exemplars

by

K C CHEUNG, KOH Wee Koon,
SOH Kay Cheng & MOOI Lee Choo

ISBN 9971-953-80-3

20-21 October 1990

ERU/16/90

Preface

In a 1970 paper, Professor Benjamin Bloom compared the power of testing on human affairs with that of atomic energy. Somewhat exaggerated, perhaps, in the comparison, but the implications are clear - hence the need for *meaningful* testing and measurement, which is the thrust of this collection of research papers. In the same paper, Professor Bloom commented that "The problems that are most alive in measurement today [ie. in 1970] are the search for better units (hopefully with properties akin to physical measurement units), the search for a parsimonious measurement system that will account for the variance of a large number of variables or measures, and the search for improved methods of sampling characteristics and individuals".

In fact, since 1970, the technology of measurement has advanced on several fronts; so has the theory of measurement. Among the theoretical advances is what Frederic Lord has called *item response theory* (or IRT), of which the Rasch model is a special case. Looking back on developments in the testing field over the last twenty years or so, as Jaeger did, he called the 1970s the *"decade of criterion-referenced measurement"* and the 1980s the *"decade of IRT methods"*.

So it is fitting that this ERU collection of research papers should plug into the technology of the 1980s by using IRT, particularly different extensions of the Rasch model. In Chapter 1, Dr K.C.Cheung explains how measurement can be made more *meaningful*, while in Chapter 2, Miss Koh Wee Koon and Dr K.C.Cheung apply the Rasch analysis in constructing a part-whole concept continuum in Mathematics. In the same way but using a different conceptual framework, Dr Soh Kay Cheng and Miss Mooi Lee Choo in Chapter 3 employ the Rasch to establish a proposed hierarchy of language skills. And in Chapter 4, Miss Mooi Lee Choo and Dr K.C.Cheung use the Rating Scale Analysis, an extension of the Rasch, to check on a computer programming anxiety instrument, newly constructed and normed on a sample of computer science students in junior colleges in Singapore in 1989. In this study, as in the other two, it is shown that it is possible to base *qualitative interpretation* on *quantitative measurement* - and this is the direction that the Rasch is heading in the area of application. The three studies demonstrate that *meaninful measurement*, in the way defined by Dr K.C.Cheung in Chapter 1, is feasible and practicable in the classroom.

The collaborative effort between staff and postgraduate students, from which Chapters 2-4 have stemmed, is most noteworthy, and I am grateful that the authors have decided to share their findings and interpretations with a larger readership through this publication.

5 October 1990

Ho Wah Kam
Head, ERU

CONTENTS

CHAPTER 1

Bamboo Stems and Pig-tails: Some Thoughts on Meaningful Measurement and Testing in the Classroom

K C CHEUNG

Introduction

A number of countries such as England and Wales are undergoing radical changes in their national curriculum. In England and Wales, the programmes of study for the four key stages of development comprise attainment targets and the associated statements of attainment. Across subject areas and grade levels, these statements of attainment are graded into ten progressive levels. New conceptions of measurement and testing are urgently required. These curricular innovations demand not only quantitative measurement on pupils' understanding but also that these measures are capable of qualitative interpretation that should be firmly rooted in the educational objectives as intended in the various cross-linked programmes of study. This paper does not provide a comprehensive solution to how pupils can be assessed within this type of national curriculum. It seeks to provide some initial thoughts on how measurement and testing can be meaningful. A revised conception in the use of measurement theories to make possible this type of measurement is discussed. The ensuing chapters provide concrete examples on how some of these thoughts can be applied for use in the classrooms.

What is meaningful measurement?

Contemporary learning theories view learning as a meaning construction process within a culture. This view of teaching and learning, known as constructivism, is implicit in a number of recently developed curriculum such as the national science curriculum in England and Wales. Knowledge with understanding is to be valued more than ranked performance in tests and examinations. According to this view, pupils enter the classrooms and bring with them a number of alternative conceptions. These conceptions are organised into cognitive schemes which are often context-dependent in order to help them make sense of the learning experiences. Misconceptions are said to arise when these alternative conceptions are at variance with those of the publicly-mediated knowledge that have been vetted by a community of scholars. Educational researchers already know that in some topic areas there are certain cognitive pathways, barriers, and helpful experiences to bridge personal conceptions and those which the teachers would like to develop in the pupils. Consequently, pupils' understanding in topic areas such as the 'part-whole' concept of fraction or language comprehension can be organised into a number of progressive levels, each of which can be interpreted qualitatively in the light of the educational objectives of the programme of study. In the cognitive psychology literature, the emotional system is perhaps a bit less developed than the cognitive system. However, emotional attributes such as anxiety towards computer programming appears to exhibit qualitatively different progressive anxiety states and traits. It is this view of measurement and testing within a constructivist philosophy that renders the measurement process a meaningful one.

What dilemma are we facing?

Learning outcomes, whether cognitive or affective, are multi-dimensional and multi-faceted in nature. The dilemma facing testing and measurement experts nowadays is that quantitative measurement usually ends up with a measure posited on a unidimensional continuum the legitimacy of which is often questionable. Thus, the more general mathematics ability and its more restricted constituents such as ability in solving 'part-whole' fraction problems can both be regarded as unidimensional. The use of factor-analytic procedures such as one-factor confirmatory factor analysis is one way to illustrate that they are really so.

Mathematically, a continuum is simply a number line and the traditional process of item calibration seeks to eliminate 'deviant' items so as to position those that can fit within certain limits on this continuum. What I would like to argue is that whether the measurement scales are unidimensional or not should be based more on informed human judgement than statistical criteria. Often, in the past, we construct scales to be unidimensional without seriously examining whether the traits to be measured are really so. At best, the continuum as defined by the order and position of the calibrated items is content-referenced, not realising that there are many factors affecting the difficulty of an item. In addition, for some who are more knowledgeable on measurement theories, the scaling procedures are simply used to massage the data. As such, it is an exercise of examining which scaling procedure works best for the measurement data. Whether the model assumptions are satisfied is often the focus of controversy when item response behaviours look weird. As a result, item parameters such as item discrimination and guessing are added, hopefully, to account for the response behaviours.

How to solve this dilemma?

Knowledge with understanding cannot normally be assumed to progress linearly along a conceptual continuum. When pupils engage in conceptual learning, there are false starts and detours before one comes close to the publicly-acknowledged conceptions. Consequently, it is the progression of the lower forms of knowing to the higher forms of knowing that should be considered to be modelled on a continuum for quantitative measurement with qualitative interpretations. Even so, if some conceptual developments do behave in a linearly progressive manner after trimming the detours, the continuum that represents this development should best be regarded as resembling a pig-tail fashioned as a bamboo stem. The different constituents of the trait to be measured intertwined to form the pig-tail. This solves the problem of how broad in scope the measured trait can be. The jointed segments of the pig-tail symbolise the stagewise manner in which the lower forms of knowing progress to the higher forms. This 'Bamboo stem/Pig-tail' metaphor is useful if one also realises that the real bamboo stem is only straight within certain length limits. Consequently, establishing the unidimensionality of a trait may become possible if one focuses on certain segments of a pig-tail. In practice, what this might mean is that we may construct a unidimensional trait of a more focused, limited scope measuring how primary three/four pupils in Singapore use the counting or partition schemes in solving 'part- whole' problems of fractions. Likewise, the different types of state and trait anxieties towards computer programming constitute the different strands of a pig-tail fashioned like a bamboo stem with the joints indicating transition thresholds from lower levels of anxiety to higher levels. The purpose of measurement and the target population to be measured determine which section of the segmented pig-tail is to be constructed.

Traditional views on the item calibration processes as described earlier need to be revised. The revised view emphasised that the item writing process is the key process in the construction of the measurement scale. The context-embedded nature of the test items and problem-solving tasks calls for a diversity of question formats and problem-solving approaches, which very often are inseparable parts of the trait to be constructed. For example, the cloze test is used to gauge pupils' comprehension within the context of the text. All these need to be considered in qualitative measurement. The calibration process is more a check of whether the requirements of a measurement model used in the scaling procedure are met or not. If they are, all the lovely properties of that model will apply to the resulting measur s. Consequently, the measurement model in the scaling procedure is analogous to some measuring tools such as calipers or micrometers in the field of engineering science in that the validity and reliability considerations will centre upon whether the design and technology of the tool satisfy specific needs and purposes. Judgements on the versatility and usability of the caliper or micrometer thus hinge upon whether the objects to be measured conform to the underlying rationale and requirements implicit in the construction of the tool.

What are the requirements of a measurement model?

The meaning of 'conformity' is that we will neither use a caliper nor micrometer to measure temperature or the diameter of a transcontinental oil-pipe. Requirements of targeting, precision, discrimination are important considerations. In the classroom testing situations there are further problems of response readiness and opportunity to learn, both of which highlight the problems of fairness and timing of test administration. Perhaps the most important requirement of all is the notion of objectivity which requires that with respect to some specific population of objects to be measured, the calibration procedure is independent of the calibration sample and the outcome of the measuring process is independent of the constituents defining the scale of the measuring tool. If quantitative measurement is valued, the measurement scale should be a linear one also. The Rasch family item response models are the <u>only</u> measurement models that possess properties satisfying the above-mentioned objectivity and linearity criteria. The ensuing chapters provide examples on how these measurement models can be applied to produce meaningful measurements in the classrooms.

CHAPTER 2

On Meaningful Measurement: Primary School Pupils' Understanding of Part-Whole Concept of Fractions

KOH Wee Koon & K C CHEUNG

Some contemporary views of teaching and learning

Primary pupils experience considerable difficulty in their learning of concepts of fractions. Despite many studies conducted in this topic area, the problems related to how to teach for conceptual understanding appear to have persisted (Nik Pa, 1989; Hope & Owens, 1987; Kerslake, 1986; Behr, et al., 1983; Haseman, 1981; Hart, 1981; Hiebert & Tonnessen, 1978; Novillis, 1976; Payne, 1976). Perhaps one major problem in spite of all these efforts is that the child has been viewed by teachers as a passive rather than an active learner, who should be actively constructing his/her knowledge schemes with understanding.

Many mathematics educators viewed concepts as essentially hierarchically structured. For example, Skemp (1971) thought that concepts of higher order than those a learner already has cannot be communicated to him/her simply by a mathematical definition. Instead, using a suitable collection of examples so as to enable him/her to construct concepts and resolve cognitive conflicts is more viable and fruitful. This view of learning, through the collaborative efforts of many educators, has been gradually developed into the contemporary constructivist view (for a discussion, for example, see von Glasersfeld, 1987). Vygotsky's theory of the Zone of Proximal Development is also of growing popularity (see Rogoff & Wertsch, 1984). The constructivist view of learning sees a teacher's rôle as not dispensing 'truths', but a diagnostician and a facilitator of learning through problem solving and negotiation of meanings. As such, a teacher needs to have an adequate idea of pupils' alternative conceptions and knowledge of how to bring these closer to the intended destinations. This view of pedagogy poses the need for meaningful measurement in the classroom for the monitoring of pupils' progress.

Part-whole concept of fractions

The concepts of fractions are amongst the most complex and important concepts that primary pupils need to grasp. A clear understanding of fractions provides a rich and solid foundation for the learning of other mathematics concepts, such as proportion, inverse and probability. A fraction can take on many more meanings than the commonly-held 'part-whole' concept, viz. a fraction is part of a whole. Kieren's (1976) initial analyses of the rational number construct, which is part of the real number concept, outlined seven basic interpretations. He contended that the 'part-whole' interpretation, which is viewed as a language-generating construct, is fundamental to all other constructs. Consequently, an adequate understanding of rational number concepts should start with this 'part-whole' concept of fractions and relate this to other rational number constructs.

In their Rational Number Project, Behr (et al., 1983) have redefined some of Kieren's categories and subdivided some others. The 'fractional measure' subconstruct addresses the question of how much a quantity is relative to a specified unit of that quantity. This revised 'part-whole' interpretation can be readily understood by primary pupils when the concepts are represented in different forms, viz. continuous (such as length, area and volume) and discrete (such as countable and partitionable objects) models. This 'fractional measure' subconstruct, which is of paramount importance in understanding the relationships between 'unit' and 'non-unit' fractions, demands cognitive and procedural abilities to partition a continuous or discrete quantity into equal-sized regions or subsets. Different types of cognitive schemes such as the counting and partition schemes are involved in construing and constructing the 'part-whole' concept of fractions in the various models of representation.

Some difficulties in learning fractions

A literature review reveals some major difficulties for primary pupils in their learning of fractions. These are briefly summarised below:

1. Despite many children realising the need to partition a quantity to illustrate a fraction as implied by the denominator of a fraction, very few understand that the partitions/subsets should be of equal-size. For those who do understand, some still lack the necessary technical, procedural skills to do the partitioning adequately (Peck & Jencks, 1981; Streefland, 1978).
2. The language used by teachers in naming and denoting a fraction

has created a cognitive barrier especially for lower primary pupils. Some pupils cannot understand words like 'a third' and 'a quarter'. Some tend to associate these names and notations with their previously acquired ordinal and cardinal meanings. As a result, when using discrete objects to represent fractions, the children tend to form groups having the same cardinality as the denominator of the given fraction (Nik Pa, 1989; Bergerson & Herscovics, 1987; Booth, 1987).

3. Children have perceptual difficulties in interpreting diagrams used to represent fractions. Perceptual cues and distractors can both facilitate and hinder children's thinking. Children normally expect no revision and restructuring of the problem context for its solution. Some children cannot anticipate and conceive alternative relationships between the parts and the whole (Lesh, et al., 1987; Payne, 1976; Piaget, et al., 1960).

4. The acquired concepts of fractions lack transfer across problem situations. For example, the ability of pupils to illustrate concepts such as 1/2 and 3/4 may not be generalised to those like 1/5 and 3/5 because the cognitive scheme of the 'part-whole' concept has not been fully developed yet (Peck & Jencks, 1981; Novillis, 1976).

5. Children have difficulties translating between different models of representation such as between the continuous and the discrete models. They are deficient in their understanding about the models and languages needed to represent (describe and illustrate) and manipulate ideas on fractions. For some children, a fraction cannot have a value greater than one. Consequently, they have

difficulties in proceeding to represent proper fractions on a number line. Generating equivalent fractions to the lowest terms of the numerator and denominator such as from 2/4 to 1/2 also poses problems because of the inadequate understanding of the desirable 'units' in the 'part-whole' concept of fractions (Lesh et al.,1987; Truran, 1987; Bright et al., 1988).

Purpose of study

Children use different cognitive strategies such as the counting and partitioning schemes in their construction and application of the 'part-whole' concept of fractions. The purpose of this paper is to introduce a 17-item multiple choice test on this concept of fractions with a well-defined concept hierarchy built into it. Some common misconceptions are also used as distractors. Using Rasch analysis, the calibrated item difficulty measures are found to behave as designed within qualitative, progressive, conceptual bands such that the levels of both conceptual and procedural understanding they represent are firmly rooted in the research literature.

Design of the multiple choice test

Pupils exhibit different levels of understanding of concepts of fractions. The difficulties they encounter have been discussed earlier. Progressive forms of knowing can be identified, ranging from the lower level of recognising names and symbolic notations of fractions, interpreting diagrammatic representations (continuous quantity versus discrete quantity models) of fractions and using diagrams to represent fractions, to the higher level of applying the 'part-whole' concept of fractions to more complex situations involving both the continuous and discrete quantity representation models (see Table 2.1). Within each level, the item difficulties are moderated by the following:

1. the different models of representation (continuous versus discrete);
2. types of fractions (unit versus non-unit);
3. types of transition from one model of representation to another, which may call for different modes of understanding (eg. interpreting a diagrammatic representation versus illustrating/representing a fraction diagrammatically, compare items 4 and 8 in Appendix 2.1);
4. the type of perceptual cues or distractors (eg. arrangement of the number of parts in a diagrammatic representation may be orderly or random, compare items 8 and 13 in Appendix 2.1).

Pupils are expected to use the counting scheme alone, or in combination with the partition scheme, to relate the parts to the whole when solving the questions. The number of shaded parts are counted, and the appropriate unit should be recognised possibly through partitioning patterns into regions and clusters. This test loads on the 'part-whole' concept of fractions only. Hence, it is imperative to include only those pupils who have the opportunity to learn this concept and have not yet received instruction on other concepts of fractions in order not to defeat the purpose of using the intended cognitive schemes that have been built into the test. Consequently, this paper used as an illustration an

Table 2.1. Levels of Understanding

Level	Progressive Forms of knowing	Task	
		Item No.	Descriptions of skills/competencies
4	Applying 'part-whole' concept of fractions to more complex situations involving continuous quantity representation model.	14 15 16 17	Group or partition (mentally) a given geometric region (where perceptual distractors are present) in order to i) name a given fraction, ii) illustrate a given fraction.
3	Applying 'part-whole" concept of fractions to more complex situations involving discrete quantity representation model.	10 11 12 13	Group or partition (physically) a given set of discrete quantity in order to name or illustrate concept of a given fraction where the total number elements in the set is not equal to the denominator of the given fraction. The elements in the set being arranged in i) an orderly manner, ii) a random manner.
2	Interpreting diagramatic representation of fractions and using diagrams to represent fractions.	4 5 6 7 8 9	Write words or symbols for fractions corresponding to diagrammatic representation of unit or non-unit fractions of i) discrete ii) continuous model. Illustrate fractions with diagrams.
1	Recognising names and symbolic notations of fractions.	1 2 3	Translate a given word to symbol or vice-versa for i) unit fractions ii) non-unit fractions.

item calibration sample of 75 primary three pupils (37 boys and 38 girls) of a typical school in a housing estate in Singapore. The timing of test administration was also carefully taken care of.

Construction of part-whole concept continuum

The item responses were scored dichotomously and subjected to Rasch analysis and were shown to satisfy all the model assumptions and requirements. Reliabilities are also excellent (Item separation reliability = 0.98, Person separation reliability = 0.92). Despite the short length of the test and the small size of the calibration sample, the standard errors of measurement were small (items' and persons' equal 0.6 and 1.0 logits approximately) compared with the span of the continuum defining the concept hierarchy, which spans a long range of thirteen logits.

Four progressive levels of understanding are unambiguously discernible. Each pupil can be measured quantitatively on both the linear logit and non-linear total score scales. Thus, norm-referenced measurements can be carried out as usual. In addition, each pupil falls into one of the four qualitative bands indicating his/her levels of understanding on the 'part-whole' concept of fractions. Figure 2.1 indicates how the quantitative achievement measures are located within qualitative bands such that the levels of both conceptual and procedural understanding they represent are firmly rooted in the research literature. The relative distance between the item difficulty and the pupil's ability level also enables us to predict his or her most probable response to that item. This meaningful hierarchical continuum and

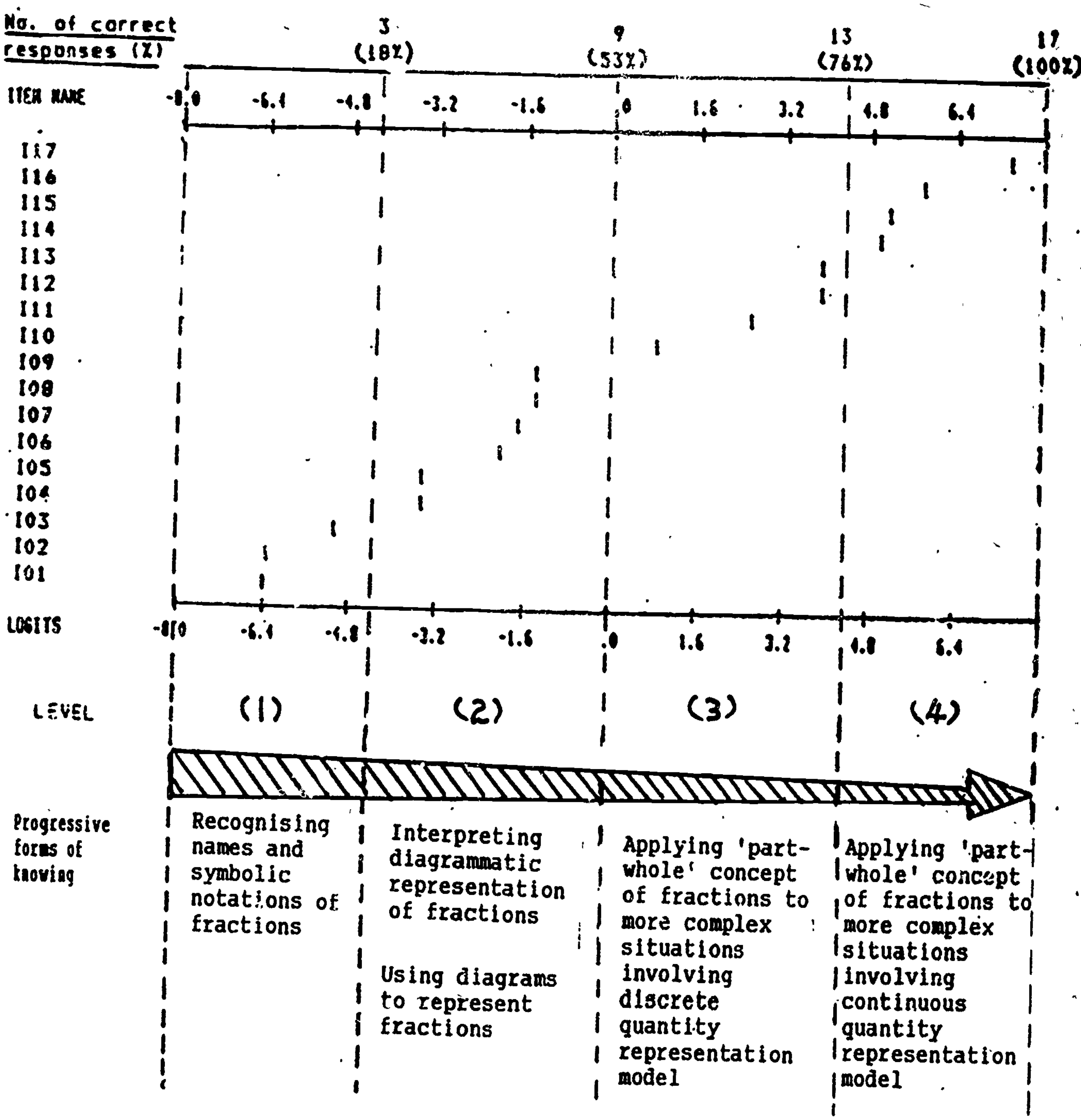


Figure 2.1. Item map of 17 items and its 4 corresponding levels of conceptual understandung of the 'part-whole' concept of fractions.

the excellent fit of data to the Rasch model establish both the external and internal validities of the test.

Quantitative and qualitative measurements

Quantitative measurement is possible so that pupils' performance can be ranked according to the linearised ability logit scores. The zero logit score in this Rasch analysis has been set to the average difficulty level of the items. This scale, if desired, can be set with a maximum score of 100 and passing mark of 50 by a simple linear transformation. Since a pupil's ability score can be mapped onto one of the qualitative bands of the continuum, qualitative measurement is also possible such that pupils' use of the cognitive schemes in solving the range of items of increasing complexity and divergent approaches can be meaningfully explained. It should be noted that the results strictly apply to the sample used in this study, although other cohorts having similar characteristics and same purpose of measurement may also apply. The following are some useful interpretations of four typical pupils whose levels of understanding lie in a different portion of the continuum. Table 2.2 shows the expected success rate of each pupil for each of the levels of understanding.

1. <u>Pupil A (Level of Understanding = 1)</u>

Pupil A (typical ability = -5.0 logits), with level of understanding equals one, has a 81.8% chance of success on the items at this level. He/she should be able to recognise names and symbolic notations of fractions and to translate from one model of

Table 2.2. Expected Success Rate (%) of Pupils A to D

	Pupils (Ability Estimates in Logits)			
	A (-5.0)	B (-1.7)	C (2.6)	D (6.1)
Level of Understanding (Average Item Difficulty Estimates in Logits)	1	2	3	4
1. Recognising names and symbolic notations of fractions. (-6.5)	81.8	99.2	100	100
2. Using diagrams to represent fractions and interpreting diagrammatic representation of fractions. (-1.9)	4.3	55.0	98.9	100
3. Applying 'part-whole' concept of fractions to more complex situations involving discrete quantity representation model. (2.4)	0.06	1.63	55.0	97.6
4. Applying 'part-whole' concept of fractions to more complex situations involving continuous quantity representation model. (5.7)	0.02	0.06	4.3	59.9

representation to another. However, he/she has only a 4.3% chance of success on items at level 2 and less than 1% chance on items at level 3 and 4. With these patterns of probability, we can infer that he/she is probably beginning to understand how to interpret diagrammatic representations of fractions. Consequently, instructions can be given to help him/her construct those knowledge at level 2 but this should not be forced onto him/her unless he/she is ready. Levels of understanding 3 and 4 are beyond him/her at this stage of his/her learning.

2. Pupil B (Level of Understanding = 2)

His/her (typical ability = -1.7 logits) chances of success on items at levels 1 and 2 are 99.2% and 55.0% respectively. Pupil B has almost mastered the ability to recognise names and symbolic notations for fractions and the ability to translate from one model of representation to another. He/she is probably able to interpret diagrammatic representations of fractions and to illustrate fractions using diagrams using both the continuous and discrete models of representation. He/she has only a 1.63% chance of success on items at level 3. He/she is just beginning to proceed to the next phase of applying the 'part- whole' concept to more complex situations involving the discrete quantity model of representation. Instructions should not be forced onto him/her until he/she is more at home with items at level 2. Level 4 is likely beyond his/her understanding.

3. Pupil C (Level of Understanding = 3)

His/her (typical ability = 2.6 logits) chances of success on items at levels 1 and 2 are approaching 100%. These indicate that he/she has mastered the skills and concepts at these two levels. With a 55% chance of success on items at level 3, he/she is currently developing his/her skills and competencies to apply the 'part-whole' concept of fractions to more complex situations involving the discrete quantity representation model. He/she is probably able to group or partition, perhaps by drawing lines to enclose subsets of a given set. This is done in order to name or illustrate concepts of a given fraction, where the total number of elements in the set is not equal to the denominator of the given fraction. Since he/she has only a success rate of 4.3% on items at level 4, instructions should not be forced onto him/her for the time being until he/she is more at home to apply such skills or concepts for items at level 3.

4. Pupil D (Level of Understanding = 4)

His/her (typical ability = 6.1 logits) chances of success for the first three levels are approaching 100%. These imply that he/she has almost perfected his/her understanding on these first three levels. His/her chances of success on items at level 4 is 59.9%. This implies that he/she is currently developing his/her skills and competencies to apply his/her understanding of the 'part-whole' concept of fractions to more complex situations involving the continuous quantity representation model. He/she is probably developing his/her competencies to mentally group and partition a given geometric region, or mentally unpartition a geometric region where perceptual distractors are present in order to name or to illustrate a given fraction.

Conclusion

This paper illustrates how meaningful measurement can be achieved through constructing a concept hierarchy of the 'part-whole' concept of fractions and illustrating how the use of a probabilistic measurement model can help to locate pupils' attainment quantitatively with qualitative interpretations that are firmly rooted in the educational objectives of the course on some concepts of fractions.

From a constructivist perspective, information of this type will enable a teacher to plan appropriate learning strategies so as to facilitate a child's construction of knowledge starting from what he/she has already attained. It is hoped that the use of meaningful measurement will enable us, as educators, to make learning more meaningful and less stressful for our children.

CHAPTER 3

On Meaningful ..asurement: Primary School P~ils' Performance in a Language Test

SOH Kay Cheng & MOOI Lee Choo

> If there is one crucial problem with the application of so-called psychological test theories in education (or in any other context), it is that both test developers and test users just don't know what they are doing.
>
> -- Wilhelm Kempf (1983)

What does it mean to have learned or acquired a language at a level appropriate to one's experiential background? Do different language abilities proceed in progressive stages and are they intertwined? How can we understand measures of language abilities in a meaningful way, taking into consideratio.. their interrelationships? These are some pertinent questions that come readily to mind when we discuss the assessment of language ability in the classroom context.

This paper, it is hoped, serves as an exemplar of an attempt at relating different language abilities within a conceptual framework. It highlights the interrelationships among these abilities for a meaningful interpretation of the results of language assessment. Although the data analysed were collected by using a test of Chinese Language for primary school children, it is believed that the concerns and concepts as discussed are general and hence to a large extent language-free.

Hierarchy of language abilities

It is a truism that in learning a new, foreign or second language, the first task is to acquire a working vocabulary which contains mainly content words and some function words. Observation of beginning language learners will verify this contention. Such word knowledge may not take one very far but here lies the foundation of more advanced and complex learning.

At a somewhat higher level, the language learner begins to emit 'telegraphic' utterances (Ellis, 1984) typically found among learners of foreign languag s or immature younger first language learners (Brown, 1973). Beyond this, the learners (young or old) begin to show the ability to string isolated words or short phrases together in a linguistically acceptable manner, thus evidencing their sensitivity to grammatical rules.

At a still higher level is the ability to make good use of contextual cues and project meanings onto the text, thus showing congruence between the meanings perceived by the reader and those intended by the writer; in short, reading comprehension. This is the level at which most language learners are expected to function effectively for most practical purposes and intent.

These abilities may develop somewhat independently but more likely interactively. When a child grows and gains more experience, the lower level ability becomes more established while ability at a higher level begins to build upon it. Consequently, there is constant feedforward to the higher levels, and a hierarchy of some kind is always maintained in the process of learning the language. This may be simply illustrated in Figure 3.1.

The test

To examine the viability of this contention, we used the Chinese Language Test for Primary Three (Soh, 1983) constructed with the hierarchy posited in Figure 3.1. The three sections of the test are meant to load on the three levels of language ability described earlier. Section A of this test has 30 items for assessing knowledge of Chinese characters. These characters were chosen by systematic sampling of all characters included in the Primary One to Primary Three word list. Pupils listen to a sentence read aloud by the tester who then specifies the character to be identified from a set of four characters printed on the answer sheet. The mental operation involved here is **recognition**, provided with auditory cues. This is a format commonly used for assessing vocabulary at the lower level of language learning.

Section B consists of 10 'scrambled sentences'. Pupils are to re-arrange the isolated characters or words and short phrases to form a meaningful sentence. This task involves the ability of **re-alignment** of the linguistic elements. This is a test of grammatical sensitivity which has been found to be most predictive of foreign and second language aptitude (Wesche, 1981). The last, Section C, comprises three short passages cast in the form of **cloze** procedures, with 30 blanks, requiring **exact response** scoring. This

LEVEL	**LANGUAGE ABILITY**
Higher (Reading comprehension)	Discourse -------------> ^ Re-construction |
Intermediate (Grammatical sensitivity)	Sentence ---------------> ^ Re-alignment |
Lower (Knowledge of characters)	Words ------------------> Recognition

Figure 3.1. Levels of language abilities

task involves **re-construction** of the text provided with only partial information. It assesses the pupils' ability to make use of contextual cues in comprehending a passage at the more complex level of integrated use of language. Although there has been disagreement among language researchers as to what language ability a **cloze** procedure tests, there is also considerable agreement that scores for **cloze** procedures predict language attainment rather well (Soh, 1984).

Calibration

The 70-item test was administered to a sample of 254 primary four pupils from a girls' school of high standing. The responses were submitted for a Rasch analysis. With respect to this sample of pupils, the test was found to be very well-targeted; only one misfitting person and four misfitting items were found. Reliability of the test is also satisfactory; the person separation reliability is .82 and item separation reliability .92. These indicate that the data fitted the Rasch model reasonably well. The logit score provides a linearised quantitative measure of the language ability.

As can be seen from Figure 3.2, notwithstanding a certain degree of overlap among the three sections of the test, most items testing knowledge of characters (Section A) were found at the easy end of the continuum and most items testing reading comprehension (Section C) at the difficult end, with those testing grammatical sensitivity (Section B) found in between these two extremes. This pattern of overlap establishes the plausibility of the postulated theoretical

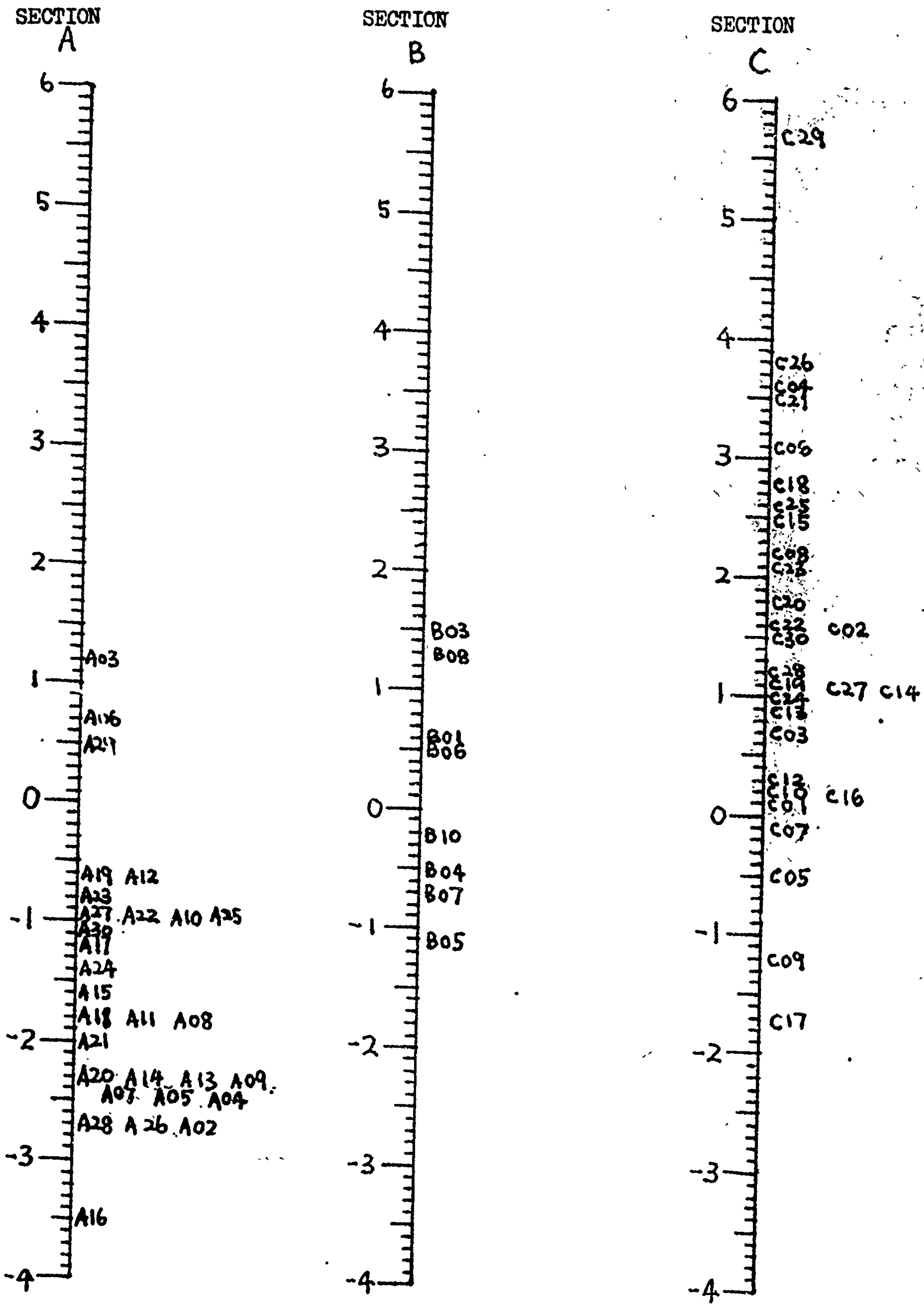


Figure 3.2. Distribution of test items

construct of language ability.

The calibrations of the three sections of the test are summarised in Table 3.1. It is apparent that the test as a whole has calibrations spanning 9.14 logits. The ranges of logits are 4.63 for the knowledge of characters sub-test, 2.55 for grammatical sensitivity sub-test, and 7.45 for reading comprehension sub-test. With respect to the calibration sample, the two thresholds, corresponding to the two transitions from one ability level to the next, are clearly discernible. Grammatical sensitivity appears to serve as a bridge between knowledge of characters and reading comprehension.

Conclusion

In this study, items purportedly measuring three language abilities were found to function to a large extent in the manner consistent with the posited hierarchy of language abilities, from the relatively simple and discrete knowledge of characters, through the intermediate level of grammatical sensitivity, to the more complex and integrated level of reading comprehension. In addition to this, as groups of items measuring a common skill can be taken as homogeneous sub-tests measuring a skill (Griffin, 1985), the sub-tests provide information of three language abilities that are qualitatively different, though intimately related in a progressive manner. Thus, the test can be used as a quantitative assessment of the presumably unitary language ability as well as for a more meaningful, diagnostic interpretation of the pupils' location on the language ability continuum.

Table 3.1. Range of item calibrations (in logits)

Test	Lowest (Easiest)	Highest (Most difficult)
Whole test	-3.45	5.69
Sub-tests:		
Knowledge of characters	-3.45	1.18
Grammatical sensitivity	-1.10	1.45
Reading comprehension	-1.76	5.69

Admittedly, the viability of the theoretical framework may have limitations in the level of pupils sampled and the language used. This points up the importance of paying attention to the purpose of measurement. It is worthy of future attempts to test the generality of this framework to sample pupils from other levels and using other languages. Moreover, language tests have a plethora of item types presumed to measure different language abilities. Whether there is a need for such a wide variety (some rather fanciful) and how efficient they are could also be studied.

The present study has been fortunate in that the test data fitted the Rasch model rather well. It is possible that the results of this study might have turned out the other way. Here, Kempf (1983: 269) wisely cautions thus,

> The fact that latent trait models do not apply ... does not necessarily mean that there is no field of application for them in educational research. But it does mean that latent trait models must not be applied routinely. Statistical theories of psychological test scores cannot substitute for psychological theorizing and for a clearly defined psychological terminology.

Replace **psychological** in the above quote with **educational** or **language**. The message will then become loud and clear to the teachers, the pupils and test developers working with, and for the schools.

<u>Note:</u> The test is available in the Institute of Education Library or from the first author.

CHAPTER 4

On Meaningful Measurement: Junior College Pupils' Anxiety Towards Computer Programming

MOOI Lee Choo & K C CHEUNG

Introduction

This paper seeks to introduce a normal diagnostic instrument on pupils' anxiety in learning computer programming. The nature of anxiety is examined using pupils' internal frames of reference and perceptions of the cause(s) of anxiety. Following this discussion, the concepts of state and trait anxiety are defined and some components of computer programming anxiety are propounded in the light of the computer programming studies in junior colleges in Singapore.

The computer programming anxiety instrument has been constructed and normed on computer science pupils in the junior colleges in Singapore in 1989. The instrument was initially validated through factor analysis and later subjected to Rating Scale Analysis. These procedures illustrate the possibility of having quantitative anxiety measures with well-defined qualitative interpretation that is rooted in the educational objectives of the computer programming course.

We hope to demonstrate the use of this instrument as a means of assessing the computer programming anxiety level of pupils. To end this paper, some ways of overcoming this anxiety are proposed.

The nature of anxiety

Pupils experience anxiety in their learning of computer programming. Take as examples the authentic cases of Violet and Sam. Violet entered junior college hoping to qualify for a university education. Her parents apparently had high expectations of her. She was the eldest of three daughters and the first among her cousins to qualify for a junior college education. She herself felt some pressure as a prospective role model for her sisters and as a symbol of hope and success for her parents. She found that she could not cope with her workload, particularly the cognitive demands of the computer programming course. She had difficulty and no confidence in doing her assignments and was often unable to meet deadlines. She did not want to seek help from her girl friends as she felt that they were also struggling through the computer programming course, while the boys had no time for her. She feared her computer programming instructor and did not want to appear foolish before her. Hence, Violet felt great anxiety towards computer programming.

Sam confessed to being rather poor at computer programming and feeling anxious at times. However, he realised his limitations and was willing to work hard and made visible attempts in trying to understand and complete his computer programming assignments. He was fortunate enough to have his good friend, Kai, to help him; Kai being rather competent with computers. Nevertheless, Sam did feel a fair amount of anxiety whenever a new computer programming assignment had to be done and new concepts learnt. He was worried about debugging. Although he might have done his assignment first, he would hold back until Kai had done his so that Kai could help him with de-ugging. Inadvertently, Sam was holding back his own progress by not seeking help from his computer programming instructor. In the two examples considered, the nature of the threat was known. The first case involved event uncertainty, or uncertainty as to whether the known threat would materialise. The second involved indecision or response uncertainty. Both pupils had one thing in common - an inability to channel the arousal produced by the threat into directed action. According to Epstein (1972), the combination of the diffuse state of arousal and avoidance tendencies evokes a unique state that can be identified as anxiety.

Some definitions of anxiety

Spielberger (1972) used the term anxiety, firstly, to describe an unpleasant emotional state or condition which is characterised by subjective feelings of tension, apprehension and worry, and by activation or arousal of the autonomic nervous system and secondly, to refer to relatively stable individual differences in anxiety proneness as a personality trait. One theory suggested by Phillips, Martin and Meyers (1972) was that Spielberger's two-part conception of anxiety includes what is referred to as "state", "objective" or "situational" anxiety, and what is called "trait", "neurotic" or "chronic" anxiety.

State anxiety (A-State) refers to individual differences in the actual response to a particular stressful situation while trait anxiety (A-Trait) refers to individual differences in anxiety proneness. Moreover, these two types of anxiety interact in a manner such that anxiety proneness (A-Trait) influences the extent of the anxiety response (A-State) (Phillips, Martin & Meyers, 1972). Researchers (Raub, 1981; Rohner & Simonson, 1981; Jordan & Straub, 1982; Maurer, 1983) generally agreed that the concept of computer anxiety fits more into the category of state anxiety than trait anxiety as A-Trait people will exhibit A-State reactions when confronted with computers. Computer programming anxiety appears to comprise several elements such as anxiety

(1) towards the content of programming learning tasks and problem-solving activities;
(2) towards some of the distinctive features of programming as an intellectual activity such as in testing and debugging programmes and in comprehension of error messages;
(3) about the connotative meanings of programming languages in our society such as the use of English words PRINT and RUN which can suggest that the computer system has the human ability to infer what is meant from what is said;
(4) about being evaluated when taking programming tests or doing programming assignments.

Computer programming studies in junior colleges in Singapore

The computer science course intends to assess 45 - 55% on areas of computer theory, algorithms, data structures, programming concepts and languages, and computer elements and architecture. Within this portion of the coursework, emphasis is placed on analysis and design work with laboratory experiences receiving a strong emphasis. In addition, pupils are exposed to a variety of programming languages and systems and must become proficient in at least one higher level language in order to complete a compulsory project.

The remaining percentages for the purpose of assessment cover information systems, data processing, the social and economic impact of computers on individuals, organisations and society, as well as characteristics of hardware elements and computer operations.

The teaching of computer programming in Singapore's junior colleges (equivalent to Grade 12) is typically a group-based, teacher-paced approach in which pupils learn in cooperation with their classmates. Instructional time and curriculum are relatively fixed and the teacher has charge of 12 or fewer pupils for each hands-on session so that no pupil needs to share a work station.

The computer learning environment is interactive in that by testing the computer programme for the given programming problem, the problem solver can get immediate feedback about how effective the programme is in solving the problem and can use this information to modify the programme or to go on to the next step.

Purpose of study

The purpose of the study is to demonstrate how meaningful measurement is possible for use in the classroom. A computer programming anxiety scale was constructed to locate both the state and trait anxieties of junior college pupils in Singapore. This five-point Likert scale was subjected to Rating Scale Analysis which is an extended version of Rasch Analysis. Four meaningful qualitative bands of anxiety levels with increasing state and trait anxieties were defined. Teachers can use this instrument to identify and diagnose pupils with extremely anxious behaviours so that more appropriate instructional strategies can be applied.

The instrument

The Computer Programming Anxiety Scale (CPAS) consists of three subscales, namely 'Errors', 'Significant Others' and 'Confidence' (see Appendix 4.1).

The 'Errors' subscale (items 2, 6, 9, 12 and 17) measures pupils' state of mind when confronted with errors or difficulties while programming. The 'Significant Others' subscale (items 3, 7, 10, 14, 15 and 18) measures how troubled and uneasy pupils are in the presence of more capable computer science pupils or when assessed by their teachers. The 'Confidence' subscale (items 1, 4, 5, 8, 11, 13, 16 and 19) measures pupils' feelings of inadequacy or lack of confidence when doing programming assignments. The former two are state anxiety measures whereas the latter is a trait anxiety measure.

For the state 'Error' and 'Significant Others' scales, the items on a five-point Likert scale from <u>never true</u> to <u>always true</u> are of the form "I am troubled by the number of bugs in my programme" and "I worry about making a fool of myself in front of my friends when I cannot write programmes". Pupils were to indicate how true each of these statements was, based on their feelings at a particular moment in time when the instrument was administered. For the trait

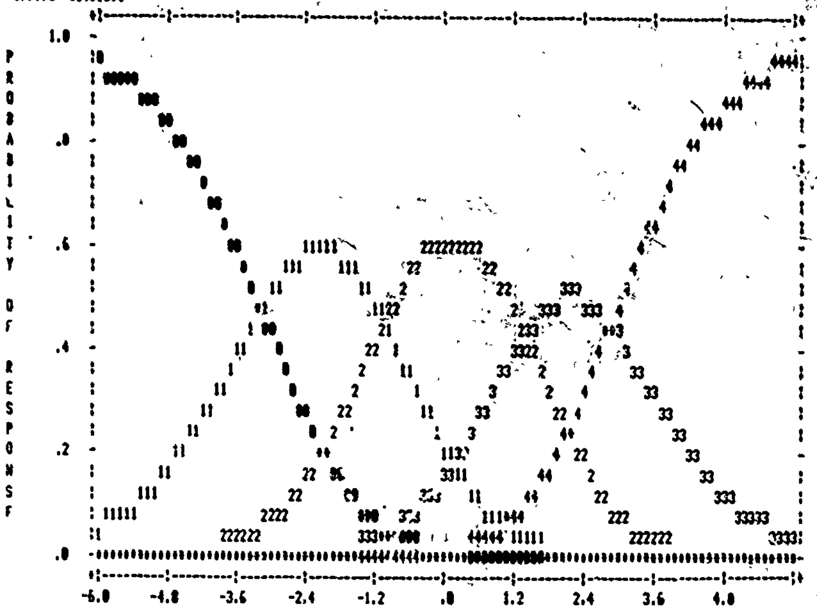
'Confidence' scale, pupils responded to statements such as "If my programme cannot run I do not know what to do next" on a five-point Likert scale from never true to always true. These items were viewed as not being affected by situational stress, and responses towards these items would be stable over time.

Item calibration and person measurement

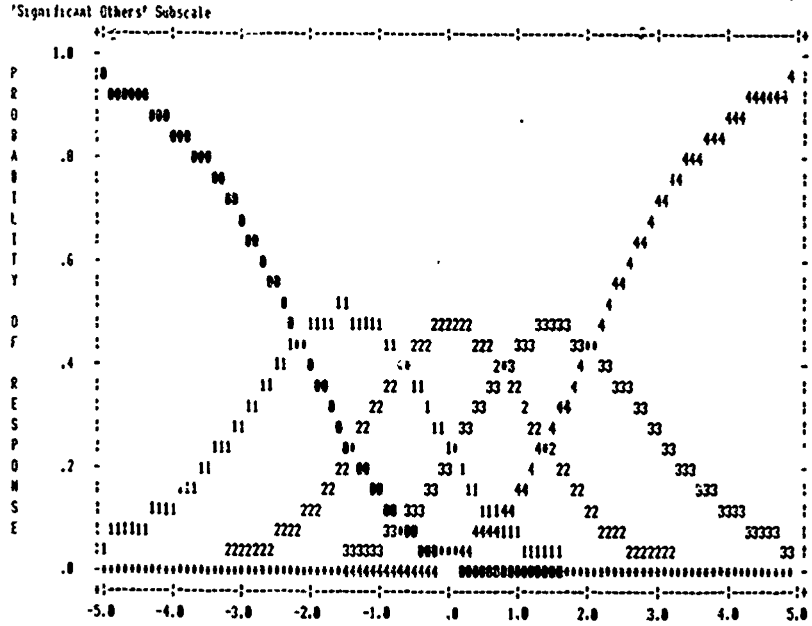
The calibration sample comprised 200 (150 boys, 50 girls) pre-university (Grade 12) computer science pupils from 8 classes in 6 junior colleges in Singapore in 1989. Details of the computer programming anxiety scale construction process and information on the normed sample can be found in Mooi (1990).

Rating Scale Analysis, an item and person calibration procedure, provides an indication of how the response patterns conform to an extended version of the Rasch Model (Wright & Masters, 1982). It is a requirement that the response categories of all items in the same subscale are understood and responded to in the same manner by all pupils in the calibration sample. As such the response category probability curves, which indicate the probability of endorsing each of the response categories, are the same for each item in the same subscale (see Figure 4.1). If the response patterns fit the Rating Scale Model as indicated by both the item and person fit statistics, then the confounding of the person's and item's anxiety level estimates by individual differences in conservatism or acquiescence is not a problem (Masters, 1980). Objective measures specific to the calibration sample or its associated target population such that the item calibration is

'Errors' Subscale



'Significant Others' Subscale



'Confidence' Subscale

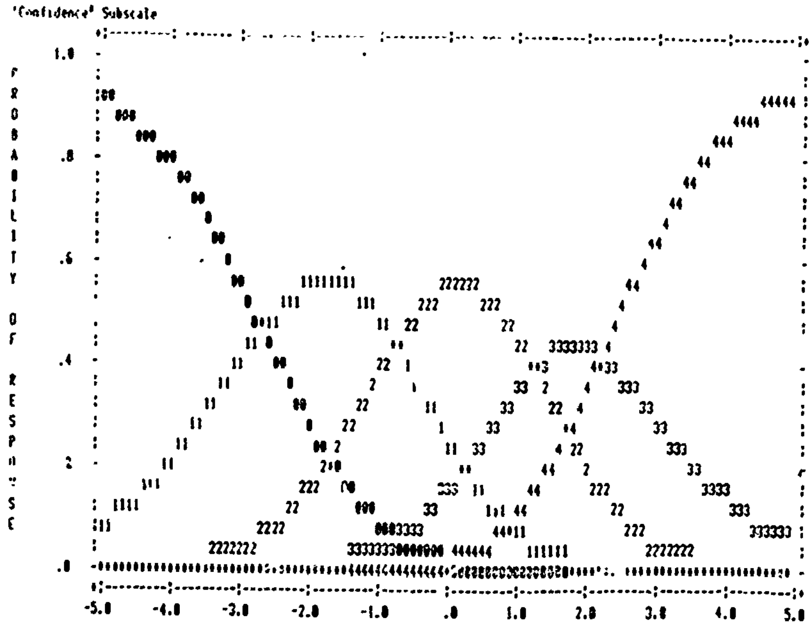


Figure 4.1. Response category probability curves

"person-free" and person measurement is "item-free" are obtained by outcome. Those pupils having zero or full scores on a subscale are not included in the analysis because their anxiety levels are not bracketed by the different categories of the items and hence are not estimable without further information or assumptions.

The three subscales, after removing only a few pupils with some unexpected response behaviours, demonstrate both excellent statistical person and item fit to the analysis model. Despite the small number of items in each subscale, the item and person reliabilities are highly satisfactory, viz. the person separation indices for the three subscales are equal to 0.85 and the item separation indices range from 0.96 to 0.99. The item calibration results and the distribution of pupil's anxiety levels are shown in Figure 4.2. The numbered horizontal lines show a conversion between the total subscale scores and equivalent logit scores. The logits are interval measures with zero centred at the average anxiety level of all the items in the subscale. As such, it is more appropriate than the non-linear total subscale measures resulting from the unequal step sizes of the response categories. The five numbered regions in Figure 4.2 represent the most probable responses to items against the spectrum of anxiety levels of the calibration sample, the central tendency of distribution of which are shown at the bottom of each most probable response curve. For example, a pupil with 'Confidence score 15 (at 0 logit) is likely to give a '1' (<u>Seldom True</u>) to question 1 and '2' (<u>Sometimes True</u>) to all other items in that subscale. Across the three subscales, q01, q17 and q03 are items associated with highest anxiety scale

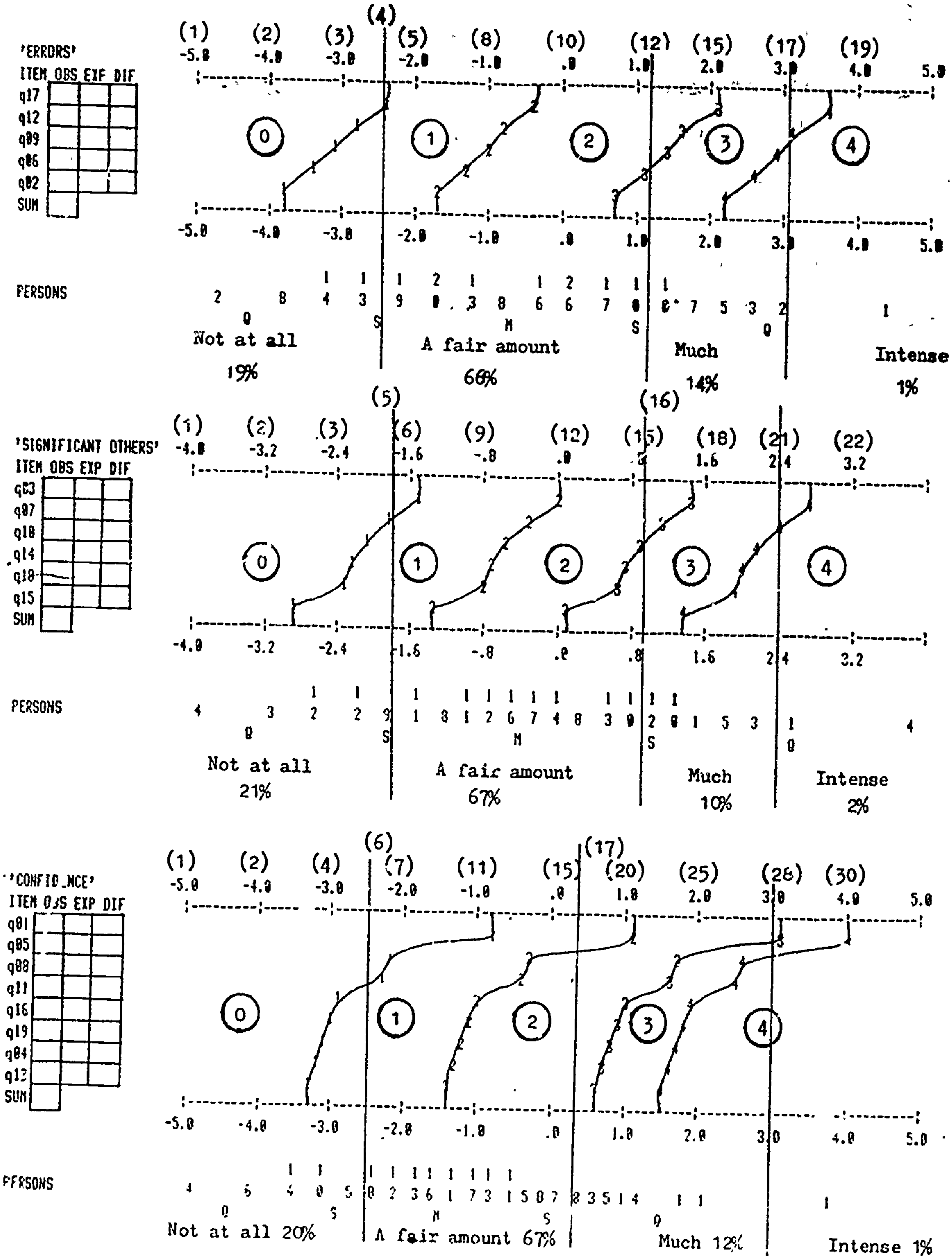


Figure 4.2. Most probable response curves
(Circled number indicates the most probable response for that region)

values whereas q13, q02 and q15 are the lowest. These imply that only the extremely anxious pupils will endorse <u>Always True</u> on the former set of questions and only the least anxious pupils will endorse <u>Never True</u> on the latter set of questions. The set of items and their categories span a continuum of 5 to 7 logits for the three anxiety subscales. This wide span is made possible by the targeted calibration sample and the correct decisions on both the number of steps and evenly-spaced (approximately) step size of categories in bracketing the full range of anxiety levels of the target population.

Not only is quantitative measurement possible, qualitative interpretation is also meaningful by examining how the set of items in a subscale span the anxiety continuum, in view of the anxiety distribution of the calibration sample and the educational objectives of the computer programming course. Four "progressive" bands of computer programming anxiety are discernible, viz. "Not At All", "A Fair Amount", "Much" and "Intense" (see Figure 4.2). This 4-band classification is more robust than the individual most probable response to items since each band size well exceeds the standard error of a person's anxiety estimate (approximately of the order of 0.5 logit). Moreover, the purpose of this instrument is not to rank pupils but to identify extremely anxious behaviours in computer programming. Approximately 20%, 66%, 12% and 2% of the pupils can be classified in this way. Table 4.1 shows how the four bands are interpreted across the three anxiety subscales, after consulting the educational objectives and classroom realities of the computer programming course. Since there are moderately high

Table 4.1. Four bands of computer programming anxiety

Subscale	NOT AT ALL 1	A FAIR AMOUNT 2	MUCH 3	INTENSE 4
CONFIDENCE	The students within this band are familiar or "at-home" with computers. They probably have previous computer programming experience and are confident of their own ability to cope with most programming tasks. They never feel anxious when programmes are complicated and are not confused by the many computer terms.	These students seldom feel nervous or uncomfortable with computers. They are likely to have been exposed to computers and have minimal or no computer programming experience. They are sometimes confused by the many computer terms. They become fairly anxious when programme lines become complicated.	These students do not actually know how the computer operates. They may become stressful when programming problems are beyond their cognitive powers. They are often not confident when writing programmes and always feel that they cannot think properly when programme lines contain errors.	Students are nervous and uncomfortable with computers, would actually suffer from "computer phobia" and avoid computers. Their minds always go blank when faced with programming problems and always cannot concentrate when programme lines become complicated. They always feel that they would not be able to understand programming.
ERRORS	Students seldom worry about debugging, never feel tense or worried when programmes have to be corrected over and over again.	The students are sometimes "bugged" by bugs in their programmes and become fairly worried when their programmes cannot run, and when they have to correct their programmes.	Students often feel troubled by the "bugs" in their programmes and have much anxiety when programmes have to be corrected over and over again.	Students are always tense whenever there are errors in their programmes and are very anxious when their programmes cannot run. They are always worried when programme lines have to be corrected several times.
SIGNIFICANT OTHERS	Students within this band are never worried about being picked by teachers and are seldom troubled by peers out-performing them. They do not feel nervous in the presence of more capable students. No peer pressure and the teacher is not likely to be a threat.	These students are sometimes troubled when their friends could programme and they could not. They are often troubled when friends discuss programming concepts that they do not understand. They are fairly anxious and worry about what teachers think about their programmes. Peer pressure is felt and the teacher poses a threat.	Students are often worried about being picked by teachers for making mistakes in programming problems, and made a fool of in front of their friends when they could not write programmes. They often feel nervous and troubled in the presence of more capable students. Great peer pressure is felt and the teacher is an element of threat.	These students are always worried that their teachers would pick on them or that their teachers would know that they could not write programmes. They are very troubled by the knowledge and programming ability of their more capable friends. Peer pressure is greatest and the teacher is an intimidating figure.

correlations amongst the three anxiety logit measures (Pearson r ranges from 0.61 to 0.73), it is highly probable that a pupil's band classifications are consistent across the three subscales. Consequently, one might not only find some typical anxiety behaviours such as reactions towards peer performance and debugging of computer programmes increasing in frequency and intensity as the band level increases but also convergent anxiety behaviours within the same band across the three subscales. These consistencies and meaningfulness establish the construct validities of the three anxiety subscales whereas the Rating Scale Model ascribes a lot of lovely properties such as specific objectivity and linearity of logit measures and homogeneity of the interpretation of the response categories in a subscale. It is this quantitative measurement with a qualitative interpretation that is firmly rooted in the educational objectives of a computer programming course and classroom situations that renders the process of measurement more meaningful and purposeful than the traditional procedures.

How classroom teachers use this instrument

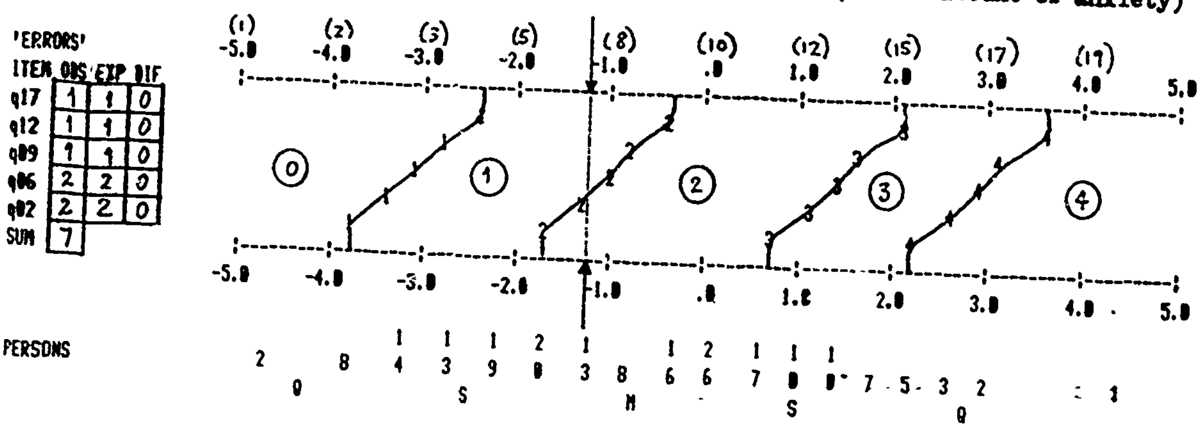
One should note that this instrument has been normed for Junior College pupils in Singapore and hence the calibrations can only be strictly applied to this target population in the computer programming course. The following steps (applicable to each subscale) are useful for teachers in order to use the instrument (please refer to Figure 4.3 for an illustration):

1. Enter the pupil's item responses (0-4) under Column OBS.
2. Add up the responses under OBS and enter the sum under SUM.

Logit Measure = -1.22 Band Level = 2 (A fair amount of anxiety)

'ERRORS'

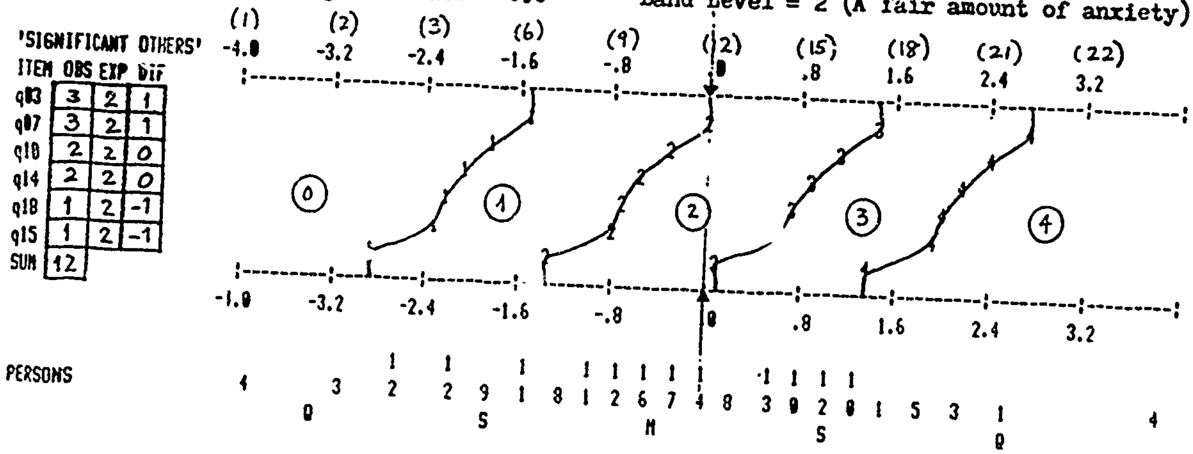
ITEM	OBS	EXP	DIF
q17	1	1	0
q12	1	1	0
q09	1	1	0
q06	2	2	0
q02	2	2	0
SUM	7		



Logit Measure = 0.0 Band Level = 2 (A fair amount of anxiety)

'SIGNIFICANT OTHERS'

ITEM	OBS	EXP	DIF
q03	3	2	1
q07	3	2	1
q10	2	2	0
q14	2	2	0
q18	1	2	-1
q15	1	2	-1
SUM	12		



Logit Measure = -2.0 Band Level = 2 (A fair amount of anxiety)

'CONFIDENCE'

ITEM	OBS	EXP	DIF
q01	1	0	1
q05	1	0	1
q08	1	1	0
q11	1	1	0
q16	0	1	-1
q19	1	1	0
q04	1	1	0
q13	1	1	0
SUM	7		

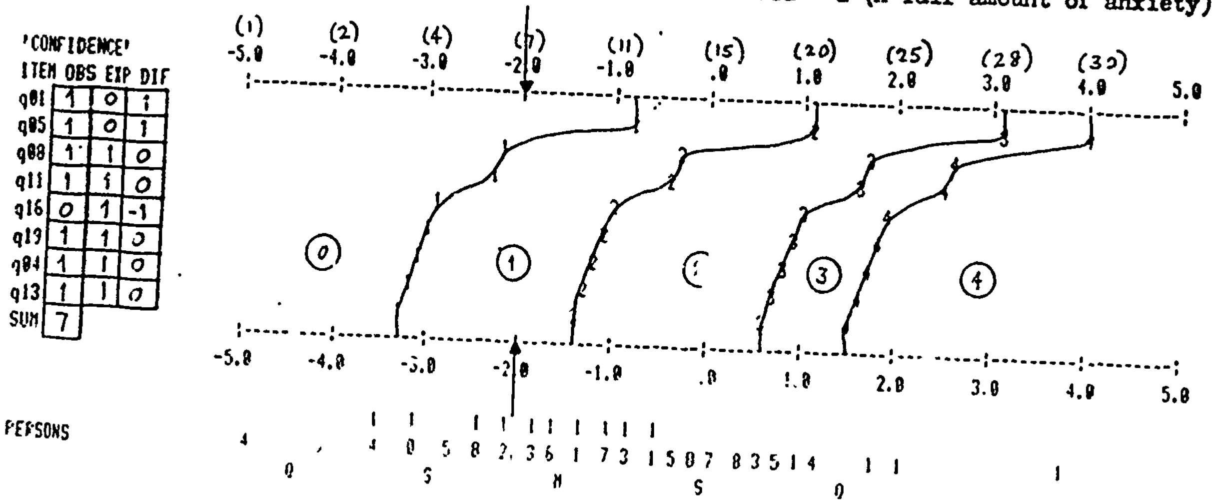


Figure 4.3. An example

3. Read off the equivalent logit score according to the ruler (the total score-logit conversion) at the top of each most probable response curve; this logit score (marked by the arrow) is the pupil's anxiety level on the subscale.
4. Note the band level and interpret this according to Table 4.1; this band level is a gross summary of the pupil's anxiety behaviours and tendencies.
5. Record under EXP the most probable response of each item based on the position of the pupil's logit score on the anxiety continuum, record the difference between OBS and EXP under DIF; this DIF provides additional information on a pupil's deviant response to an item should the difference be a large one (say plus or minus 2).

Overcoming anxiety

Much work needs to be done to really map out the effects of computer programming anxiety. What we can do here is to recognise the existence of anxiety and its potency so that we may develop strategies to overcome it. We do not think that one can "cure" computer programming anxiety by simply writing on how-to-do-it. In fact, the famous Murphy's Law is an expression of this anxiety: "If anything can go wrong, it will!".

Some techniques for reducing anxiety include development of learning skills and discussion of past bad experiences so as to change negative self-images.

Development of learning skills: Computer programming problems are normally illustrated and solved by a combination of convergent and divergent thinking. However, these illustrative problems are usually presented as if they had been solved entirely by convergent thinking, thus giving pupils a deceptive view of how the solutions are arrived at. It is worthwhile therefore for teachers to discuss the process that went into solving a computer programming problem.

The teacher may divide the pupils into cooperative groups to develop their own theories on the computer programming problem and its possible solutions. Pupils who are frustrated at a certain stage in the problem solving process will then have the opportunity to retrace and clarify the process with peers. When the teacher assumes the role of a guide rather than director, pupils may be more willing to take risks and to try different approaches. Hopefully, the process of learning will in fact produce learning.

The computer programming laboratory sessions in Singapore give pupils hands-on, concrete experience with the concepts discussed in lectures. Pupils must be educated to be prepared beforehand in order to get the most out of the laboratory experience. They may be allowed to share workstations in working on a sample set of programmes, studying programme outputs and analysing error messages.

<u>Discussion of past experiences</u>: The teacher was once a learner. It is desirable that teachers discussed their past experiences on computer programming. Any experiences that parallel those of the pupils should be emphasised. Revealing these to pupils makes a very strong impact and can communicate to them that they too can overcome their anxiety.

Pupils must learn to recognise the causes of their anxiety.

Just hearing another pupil say "I feel the same way" is a revelation that takes away some of the fear of having to fight anxiety alone. Pupils should be encouraged to talk about their earlier experiences with computers and how they think these may have affected their attitudes towards learning how to programme. The fact that most pupils are in the same boat is a giant first step towards reducing computer programming anxiety.

In challenging anxieties that are related to how pupils think others are doing in tests and assignments, pupils may transfer their own coping mechanisms for non-computer-related anxiety to computer programming anxiety. Pupils may produce coping statements like "Is everyone really doing better than I do?". At the same time pupils must recognise that every programming task could be accomplished in small, discrete steps and realise that it is every programmer's lament that the computer will never do quite what you want but only what you tell it to do.

Conclusion

Rating Scale Analysis of pupils' responses to the Computer Programming Anxiety Scale instrument provides a method of systematically looking at pupils' anxiety in learning to programme. Detailed outputs normed on junior college computer science pupils provide teachers with a means of identifying pupils with unusual response patterns for further diagnoses.

This study demonstrates the feasibility and potential of using the instrument for both prescriptive and predictive evaluation of pupils' performance in computer programming. It is certainly premature to propose a conceptual and measurement panacea, but like the Chinese journey of a thousand miles, we shall approach it one step at a time. Shall we go? Now?

AUTHORS

Dr K. C. CHEUNG joined the Educational Research Unit, Institute of Education, Singapore in 1989 after obtaining his Ph D from the Centre of Educational Studies, King's College, London. His thesis on science achievement in Hong Kong won him the Bruce H Choppin Memorial Award. He was lecturing in the University of Hong Kong before going on to London on a Commonwealth Academic Staff Scholarship for Science Education. His research interests cover constructivism in learning and teaching, science education and modelling of educational data.

Miss KOH Wee Koon is in the midst of writing her dissertation, "Primary school pupils' understanding of fractions", for the Master's degree in Education. Her several years of experience as a secondary school mathematics teacher convinced her of the need for greater understanding of how pupils learn mathematics. Her interests include understanding how pupils acquire knowledge and meaningful measurement and classroom evaluation.

Dr SOH Kay Cheng is currently Head, Postgraduate Programme and Assistant Head, Educational Research Unit, Institute of Education, Singapore. He obtained his Ph D from the National University of Singapore with his thesis on Chinese-English code-switching of primary school children. His main research interests include attitudinal measurements, teacher personality, language testing and bilingualism.

Miss MOOI Lee Choo has submitted for examination her dissertation, "Factors affecting computer programming ability among junior college students", for the Master's degree in Education and is teaching Computing (Advanced Level) in a junior college. Her main interests are educational computing, modelling, testing and measurement.

References

Behr, M., Lesh, R., Post, T., & Silver, E. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.). The Acquisition of Mathematical Concepts and Processes. New York: Academic Press.

Bergeron, J. C., & Hercovics, N. (1987). Unit fractions of a continuous whole. In J. C. Bergeron, N. Hercovics & C. Kieran (Eds.). Proceedings of the Eleventh International Conference for the Psychology of Mathematics Education, Vol. I, Montreal, 357-365.

Bright, G. W., Behr, M. J. Post, T., & Wachsmuth, I. (1988). Identifying fractions number lines. Journal for Research in Mathematics Education, 19, 215-232.

Brown, R. (1973). A First Language: The Early Stages. Boston, Mass: Harvard University Press.

Booth, L. R. (1987). Primary school children's conceptions of fractions. Proceedings of Fourth South-East Asia Conference on Mathematical Education. Singapore: Institute of Education.

Ellis, R. (1984). Classroom Second Language Development. London: Pergamon Press.

Epstein, S. (1972). The nature of anxiety with emphasis upon its relationship to expectancy. In C. D. Spielberger (Ed.). Anxiety: Current Trends in Theory and Research (Vol. II). New York: Academic Press.

Griffin, P. E. (1985). The use of latent trait models in the calibration of tests of spoken language in large-scale selection-placement programs. In Y. P. Lee, Angela C. Y. Y. Fok, Ray. Lord and G. Low (Eds.). New Directions in Language Testing. London: Pergamon press.

Hart, K. (1981). Fractions. In K. M. Hart (Ed.). Children's Understanding of Mathematics: 11-16 (pp. 66-81). London: John Murray.

Hasemann, K. (1981). On difficulties with fractions. Educational Studies in Mathematics, 12, 71-87.

Hiebert, J., & Tonnessen, L. H. (1978). Development of the fraction concept in two physical contexts: An exploratory investigation. Journal for Research in Mathematics Education, 9(5), 374-378.

Hope, J. A. & Owens, D. T. (1987). An analysis of the difficulty of learning fractions. Focus on Learning Problems in Mathematics, 9(4), 25-40.

Jordan, E. W. & Stroup, D. F. (1982). The benavioural antecedent of computer fear. Journal of Data Education, 22, 7-8.

Kempf, W. (1983). Some theoretical concerns about applying latent trait models in educational testing. In S. B. Anderson and J. S. Helmick (Eds.). On Educational Testing. New York: Jossey-Bass Publishers.

Kerslake, D. (1986). Fractions: Children's Strategies and Errors. Windsor: NFER-NELSN.

Kieren, T. (1976). On the mathematical, cognitive and instructional foundations of rational numbers. In R. Lesh (Ed.). Number and Measurement: Papers from a research workshop (pp. 101-144). Columbus, OH: ERIC/SMEAC.

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematical learning and problem solving. In C. Janvier (Ed.). Problems of Representation in the Teaching and Learning of Mathematics. New Jersey: Lawrence Erlbaum Associates Inc.

Lesh, R., Behr, M., & Post, T. (1987). Rational number relations and proportions. In C. Janvier (Ed.). Problems of Representation in the Teaching and Learning of Mathematics. New Jersey: Lawrence Erlbaum Associates Inc.

Masters, G. N. (1980). "A Rasch Model for Rating Scales." Unpublished doctoral dissertation, University of Chicago

Maurer, M. M. (1983). "Development and Validation of a Measure of Computer Anxiety." Unpublished master's thesis, Iowa State University.

Mooi, L. C. (1990). "Factors Affecting Computer Programming Ability of Junior College Students." Unpublished master's thesis, National University of Singapore.

Nik Pa, N. A. (1989). Research on children's conceptions of fractions. Focus on Learning Problems in Mathematics, 11(3), 3-25.

Novillis, C. F. (1976). An analysis of the fraction concept into a hierarchy of selected subconcepts and the testing of the hierarchical dependence. Journal for Research in Mathematics Education, 7, 131-144.

Peck, D. M., & Jencks, S. M. (1981). Conceptual issues in the teaching and learning of fractions. Journal for Research in Mathematics Education, 12, 339-348.

Phillips, B. N., Martin, R. P. & Meyers, J. (1972). Interventions in relation to anxiety in school. In C. D. Spielberger (Ed.). Anxiety: Current Trends in Theory and Research (Vol. II). New York: Academic Press.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). The Child's Conception of Geometry. New York: Basic Books.

Raub, A. C. (1981). "Correlates of Computer Anxiety in College Students." Unpublished doctoral dissertation, University of Pennsylvania.

Rogoff, B. & Wertsch, J. V. (1984). Children's Learning in the "Zone of Proximal Development." New Directions for Child Development, no.23. San Francisco: Jossey-Bass.

Rohner, D. J. & Simonson, M. R. (1981). "Development of an Index of Computer Anxiety." ERIC Doc. No. 207487.

Skemp, R. R. (1971). The Psychology of Learning Mathematics. Harmondsworth, Middx: Penguin.

Soh, K. C. (1983). A Standardized Chinese Language Test for Primary Three (Research Version). Occasional Paper No. 14. Institute of Education, Singapore.

Soh, K. C. (1984). Taking a Cloze Look Psychometrically. Occasional Paper No.20. Institute of Education, Singapore.

Spielberger, C. D. (1972). Conceptual and methodological issues in anxiety research. In C. D. Spielberger (Ed.). Anxiety: Current Trends in Theory and Research (Vol. II). New York: Academic Press.

Streefland, L. (1982). Some observational results concerning the mental constitution of the concept of fraction. Educational Studies in Mathematics, 9, 51-73.

Truran, J. M. (1987). The teaching of fractions. Proceedings of Fourth South-East Asia Conference on Mathematical Education. Singapore: Institute of Education.

von Glasersfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.). Problems of Representation in the Teaching and Learning of Mathematics. New Jersey: Lawrence Erlbaum Associates Inc.

Wesche, M. B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In Karl C. Diller (Ed.). Individual Differences and Universals in Language Learning Aptitude. New York: Newbury House Publishers.

Wright, B. D., Linacre, J. M., & Schultz, M. (1990). Bigscale, A Rasch Program for Rating Scale Analysis. Chicago, Illinosis.

Wright, B. D. & Masters, G. N. (1982). Rating Scale Analysis: Rasch Measurement. Chicago: Mesa Press.

Wright, B. D., & Stone, M. H. (1980). Best Test Design. Chicago: MESA Press.

Appendix 2.1

The multiple-choice test on fractions

1. Which fraction says $\frac{1}{5}$?

(A) one (B) one-sixth (C) one-fifth

(D) Not given (E) I don't know

---

2. Which fraction says "one-third"?

(A) 4 (B) $\frac{1}{3}$ (C) $\frac{1}{4}$

(D) Not given (E) I don't know

---

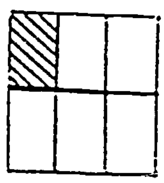
3. Which fraction says "three-fifths"?

(A) $\frac{3}{5}$ (B) $\frac{3}{8}$ (C) 8

(D) Not given (E) I don't know

---

4. What fraction of the square is shaded?

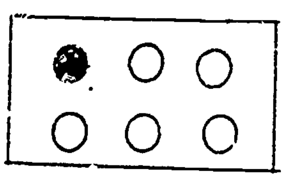


(A) $\frac{1}{6}$ (B) $\frac{1}{5}$ (C) 1

(D) Not given (E) I don't know

---

5. What fraction of the set is shaded?



(A) $\frac{1}{5}$ (B) $\frac{1}{6}$, 1

(D) Not given (E) I don't know

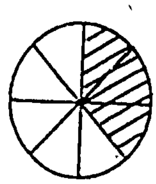
6. What fraction of the set is shaded?

(A) one-third (B) two-thirds (C) one-half

(D) Not given (E) I don't know

---

7. What fraction of the circle is shaded?

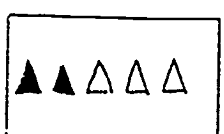
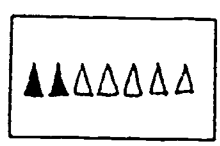


(A) two-fifths (B) three-fifths (C) three-eighths

(D) Not given (E) I don't know

---

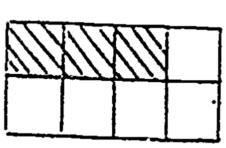
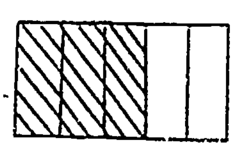
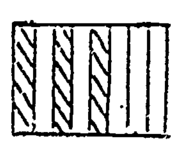
8. Which picture below shows $\frac{2}{5}$ of the set shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

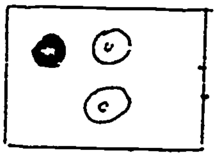
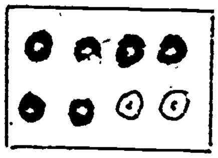
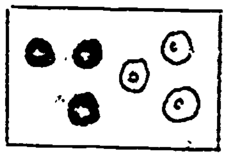
---

9. Which rectangle below shows three-fifths shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

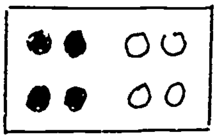
10. Which picture below shows one-half of the set shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

---

11. What fraction of the set is shaded?

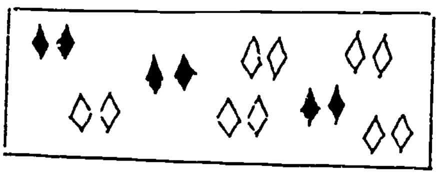


(A) one-half (B) one-fifth (C) one-fourth

(D) Not given (E) I don't know

---

12. What fraction of the set is shaded?

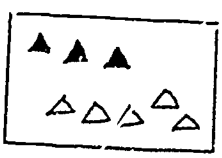
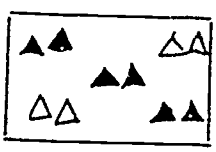
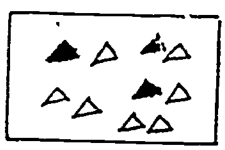


(A) three-fifths (B) three-eighths (C) six-tenths

(D) Not given (E) I don't know

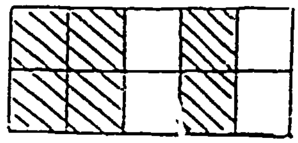
---

13. Which picture below shows $\frac{3}{5}$ of the set shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

14. What fraction of the rectangle is shaded?

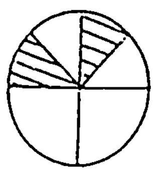


(A) one-quarter (B) three-fifths (C) three-fourths

(D) Not given (E) I don't know

---

15. What fraction of the circle is shaded?

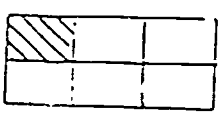
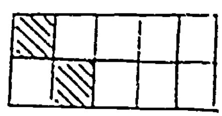
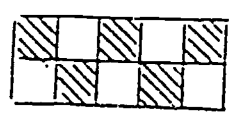


(A) one-quarter (B) two-fourths (C) two-sixths

(D) Not given (E) I don't know

---

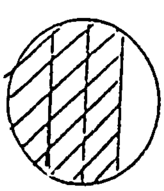
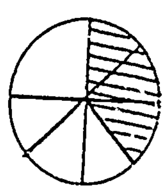
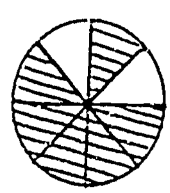
16. Which rectangle below shows $\frac{1}{5}$ shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

---

17. Which circle below shows $\frac{3}{4}$ shaded?

(A)  (B)  (C) 

(D) Not given (E) I don't know

Computer Programming Anxiety Scale

Please respond to each statement by circling the number that best describes how you feel. Thank you.

CODE:	*Never true* 0	*Seldom true* 1	*Sometimes true* 2	*Often true* 3	*Always true* 4

1.	Computers make me feel nervous and uncomfortable.	0	1	2	3	4
2.	Debugging of programmes is a great worry to me.	0	1	2	3	4
3.	I worry about making a fool of myself in front of my friends when I cannot write programmes.	0	1	2	3	4
4.	My mind seems to be confused with so many computer terms.	0	1	2	3	4
5.	I think I would not be able to understand programming.	0	1	2	3	4
6.	I am troubled by the number of bugs in my programme.	0	1	2	3	4
7.	I worry about being picked by my teachers when I make mistakes in programming.	0	1	2	3	4
8.	If my programme cannot run I do not know what to do next.	0	1	2	3	4
9.	I get worried when my programmes cannot run.	0	1	2	3	4
10.	It worries me that my teacher knows I cannot programme well.	0	1	2	3	4
11.	I cannot think properly when there is an error in my programme lines.	0	1	2	3	4
12.	I feel tense when I have to correct my programme lines over and over again.	0	1	2	3	4
13.	As the programme lines become more complicated I feel that I cannot concentrate.	0	1	2	3	4
14.	I feel nervous in the presence of more capable computer science students.	0	1	2	3	4
15.	I feel troubled when my friends could programme and I could not.	0	1	2	3	4
16.	My mind seems to go blank when I cannot solve a computer problem.	0	1	2	3	4
17.	I feel tense whenever there are errors in my computer programme.	0	1	2	3	4
18.	I am troubled when friends discuss programming concepts that I do not understand.	0	1	2	3	4
19.	I do not feel confident whenever I work on the computer.	0	1	2	3	4